

Detection of snow disturbances in boreal forests using unitemporal airborne lidar data and aerial images

Janne Rätty^{1,*}, Mikko Kukkonen², Markus Melin³, Matti Maltamo⁴, Petteri Packalen⁵

¹Bioeconomy and Environment, Forest Inventory and Planning, Natural Resources Institute Finland, Yliopistokatu 6, 80100, Joensuu, Finland

²Bioeconomy and Environment, Forest Inventory and Planning, Natural Resources Institute Finland, Paavo Havaksen tie 3, 90570 Oulu, Finland

³Natural Resources, Forest Health and Biodiversity, Natural Resources Institute Finland, Yliopistokatu 6, 80100, Joensuu, Finland

⁴School of Forest Sciences, University of Eastern Finland, P.O. Box 111, 80101 Joensuu, Finland

⁵Bioeconomy and Environment, Forest Inventory and Planning, Natural Resources Institute Finland, Latokartanonkaari 9, 00790 Helsinki, Finland

*Corresponding author. Email: janne.ratty@luke.fi

Abstract

Snow is among the most significant natural disturbance agents in Finland. In silviculture, maps of snow disturbance are needed to recognize severely disturbed forests where the risk of subsequential disturbances, such as insect outbreaks, is high. We investigated the potential of unitemporal airborne lidar (light detection and ranging) data and aerial images to detect snow disturbance at the tree level. We used 81 healthy and 128 snow-disturbed field plots established in a 63 800 ha study area in Eastern Finland. A subset of trees ($n = 675$) was accurately positioned in the field plots. We carried out individual tree detection (ITD) using airborne lidar data (5 p/m^2), and a random forest classifier was used to classify healthy and broken trees. Tree features were extracted from a terrain elevation model, lidar data, and aerial imagery. We compared canopy height model-based (ITD_{CHM}) and point cloud-based (ITD_{PC}) ITD approaches. We explored random forest variable importance scores and evaluated the classification performance by an F1-score and its components (precision and recall). Performance was also evaluated at the plot level to investigate errors associated with the predicted number of broken trees. We achieved F1-scores of 0.66 and 0.85 for the tree- and plot-level classifications, respectively. The variable importance scores showed that elevation above sea level was the most important predictor variable followed by ITD-based features characterizing the neighborhood of trees. The ITD_{CHM} slightly outperformed the ITD_{PC} at the tree level, while they both underestimated the number of broken trees at the plot level. The proposed approach can be carried out alongside lidar-assisted operational forest management inventories provided that a set of positioned broken and healthy trees are available for model training. Since airborne lidar data often have a temporal resolution of several years for the same areas, future research should consider the utilization of other remotely sensed data sources to improve the temporal resolution.

Keywords: forest damage; airborne laser scanning; natural forest disturbance; snow disturbance; snow damage

Introduction

Forest disturbances caused by snow are an inherent part of forest dynamics in northern Europe. Snow disturbance is induced by the accumulation of a snow load that typically results in crown bending or breakage of the tree. Snow load may also uproot trees, although this is not common in northern Europe due to the frozen soil in winter (Nykänen et al. 1997). While snow disturbance rarely damages all trees in a forest, the negative effects of snow breakages can be considerable at the level of the forest stand. Snow breakages reduce the quality of the commercial part of the stem, which causes immediate economic losses. In addition, the broken trees also have a limited ability to maintain photosynthesis, which reduces their biomass growth in the future. It is worth noting that snow disturbance can expose forests to subsequential disturbances caused by insect and fungal attacks (Juutinen 1953, Schroeder and Eidmann 1993). From the perspective of forest dynamics, snow disturbance contributes to the creation of deadwood, which is crucial for the maintenance of boreal forest biodiversity (Siitonen 2001).

Snow is the most harmful disturbance agent associated with Finnish timber production when the degree of harmfulness is assessed by area (Korhonen et al. 2021). According to the Finnish National Forest Inventory (NFI), the estimated areal extent of snow disturbance in timber production forests in 2021 was 1.5 million ha (Aarnio et al. 2022). The likelihood of snow disturbance varies considerably across geographical areas due to climatic factors, local topography, and the stand characteristics of the forest (Nykänen et al. 1997). This means that some geographical areas are inherently more vulnerable to snow disturbance than others. Previous studies have also found a relationship between snow disturbance and terrain elevation above sea level (asl). For example, Suominen (1961) reported an increased likelihood of snow disturbance events for forests in southern Finland located higher than 90 m above asl, while Heikinheimo (1920) noted that, overall, snow disturbance mainly occurred at altitudes above 300 m asl. The weather conditions that are linked to the occurrence of snow disturbance have been reported to become more common above altitudes of ca. 200 m asl (Ahti 1978).

Handling editor: Dr. Ewa Grabska-Szwagrzyk

Received 1 July 2024; revised 25 October 2024; accepted 5 November 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Institute of Chartered Foresters.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

While snow-induced disturbances cannot be fully prevented, timely forest treatments, such as thinnings, can mitigate the vulnerability of forests to snow-induced stem breakages (Päätaalo 2000). On the contrary, forest treatments, such as fertilization, have been found to increase the risk of snow disturbance (Valinger et al. 1993). Moreover, climate change has already led to increased winter temperatures in Finland and is expected to further raise temperatures in the next decades, which will increase the likelihood of heavy snow accumulation in the northern and eastern parts of the country that have typically experienced cold winters and thick snow cover (Lehtonen et al. 2016).

Traditionally, snow disturbance has been monitored in sample-based NFIs, and several studies have used NFI data to foster an understanding of the dynamics of snow disturbance under different growing conditions and stand characteristics (Valinger and Fridman 1999, Jalkanen and Mattila 2000, Díaz et al. 2019, Suvanto et al. 2021). The snow disturbance events recorded in NFI plots can be linked to existing forest resource information, climate, and environmental data to model the likelihood of snow disturbance with different stand characteristics. Such risk models promote an understanding of the dynamics of snow disturbance, although the models cannot be used to map snow disturbance that has occurred. Such maps are, however, needed to implement timely actions for the mitigation of subsequential disturbances that impede future forest growth and vitality. In Finland, the timely detection of large-scale snow disturbance is also important from an operational point of view. This is because Finnish forest legislation directed to prevent forest damage states that freshly damaged coniferous wood must be removed from the forest due to the potential for subsequential insect damages (“Forest Damages Prevention Act 1087/2013” 2013).

Remotely sensed (RS) data are potential auxiliary data sources that can be used to map snow disturbance because the anomaly in forest structure created by snow damage is relatively intense (Fig. 1). Figure 1 shows typical examples of snow disturbances in boreal, coniferous forests. In boreal forests, the snow breakages are distinctive and easily distinguishable by the naked eye from the other forest disturbances, such as defoliation caused by insects or wind damages, which supports the use of RS data for the detection of snow disturbances. Previous studies have demonstrated that multitemporal RS data enable the detection of canopy changes with good accuracy (Vastaranta et al. 2012, Tomppo et al. 2019). In particular, the use of bitemporal airborne lidar (light detection and ranging) is a promising data source for the detection of snow disturbance (Vastaranta et al. 2012). However, the acquisition of airborne lidar data at several time points within a short time interval is often not feasible from an operational point of view. Instead, RS data collected by space-borne instruments provide greater temporal resolution at low cost, although the spatial resolution of such datasets is typically so poor that the detection of canopy changes associated with individual trees is out of the question. Tomppo et al. (2019) mapped snow disturbance using stand-level snow disturbance observations and multitemporal imagery collected by space-borne synthetic aperture radar (SAR) and achieved a detection accuracy of 90% associated with the stand-level snow disturbance. Nevertheless, they concluded that further research was needed to support their results because their reference data did not comprise actual field measurements to reliably verify, for example, the severity of the snow disturbance. Puliti and Astrup (2022) proposed a convolutional neural network-based framework for the detection of snow disturbance at the tree level by means of unitemporal aerial imagery collected by an unoccupied aerial vehicle (UAV)

operated at a flight altitude of ~110 m above ground level. Their model achieved promising results in terms of classification accuracy, but it is worth noting that the detection of snow disturbances with the current UAV technology is best suitable for inventories that are targeted at specific small areas because of the high time demand and inventory costs per hectare in large-area mapping applications. Zubkov et al. (2024) also proposed a tree-level approach for snow disturbance mapping, but their approach relied on mechanistic snow disturbance risk modeling. They extracted tree heights and crown radii from airborne lidar data, but otherwise, their modeling framework was largely built on snow accumulation and critical snow load models.

Research associated with individual tree detection (ITD) and tree-level inventories has been intensive during the era of lidar-based forest inventories, although ITD has not matured into an operational standard mostly due to challenges in the detection of suppressed trees and the application of tree allometry for the prediction of diameter by tree height (Hyyppä and Inkinen 1999, Vauhkonen and Mehtätalo 2015). The most common way to detect trees from lidar data is to rely on canopy height models (CHMs) and image processing techniques. In general, the CHM-based workflow has received criticism concerning the sensitivity to the parameters, such as smoothing and interpolation, used in the generation of a CHM. An alternative to the CHM-based approach is to segment trees directly from a 3D point cloud (Li et al. 2012). Such an approach avoids the loss of information due to CHM processing, which ensures that the ITD algorithm can also utilize the details that would potentially be smoothed out during the generation of the CHM. The performance differences associated with the ITD approaches are not known in snow-disturbed forests where the crown shapes of broken trees can be very different compared to healthy trees.

In this study, we examined the detection of snow disturbance at the level of individual trees using airborne lidar data (5 p/m²) and aerial images. The performance assessment was also carried out at the level of field plots by aggregating the predicted class labels of detected individual trees. The objectives of this study were:

- (1) To compare the classification accuracy of healthy and broken trees detected by CHM-based and point cloud-based ITD algorithms.
- (2) To analyze the features of lidar data and aerial images that are useful in the classification of healthy and broken trees.
- (3) To investigate how the proposed approach performs for the detection of snow disturbance at the level of field plots and thereby gain insights into how the developed system could aid operational forestry.

Materials and methods

Study area

The 63 800 ha study area (63.2°N 28.9°E) was located in Juuka, eastern Finland. The study area is a subset of an operational inventory block that was used in the RS-based forest inventory organized by the Finnish Forest Centre. The study area is representative of typical Finnish forests that are used for multiple purposes in addition to timber production. For example, there are conservation and recreation areas in the study area. Coniferous tree species, such as Scots pine (*Pinus sylvestris* [L.]) and Norway spruce (*Picea abies* [L.] Karst.), are dominant in the forests, whereas broadleaf species, such as birch (*Betula* spp.), are commonly found as admixtures. The terrain elevation varies between 97 and 319 m asl.

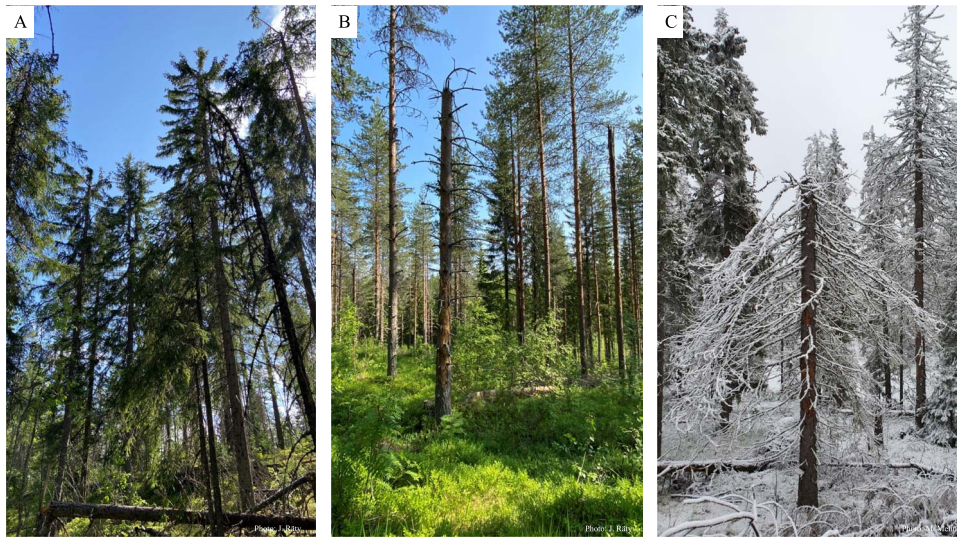


Figure 1. Examples of snow disturbance in boreal, coniferous forests: A) top breakages in a spruce forest, B) stem breakages in a pine-dominated forest, and C) stem breakage in a spruce-dominated forest.

Snow disturbance plots

In total, 128 field plots were measured during the summers of 2022 and 2023 in coniferous-dominated forests where snow disturbance had occurred. The snow-disturbed forests were middle-aged or mature forest stands with a typical single-layered forest structure. Snow disturbances cause distinctive changes in boreal canopy structures, which enables the separation of tree damages caused by snow and other disturbance agents in the field. The plots were located at 189–307 m asl. The exact years that the snow disturbance occurred were not possible to define, but most of the disturbance events occurred during the heavy snow winters of 2018 and 2021. The locations of the snow disturbance plots were subjectively selected during field visits so that the set of plots represented a variety of disturbance magnitudes from plots with minor damage to plots with 68% damaged trees.

The snow disturbance plots were circular plots with a fixed radius of 10 m. The centers of the plots were accurately positioned using a dGNSS (differential global navigation satellite system) device (Trimble R2). The plot positions were postcorrected using reference positions provided by the closest Trimble virtual reference station. The measurement protocol comprised a total enumeration of trees with a diameter at breast height (DBH) ≥ 50 mm. The measurements comprised recording of DBH, height (H), and tree species. All trees were categorized into healthy or broken categories. Total tree volumes were computed using the species-specific volume equations that employ DBH and H as predictor variables (Laasasenaho 1982). For broken trees, total H was predicted using the Näslund curve (Näslund 1936) where parameters were estimated by tree species using the nonlinear least square estimation in the R environment (R Core Team 2024) and by H and DBH measurements of healthy trees.

In each plot, up to five trees broken by snow were accurately positioned with similar positioning equipment as used for the centers of the field plots. In addition to the broken trees, one to two healthy trees were positioned in each plot during the field season of 2023. Our aim was to take into account the fact that snow causes scattered disturbances, which means that there are also healthy trees in the snow-disturbed plots of the modeling data. In total, 490 broken and 185 healthy trees were accurately positioned in snow disturbance plots.

Healthy plots

Data from the healthy plots were collected in two field campaigns. First, a set of 131 field plots without snow-induced stem breakages were taken from a database of field plots acquired for the operational forest management inventories operated by the Finnish Forest Centre (henceforth, these plots are referred to as FFC plots). The field plots were measured during the summer of 2022, and they were circular plots with either a radius of 9.0 or 12.62 m. The larger radius was used when a plot with a radius of 9 m comprised < 20 trees. Field measurements of healthy plots did not differ from the measurements of the snow disturbance plots. Since the original set of FFC healthy plots comprised a variety of forest structures (in terms of species composition and development class), we filtered out field plots that were dissimilar compared to the forest attribute distribution of the snow disturbance plots. Homogenization between FFC and snow disturbance data was needed to avoid overly optimistic results in the detection of snow disturbances. Homogenization of field datasets was performed using a principal component analysis (PCA) in the R environment. The space of plot-level forest attributes was reduced into two dimensions using the first two principal components. The plot-level attributes were number of stems, dominant height, basal area, the proportion of deciduous volume to total volume, and species-specific (pine, spruce, broadleaf) volumes. For the snow disturbance plots, volume and dominant heights were representative of the status prior to the disturbance event. The space of the first two principal components associated with the snow disturbance plots was delineated using the concave hull algorithm implemented in the *concaveman* R package (Gombin et al. 2020). The same dimensionality reduction was carried out for the FFC healthy field plots, and they were projected into the same space as the snow disturbance plots. The FFC field plots that did not fall into the delineated space of snow disturbance plots were omitted. In total, 60 FFC field plots were used in the training data.

Next, a set of 21 healthy field plots were measured in 2023 to supplement the FFC plots. The healthy plots were measured from snow-disturbed forest stands in a similar manner to snow disturbance plots, with the exception that the location of the field plot was selected so that it contained no trees broken by snow. The measurement of healthy plots near the snow disturbance

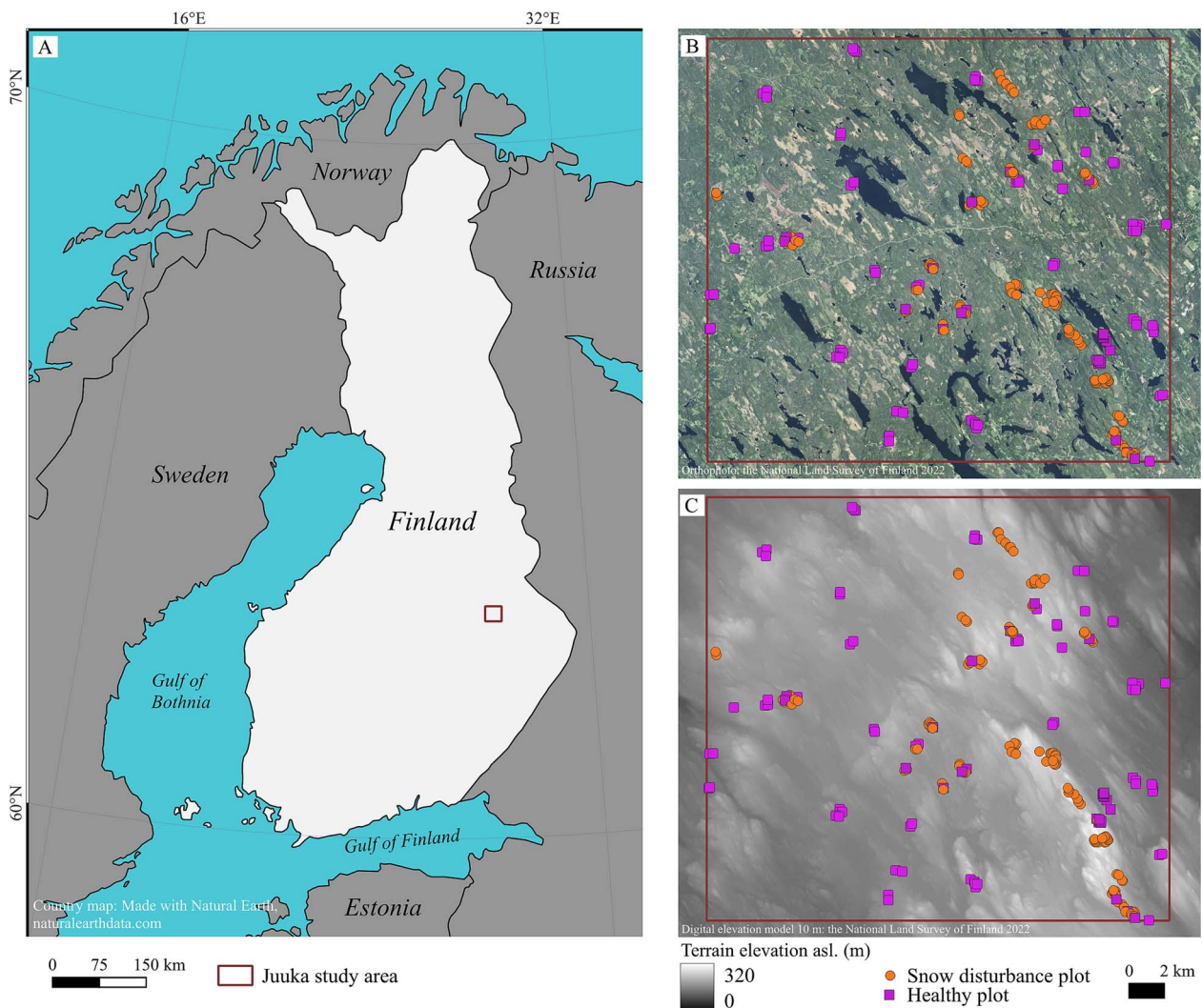


Figure 2. (A) Location of the study area in eastern Finland, (B) locations of the field plots in the study area, and (C) locations of the field plots overlaid on the terrain elevation model. asl, above sea level.

plots was motivated by the disposition of the FFC plots that favored productive low-elevation forests (Fig. 2). The main purpose of these supplementary healthy plots was to ensure that the reported results were not overly optimistic due to the dissimilar disposition of healthy and snow disturbance plots in terms of terrain elevation.

The total number of healthy plots was 81 and comprised 2832 trees. No individual trees were positioned with the dGNSS device from the healthy plots as the absence of snow-disturbed trees from these plots was ensured in the field, and all trees were assumed to be healthy. It is worth noting that the ultimate number of trees used in the modeling depended on the ITD algorithm. The positions of the healthy plots and snow disturbance plots are shown in Fig. 2, and the associated forest attributes are presented in Table 1.

Remotely sensed data

We used airborne lidar data and multispectral aerial images. Both airborne lidar data and aerial images belong to the national data acquisition programs organized by the National Land Survey of Finland. The main purpose of the nationwide acquisitions is to provide remotely sensed data for forest management inventories and terrain mapping.

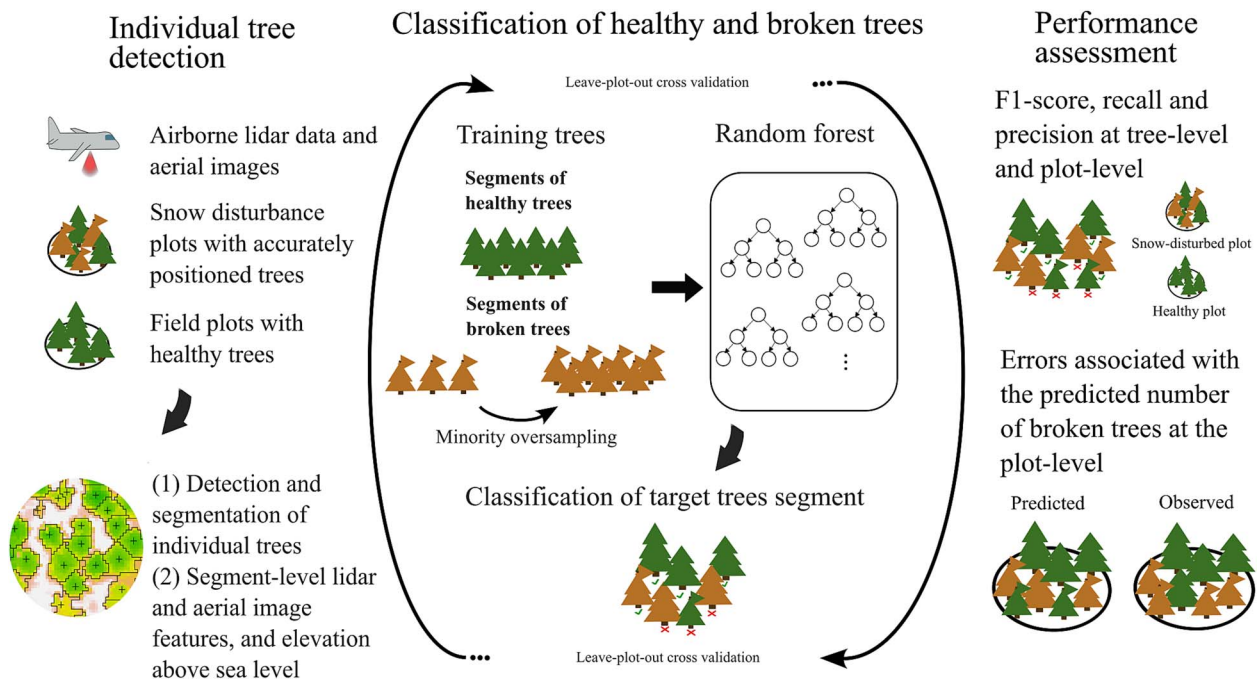
The airborne lidar data were collected between 5 and 12 August 2022 from a fixed-wing airplane equipped with a Riegl VQ1560II-S sensor operating in the near-infrared wavelength. The airplane flew at an altitude of 2100 m with a flight speed of 287 km/h. The lateral overlap between adjacent scanning lines was 20%. The Riegl VQ1560II-S sensor was used with a half angle of 20°, a beam divergence of 0.25 ($1/e^2$), and a pulse repetition frequency of 1620 kHz. The flight and scanning parameters resulted in an average density of 5.5 submitted pulses/m². The echoes were categorized into *last-of-many*, *intermediate*, *first-of-many*, and *only* echoes. The echoes were classified as ground and other returns (Axelsson 2000), and the ground returns were used to interpolate a digital terrain model using a Delaunay triangulation. The lidar echoes were height-normalized by subtracting the digital terrain model from the original echo heights.

The aerial images were captured at a nominal flight altitude of 7680 m above ground level using an UltraCam Eagle M3 digital camera. The flight campaign was carried out on 24 June 2022. The lateral and forward overlap were 30% and 80%, respectively. The digital camera had a focal length of 100.5 mm and was equipped with a 5668 × 8820-pixel sensor with a pixel size of 12 μm for color images (red, green, blue, and near-infrared). This configuration resulted in a ground sampling distance of ~92 cm. External

Table 1. Statistics associated with the healthy ($n = 81$) and snow disturbance plots ($n = 128$).

	Mean		Minimum		Maximum		Standard deviation	
	Healthy	Disturbed	Healthy	Disturbed	Healthy	Disturbed	Healthy	Disturbed
V_{theo} (m ³ /ha)		190.1		64.8		446.2		70.1
V_{cur} (m ³ /ha)	171.8	184.0	60.8	56.5	474.8	445.8	74.1	70.2
D (cm)	19.2	21.9	11.2	12.8	29.6	35.4	4.6	3.6
N (stems/ha)	1246	981	287	350	3504	2419	712	455
P_{broken} (%)	0.0	28.2	0.0	2.0	0.0	68.2	0.0	15.1
P_{decv} (%)	4.9	0.0	0.0	0.0	38.6	0.3	8.1	0.1
Elev_{asl} (m)	192.4	234.0	134.6	188.9	258.7	306.7	28.3	30.0

V_{theo} , theoretical timber volume for broken trees, i.e. timber volume prior to disturbance; V_{cur} , current timber volume; D , mean diameter at breast height weighted by basal area; N , number of stems; P_{broken} , proportion of trees broken by snow; P_{decv} , proportion of deciduous timber volume in terms of total volume; Elev_{asl} , terrain elevation above sea level.


Figure 3. Workflow of the study.

orientations were determined by a bundle block adjustment using ground and tie points. The optical information of the nonrectified images was attached to the airborne lidar data by linking the digital number (DN) values of the different bands to the first echoes of lidar data using collinearity equations (Packalén et al. 2009). The average of the DN values was used when several pixels were attached to a single lidar echo.

Workflow

The methodological workflow of this study is shown in Fig. 3. The subsequent sections describe the methodological workflow steps in detail: ITD (section [Individual Tree Detection](#)), classification of healthy and broken trees (section [Classification of Healthy and Broken Trees](#)), and the performance assessment (section [Performance Assessment](#)).

Individual tree detection

Individual trees were detected and segmented using CHM-based and PC-based algorithms. The field-positioned trees were linked with the ITD trees by searching for the nearest treetop using the 3D Euclidean distance. The ITD approach was carried out for each

field plot using a 60 m tile because trees beyond the extent of field plots were needed for the calculation of neighborhood metrics.

The CHM-based ITD (henceforth ITD_{CHM}) approach follows the principles proposed by Hyypä and Inkinen (1999) and Pitkänen et al. (2004). ITD_{CHM} consists of four steps: (i) generation of the CHM, (ii) low-pass filtering of the CHM, (iii) detection of local maxima, and (iv) segmentation of the trees. The CHM was generated with a cell size of 0.5 m using the normalized echo heights within the cells. Low-pass filtering was implemented by a Gaussian kernel in which a sigma parameter defines the magnitude of smoothness. Here, we show results using sigma values of 0.4 and 0.7, which are referred to as $\text{ITD}_{\text{CHM}4}$ and $\text{ITD}_{\text{CHM}7}$, respectively. The subsequent step was to detect local maxima in the low-pass-filtered CHMs and segment the detected trees using the marker-controlled watershed algorithm (Narendra and Goldberg 1980). Unrealistically small segments with an area $< 1 \text{ m}^2$ were removed. The implementation of ITD_{CHM} was based on in-house algorithms that exploited the Geospatial Data Abstraction Library (GDAL) library (GDAL/OGR contributors 2024).

The ITD_{PC} detects and delineates individual trees directly from the point cloud using the algorithm proposed by Li et al. (2012) through the lidar package (Roussel et al. 2020, Roussel and Auty

Table 2. Features extracted from remotely sensed data for tree segments.

Data	Feature	Description
Lidar data ^a	hmin, hmax	Maximum and minimum of echo heights (m).
	hmean, hmed, hstd	Mean, median, and standard deviation of echo heights (m).
	h5, h10, ..., h95	Percentiles (5%, 10%..., 95%) associated with the distribution of echo heights (m).
	d0.5, d2, d5, d10, d15	Proportion of echo heights below a fixed threshold of 0.5, 2, 5, 10, or 15 m (%).
	fprop	Proportion of first-of-many and only echoes to all echoes (%).
	hratio _{hmax/hP}	Ratio of hmax and height percentile. P refers to height percentiles (h5, h10, ..., h95) (%).
	intmean, intmed, intstd	Mean, median, and standard deviation of echo intensity recordings.
	int5, int10, ..., int95	Percentiles (5%, 10%..., 95%) associated with the distribution of echo intensity recordings.
	nb _{itdn}	Number of detected trees
	nb _{itdgr}	Proportion of detected trees taller than a target tree in a 10 m-radius neighborhood.
	se _{garea}	Area of tree segment (m ²).
	nb _{segarea}	Mean area of tree segments in the 10 m-radius neighborhood (m ²).
	nb _{hnm}	Height of target tree in relation to the mean of detected tree heights in the 10 m-radius neighborhood (%).
Aerial images ^b	nb _{hstd}	Standard deviation of detected tree heights in the 10 m-radius neighborhood (m ²).
	aimin, aimax, aistd, aimean	Minimum, maximum, standard deviation, and mean of digital number values linked to first lidar echoes from aerial images. Computed by bands red, green, blue, and near-infrared, e.g. Rmin, Gmax, and NIRmean.
	airatio _{X/Y}	Ratio of mean digital number values from X and Y bands. X and Y refer to the band prefixes ^b .
Terrain elevation model	elev _{asl}	Terrain elevation above sea level.

Lidar, light detection and ranging. ^aPrefixes: *f* for first echoes, and *a* for all echoes. For example: *f_hmean*. ^bBand prefixes: R—red, G—green, B—blue, and NIR—near-infrared.

2022) in the R environment. The ITD_{PC} algorithm is controlled by three thresholds (*dt1*, *dt2*, *Zu*) that determine how the relative spacing between trees is considered in the assignment of lidar echoes to clusters. We used the default parameter settings provided in the *lidR* package. Since the ITD_{PC} relies on the point cloud and on the clustering of lidar echoes into individual trees, the clusters of tree-level points were transferred into 2D crown segments by means of the concave hull algorithm implemented in the *concaveman* R package (Gombin et al. 2020). Unrealistically small segments with an area <1 m² were removed.

Feature extraction for the ITD trees involved statistics associated with H and intensity measurements of lidar echoes, aerial image features, neighborhood characteristics derived from ITD, and terrain elevation. The extracted tree-level features are listed and described in Table 2. The lidar features were used without a height cutoff, and the features were computed for *first* and *all* echoes. The *first* echo category comprises both *first-of-many* and *only* echoes. Terrain elevation *asl* was computed for each tree based on a national 2 m resolution terrain elevation model (National Land Survey of Finland 2023).

Classification of healthy and broken trees

The random forest (RF) classifier is an ensemble machine learning method that relies on a group of decision trees to predict class labels for the target observations based on the training data (Breiman 2001). The decision trees of the forest are grown using random subsets of the predictor variables and bootstrapped samples of the training data. The RF classifier has an inherent capability to provide *importance* scores for predictor variables included in the training data. We used the RF implementation of the *ranger* package (Wright and Ziegler 2017) in the R environment. The number of decision trees was set at 500, whereas the hyperparameters associated with the number of predictor variables to be randomly selected at each split (*mtry*), the minimum size of terminal nodes (*min.node.size*) and the proportion of observations to sample (*sample.fraction*) were optimized with the *tuneRanger* R package (Probst et al. 2019) using out-of-bag predictions and the F1-score (see Performance assessment) as a performance measure. The split rule associated with the decision trees was Gini impurity, and the

variable importance scores were computed using the permutation technique as implemented in the *ranger* package.

The number of healthy trees was greater than the number of broken trees in our field data, which poses an imbalance between the two classes to be classified. The class imbalance is a potential source of systematic errors with the RF classifier, which relies on bootstrap sampling. We used the synthetic minority over-sampling technique (SMOTE) (Chawla et al. 2002) to account for the class imbalance and augmented our training data by the generation of new synthetic observations for the minority class, i.e. broken trees based on real observations. The SMOTE algorithm identifies *k*-nearest neighbors for real observations in the feature space and generates new synthetic observations based on the information retrieved from the neighborhood. The generation of synthetic observations using SMOTE is controlled by the *k*-value and by the percentage of minority increment. The *k*-value was optimized so that it maximized the F1-score in the training data. The percentage of minority was iteratively increased from 1 until the approximate class balance was achieved. We used the SMOTE implementation of the *performanceEstimation* package (Torgo 2014) in the R environment. For cross-validation, the training data of each cross-validation iteration were balanced using SMOTE, prior to the fitting of the RF classifier.

Performance assessment

The tree-level classifications were carried out by following the principle of leave-plot-out cross-validation. Since some of the field plots were very close to each other, the cross-validation was set to exclude training plots located closer than 200 m to the target plot. The accuracy associated with the classification of the healthy and broken trees was quantified by the F1-score [equation (1)] and by the precision and recall of its components.

$$F1 - score = \frac{(2 * recall * precision)}{recall + precision} \quad (1)$$

in which $recall = \frac{TP}{TP+FN}$, and $precision = \frac{TP}{TP+FP}$ with TP, FP, and FN denote true positives, false positives, and false negatives, respectively.

The performance was also evaluated at the level of field plots, which means that the predicted classes of ITD trees were aggregated to the level of field plots. Recall and precision were also investigated in terms of a classification threshold that determines the proportion of ITD trees in a plot that must have a predicted broken tree label to classify the plot as a snow-disturbed plot. At the plot level, the prediction performance associated with the number of broken trees was evaluated using the root-mean-squared error [RMSE, equation (2)] and mean difference [MD, equation (3)]. The %RMSE and %MD values were computed by dividing the absolute values by the observed mean.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x}_i)^2}{N}} \quad (2)$$

$$\text{MD} = \frac{\sum_{i=1}^N (x_i - \hat{x}_i)}{N} \quad (3)$$

Results

Classification of broken and healthy trees

The greatest classification accuracy was achieved with the trees detected by ITD_{CHM7} (Fig. 4). The recall scores were similar among all the compared ITD alternatives when all metrics were used as predictor variables in the RF classifier. However, the precision score associated with ITD_{CHM4} was smaller, i.e. the accuracy associated with the predicted broken class was lower, compared to ITD_{PC} and ITD_{CHM7}. The ITD_{CHM7} was also the best-performing approach when terrain elevation was not used as a predictor variable. The exclusion of terrain elevation from the classification associated with the trees of ITD_{CHM7} decreased the F1-score by 8%, which indicated that terrain elevation was an important factor to consider in the prediction of snow disturbance.

The five most important variables associated with the RF classifiers are shown in Fig. 5. Terrain elevation was the most important variable, followed by the neighborhood features. The most important neighborhood features were the number of ITD trees (nb_{itdn}) and the height of the target tree in relation to the neighbor trees (nb_{hm}). In importance order, terrain elevation and neighborhood features were followed by intensity features (f_{int75}) and ratio features computed using the bands of aerial images.

Error rates associated with the number of broken trees at the level of field plots

While the F1-scores associated with the tree-level classifications suggested satisfactory accuracy at the tree level, the predictive performance at the plot level was more relevant in the operational application. The smallest error rates associated with the predicted number of broken trees were achieved using ITD_{CHM4} with all features (Table 3). The MD values were always positive, which indicated that the number of broken trees was systematically underpredicted regardless of the ITD setup. As with the tree-level results, the error rates associated with the detection of snow disturbances also increased when terrain elevation was not used as a predictor variable. For the sake of clarity and because it produced the smallest error rates, the remainder of the Results section will focus on the RF classifier trained with the trees of ITD_{CHM4}. The detection of broken and healthy trees in a selected plot is illustrated in Fig. 6.

Classification of healthy and snow disturbance field plots

There was a trade-off between precision and recall in the plot-level classification, which is demonstrated in terms of

a classification threshold value in Fig. 7. The classification threshold value refers to the plot-level proportion of ITD trees with predicted class “broken” required to classify a plot as a snow disturbance plot. The smallest threshold value of the plot-level classification produced large recall values, which also resulted in a greater number of false positives than larger threshold values. The finding showed that precision and recall curves intersected when the threshold value of plot-level classification was set at 11%. This implies that 11% of detected ITD trees must be classified as broken in order to label the plot as a snow disturbance plot. The precision and recall values associated with the classification of field plots were 0.85 at the intersect of the curves, which implied that the F1-score was also 0.85. Figure 8 shows the relationship between the severity of observed snow disturbances and the recall scores associated with the disturbance detection. Results show that 93% of snow disturbance plots with at minimum every fifth tree broken were identified as snow disturbance plots, whereas the corresponding figure associated with the snow disturbance plots with at minimum 5% broken trees was 80%.

Discussion

Boreal snow disturbances rarely damage all the trees in a forest stand, which motivated us to use a tree-level approach instead of an area-level approach for the detection of snow disturbances. It is not realistic to assume that minor snow disturbances, such as a few broken tops, could be distinguishable based on the RS features computed for grid cells or plots. The well-documented drawback of the ITD approach is the deficient detection rate of suppressed trees in boreal forest conditions because suppressed trees grow below or inside the crowns of dominant trees, thereby rendering them invisible to sky-borne sensors (Räty et al. 2020, Kostensalo et al. 2023). In the case of snow disturbance, poor detection of suppressed trees is not a major issue because the snow breakages that occur in the dominant canopy layer are typically of major interest in managed forests. Another aspect to note is that the ITD algorithms are developed for the detection of trees with a treetop, whereas trees broken by snow do not often have a treetop at all. For this reason, we compared the ITD_{PC} and ITC_{CHM} approaches for the detection of snow-damaged trees because the processing of CHMs may remove relevant information associated with the tree crowns to be delineated (Li et al. 2012). However, our results did not indicate strong superiority of either ITD approach. This is likely due to the fact that, regardless of the ITD algorithm, trees with overlapping crowns are very difficult to identify correctly prior to the actual classification of the disturbance condition of the tree. We expect that the increased point density of the lidar data will improve the detection of trees that have completely lost their crowns (Fig. 1B).

Terrain elevation was the most important predictor variable in the classification of healthy and broken trees, which is in line with earlier studies (Heikinheimo 1920, Suvanto et al. 2021). It should be noted that the terrain elevation is an area-specific predictor variable that limits the transferability of the classification model among mapping domains in large area application. Therefore, it is essential to have a geographically comprehensive training data in the target domain to be mapped. The features that characterize the neighborhood of the trees showed the strongest predictive power among the features extracted from the RS data. The relationship between the neighborhood features and the probability of snow disturbance was reported by Vastaranta et al. (2012). We also noticed this in the field as the snow-disturbed trees were often located on the edge of a canopy gap. Indeed, the

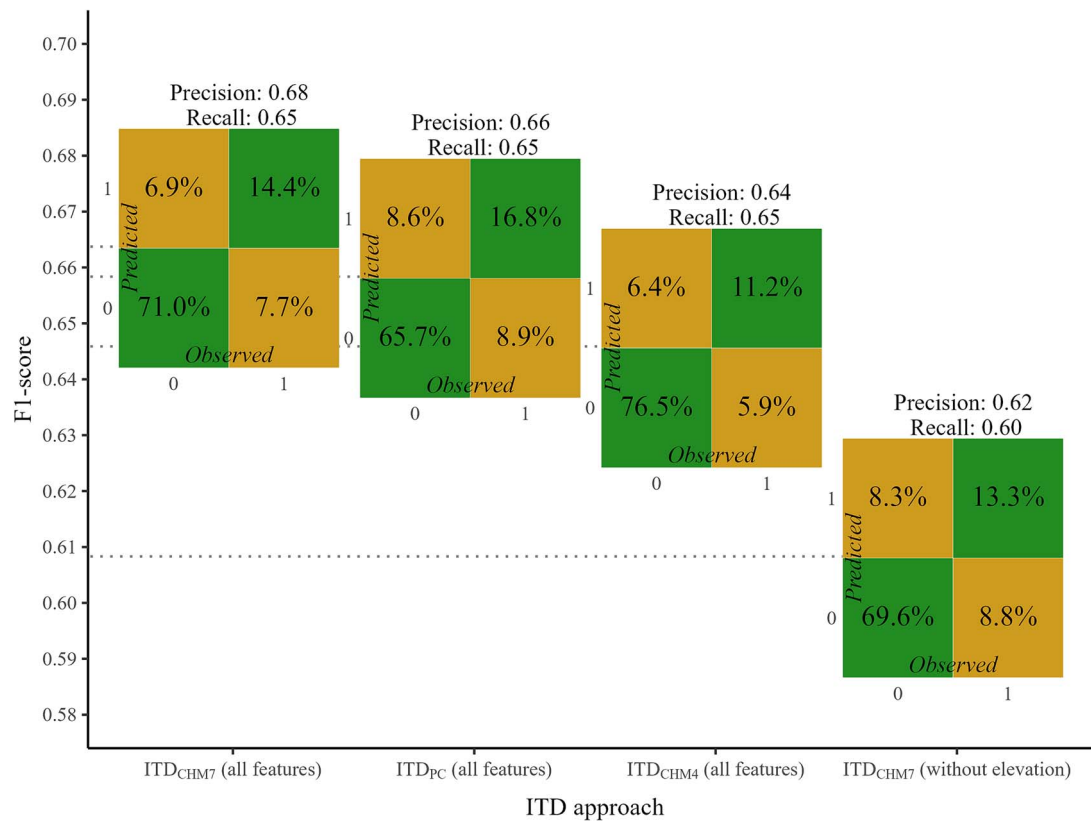


Figure 4. Confusion matrices, F1-scores (dotted line), precision, and recall associated with the tree-level classification of healthy and broken trees by different ITD setups. Note that the number of trees is presented as a proportion because the resulting total number of trees is dependent on the ITD setup. 1—tree broken by snow, 0—healthy tree.

Table 3. RMSE and MD values associated with the predicted number of trees in the field plots that were broken by snow.

	Individual tree detection (ITD)	RMSE (stems/ha)	%RMSE	MD (stems/ha)	%MD
All features	ITD _{CHM4}	167.3	101.7	25.0	15.2
	ITD _{CHM7}	190.8	116.0	56.8	34.5
	ITD _{PC}	191.7	116.6	57.6	35.0
Without terrain elevation	ITD _{CHM4}	200.8	122.1	61.0	37.1
	ITD _{CHM7}	201.1	122.3	61.8	37.6
	ITD _{PC}	206.3	125.5	63.7	38.7

presence of canopy gaps promotes the direct accumulation of snow on individual trees. In the managed forests of the study area, typical canopy gaps were forwarder tracks made during thinning operations. It is also worth noting that neighborhood features, such as the number of detected trees (nb_{itdn}), are not as sensitive to the artefacts associated with the segmentation of ITD trees as the features computed based on lidar echoes linked to the ITD segments. This may explain why structural differences between healthy and broken trees cannot always be captured with the segment-level features extracted from the distribution of lidar echo heights or aerial images.

The predictive power of aerial image and lidar intensity features remained weaker than that of the neighborhood and elevation features. The variable importance scores indicated that aerial image ratio features including RGB and NIR (near-infrared) bands and the f_{int75} intensity feature explained snow disturbance to some extent. The RGB features alone were not important. This is presumably linked to the amount of

chlorophyll, i.e. living needles or leaves, in the standing trees, which affect the reflectance in the near-infrared spectrum (Jutras-Perreault et al. 2023, Polvivaara et al. 2024). In our data, some of the snow-disturbed trees had already lost their living needles. Although the lidar intensity features showed some potential in the detection of snow-disturbed trees, the use of lidar intensity must be carefully considered in target domains comprising lidar data collected with several sensors. This is because intensity recordings may not be comparable among sensors (Wagner et al. 2006).

We reported F1-scores of 0.66 and 0.85 for the tree-level and plot-level classifications, respectively. The results of the tree-level classification reported here are poorer than the values reported in Puliti and Astrup (2022) who used high-resolution UAV imagery for the classification of healthy and broken trees (precision: 0.76, recall: 0.78). However, a direct comparison of the two studies is not possible due to the use of high-resolution UAV imagery acquired at low flight altitudes in Puliti and Astrup (2022). Our study was

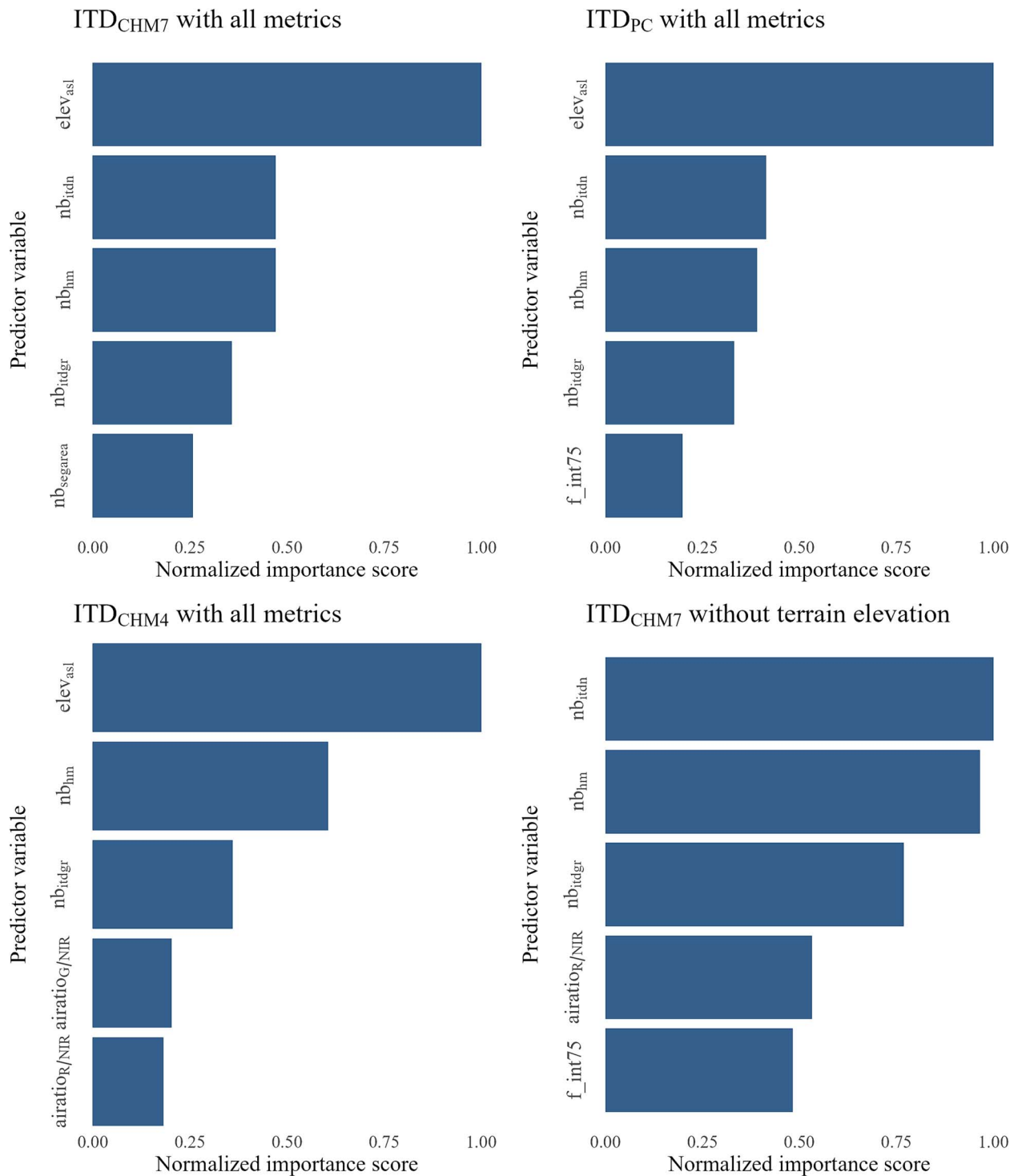


Figure 5. Five most important predictor variables associated with the RF classifier of healthy and broken trees by the ITD alternatives.

targeted at an operational application of several hundreds or thousands of hectares at a time and such an area cannot be covered with the current UAV technology in an operational application. From the point of view of operational application, performance assessment at the plot- or stand level is relevant. Tomppo et al. (2019) reported an overall accuracy of 0.9 for the detection of snow disturbance stands based on C-band SAR scenes, although their method was based on a time series of data, whereas we relied on RS data at a single time point. In addition, the exact magnitude of snow disturbance was not known in their reference

data, which originated from stand-level operational reports on cuttings caused by snow disturbance. It is realistic to assume that such cuttings are reported and implemented only when the magnitude of snow disturbance is so severe that the vitality and productivity of the forest is threatened. In our analyses, we included snow disturbance plots where the proportion of trees broken by snow ranged from 0% and 68%.

Our results indicate that the proposed detection of snow disturbance is prone to systematic errors, which must be considered when using the resulting maps for decision-making.

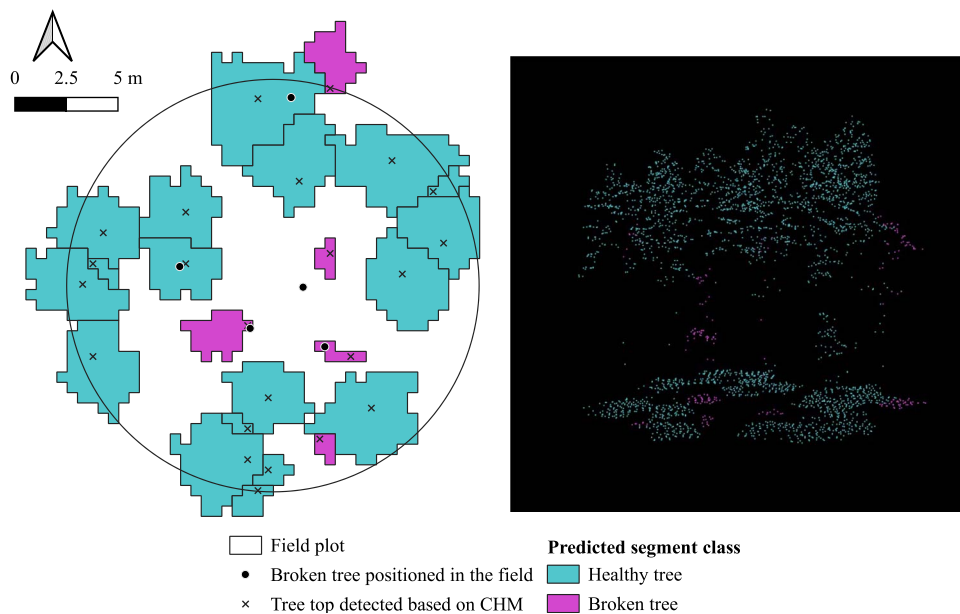


Figure 6. Left: Example of the detection of the trees damaged by snow in a pine forest. Right: Airborne lidar (light detection and ranging) data colored by segmentation and predicted classes. In this plot, all trees broken by snow were positioned in the field ($n = 5$). CHM – canopy height model.

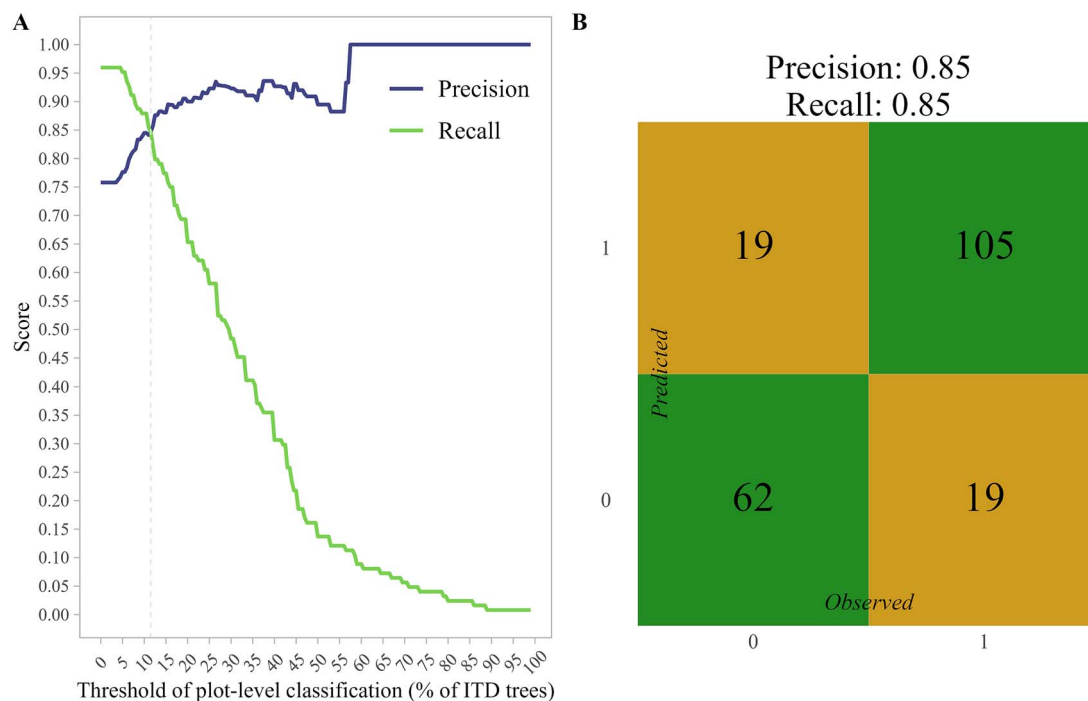


Figure 7. (A) Precision and recall curves with different thresholds of plot-level classification. The threshold implies the proportion of ITD trees classified as broken required to classify the field plot as a snow disturbance plot. (B) Confusion matrix of the plot-level classification with the threshold (11%) at the intercept of the precision and recall curves. 1—tree broken by snow, 0—healthy tree.

A recommended application of the RS-based snow disturbance map is that it would guide field groups to target stands that most likely have snow disturbance. This reduces inventory costs due to the improved efficiency associated with the fieldwork. Alternatively, RS-based forest disturbance maps may also be useful in the context of sample-based forest inventories in which auxiliary information are used in the planning of additional field measurements (Roberge et al. 2017) or in the design-based estimation of forest disturbance areas (Schroeder et al. 2014).

Our approach underpredicted the number of broken trees, which presumably implies that minor disturbance events cannot

be detected as accurately as severe disturbance events. However, the detection of severe, not minor, snow disturbance is what is crucial for operational forestry because disturbances caused by snow may trigger the requirement (proposed by Finnish legislation) to remove damaged trees in order to prevent subsequent insect outbreaks. The Forest Damages Prevention Act 1087/2013 (2013) obliges forest owners to remove trees from the forest if the amount of damaged timber exceeds 10 m³/ha for spruce or 50m³/ha for pine.

The proposed detection method for snow disturbance requires lidar data from which individual trees can be identified. This

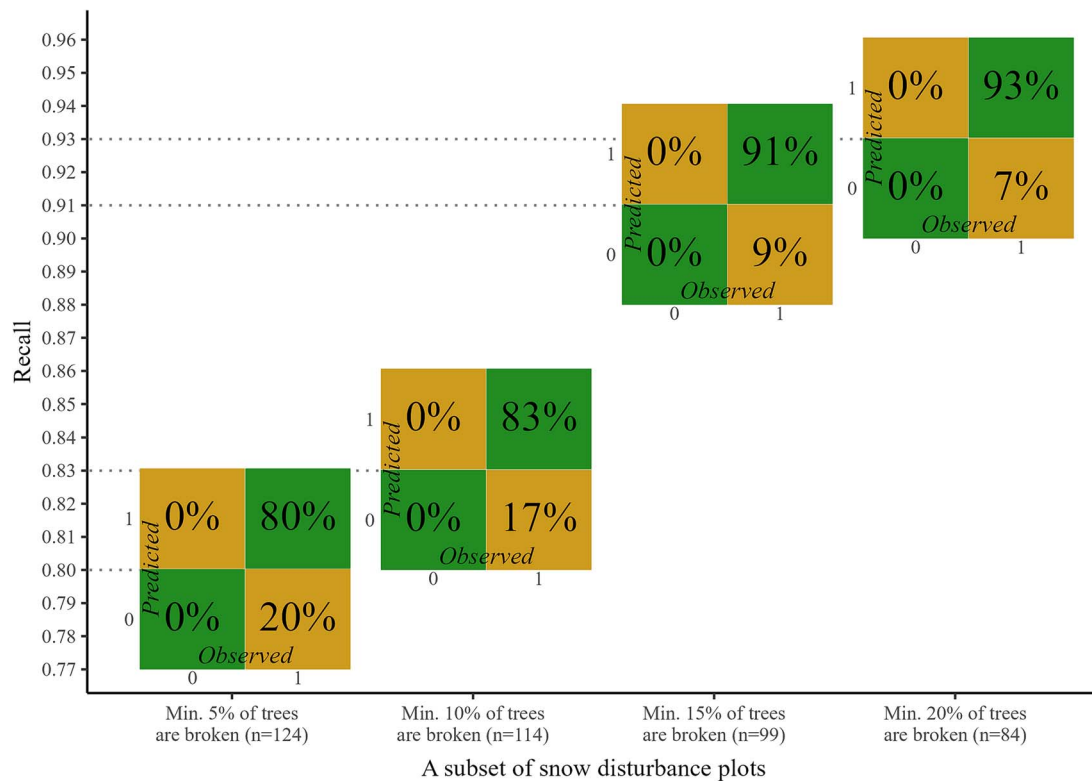


Figure 8. Recall scores (dotted line) and confusion matrices associated with the classification of snow disturbance plots by disturbance severities. The threshold value of plot-level classification was set at 11%.

means that the mapping of snow disturbance can only be carried out when new airborne lidar data are available. Typically, the acquisition cycle of airborne lidar data is several years, which causes temporal challenges given that forest owners are expected to react promptly to snow disturbance events in managed forests. While the temporal resolution of several years is a sufficient mapping interval in some applications, it may not be acceptable for the detection of snow disturbance.

While this study showed that lidar data at a single time point are appropriate for the detection of snow disturbance, previous studies have also reported the potential of multitemporal RS datasets for the detection of snow-induced canopy changes (Vashtaranta et al. 2012, Tomppo et al. 2019). It is also worth noting that 3D RS data could be collected using UAVs whenever and wherever needed. However, the operational applications of UAVs for data collections in large areas are currently missing due to higher price tag per hectare compared with national airborne-based data collections. In large-area mapping, it would also make sense to account for the interaction between weather conditions and the risk of snow disturbance, as suggested by Suvanto et al. (2021) and Zubkov et al. (2024).

Conclusions

We draw the following conclusions based on our results:

- (1) While the CHM-based ITD performed slightly better than point cloud-based ITD in the classification of healthy and broken trees, both underestimated the number of broken trees at the plot level.
- (2) Terrain elevation above sea level and the neighborhood features computed based on ITD (e.g. number of neighbor trees and deviation of neighbor tree heights) were among the most

important predictor variables used to distinguish between broken and healthy trees.

- (3) Although a tree-level F1-score of 0.66 indicated room for improvement, the proposed approach achieved a plot-level F1-score of 0.85, which highlights the potential of unitemporal remotely sensed data for the detection of snow disturbances alongside operational forest inventories relying on similar remotely sensed data sources.

Acknowledgements

We acknowledge the work of Aki Suvanto in the processing of the image data.

Author contributions

Janne Rätty (Conceptualization, Formal analysis, Investigation, Methodology, Software, Writing—original draft), Mikko Kukkonen (Conceptualization, Writing—review & editing), Markus Melin (Conceptualization, Funding acquisition, Project administration, Writing—review & editing), Matti Maltamo (Conceptualization, Funding acquisition, Writing—review & editing), Petteri Packalen (Conceptualization, Funding acquisition, Project administration, Software, Supervision, Writing—review & editing).

Conflict of interest

None declared.

Funding

This work is a contribution to the *LumiLaser* project (decision 5145/2021) funded by the Ministry of Agriculture and Forestry of

Finland. The work was also supported by the Research Council of Finland through the Finnish Flagship Programme for the Forest-Human-Machine Interplay—Building Resilience, Redefining Value Networks and Enabling Meaningful Experiences (UNITE) [337655], and the project Is climate smart forestry a utopia if the preferences of landowners are not considered? (UTOPIA) [352782].

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

References

- Aarnio L, Nuorteva H, Ylioja T. Kansalaisten tekemien metsätuhoilmoitusten satoa vuodelta 2021. In: Melin M, Terhonen E, (eds.). *Metsätuhot Vuonna 2021, Luonnonvara- ja Biotalous Tutkimus*. Helsinki: Luonnonvarakeskus, 2022, 7–14.
- Ahti K. *Huurteen Muodostumiseen Ja määrään Vaikuttavista tekijöistä Suomessa* (MSc thesis). Helsinki: University of Helsinki, 1978.
- Axelsson P. DEM generation from laser scanner data using adaptive TIN models *Int Arch Photogramm Remote Sens*. 2000;**33**:110–7.
- Breiman L. Random forests *Mach Learn*. 2001;**45**:5–32. <https://doi.org/10.1023/A:1010933404324>.
- Chawla NV, Bowyer KW, Hall LO. et al. SMOTE: synthetic minority over-sampling technique *J Artif Intell Res*. 2002;**16**:321–57. <https://doi.org/10.1613/jair.953>.
- Díaz-Yáñez O, Mola-Yudego B, González-Olabarria JR. Modelling damage occurrence by snow and wind in forest ecosystems *Ecol Model*. 2019;**408**:108741, 1–12. <https://doi.org/10.1016/j.ecolmodel.2019.108741>.
- Forest Damages Prevention Act 1087/2013 [WWW Document], 2013. URL <https://www.finlex.fi/fi/laki/smur/2013/20131087> (accessed 1.20.24).
- GDAL/OGR contributors. GDAL/OGR geospatial data abstraction software library *Open Source Geospatial Foundation*. 2024. <https://doi.org/10.5281/zenodo.5884351>.
- Gombin J, Vaidyanathan R, Agafonkin V. *Concaveman: A Very Fast 2D Concave Hull Algorithm*, 2020. R package version 1.1.0, <https://CRAN.R-project.org/package=concaveman>.
- Heikinheimo O. *Suomen Lumituhoalueet ja Niiden metsät*, Vol. 3. Metsäntutkimuslaitos: Communicationes Instituti Forestalis Fenniae, 1920.
- Hyypä J, Inkinen M. Detecting and estimating attributes for single trees using laser scanner *Photogramm J Finland*. 1999;**16**:27–42.
- Jalkanen A, Mattila U. Logistic regression models for wind and snow damage in northern Finland based on the National Forest Inventory data *For Ecol Manag*. 2000;**135**:315–30. [https://doi.org/10.1016/S0378-1127\(00\)00289-9](https://doi.org/10.1016/S0378-1127(00)00289-9).
- Jutras-Perreault M-C, Gobakken T, Ørka HO. Detecting the presence of standing dead trees using airborne laser scanning and optical data *Scand J For Res*. 2023;**38**:208–20. <https://doi.org/10.1080/02827581.2023.2211807>.
- Juutinen P. Männyn Toipuminen Kolilla Talven 1947–48 Lumituhojen jälkeen. Vol. 41. *Metsäntutkimuslaitos: Communicationes Instituti Forestalis Fenniae*, 1953, 43.
- Korhonen KT, Ahola A, Heikkinen J. et al. Forests of Finland 2014–2018 and their development 1921–2018 *Silva Fenn*. 2021;**55**:1–49. <https://doi.org/10.14214/sf.10662>.
- Kostensalo J, Mehtätalo L, Tuominen S. et al. Recreating structurally realistic tree maps with airborne laser scanning and ground measurements *Remote Sens Environ*. 2023;**298**:113782, 1–20. <https://doi.org/10.1016/j.rse.2023.113782>.
- Laasasenaho J. *Taper Curve and Volume Functions for Pine, Spruce and Birch*. Metsäntutkimuslaitos: Communicationes Instituti Forestalis Fennica, 1982.
- Lehtonen I, Kämäräinen M, Gregow H. et al. Heavy snow loads in Finnish forests respond regionally asymmetrically to projected climate change *Nat Hazards Earth Syst Sci*. 2016;**16**:2259–71. <https://doi.org/10.5194/nhess-16-2259-2016>.
- Li W, Guo Q, Jakubowski MK. et al. A new method for segmenting individual trees from the lidar point cloud *Photogramm Eng Remote Sens*. 2012;**78**:75–84. <https://doi.org/10.14358/PERS.78.1.75>.
- Narendra PM, Goldberg M. Image segmentation with directed trees *IEEE Trans Pattern Anal Mach Intell*. 1980;**PAMI-2**:185–91. <https://doi.org/10.1109/TPAMI.1980.4766999>.
- Näslund M. Skogsförsöksanstaltens gallringsförsök i tallskog *Meddelanden från Statens Skogsförsöksanstalt*. 1936;**29**:169.
- National land survey of Finland. Elevation model 2m. 2023. Available at <https://www.maanmittauslaitos.fi/en/maps-and-spatial-data/datasets-and-interfaces/product-descriptions/elevation-model-2-m>. Accessed 1/2024.
- Nykänen M-L, Peltola H, Quine C. et al. Factors affecting snow damage of trees with particular reference to European conditions, Vol. 31. *Silva Fennica*, 1997. <https://doi.org/10.14214/sf.a8519>.
- Päättälö M-L. Risk of snow damage in unmanaged and managed stands of scots pine, Norway spruce and birch *Scand J For Res*. 2000;**15**:530–41. <https://doi.org/10.1080/028275800750173474>.
- Packalén P, Suvanto A, Maltamo M. A two stage method to estimate species-specific growing stock *Photogramm Eng Remote Sens*. 2009;**75**:1451–60. <https://doi.org/10.14358/PERS.75.12.1451>.
- Pitkänen J, Maltamo M, Hyypä J. et al. Adaptive methods for individual tree detection on airborne laser based canopy height model *Int Arch Photogramm Remote Sens Spat Inf Sci*. 2004;**36**:187–91.
- Polvivaara A, Korpela I, Vastaranta M. et al. Detecting tree mortality using waveform features of airborne LiDAR *Remote Sens Environ*. 2024;**303**:114019. <https://doi.org/10.1016/j.rse.2024.114019>.
- Probst P, Wright M, Boulesteix A-L. Hyperparameters and Tuning Strategies for Random Forest. *WIREs Data Mining Knowl Discov*. 2019;**9**:e1301. <https://doi.org/10.1002/widm.1301>.
- Puliti S, Astrup R. Automatic detection of snow breakage at single tree level using YOLOv5 applied to UAV imagery *Int J Appl Earth Obs Geoinf*. 2022;**112**:102946. <https://doi.org/10.1016/j.jag.2022.102946>.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2024.
- Rätty J, Packalén P, Kotivuori E. et al. Fusing diameter distributions predicted by an area-based approach and individual-tree detection in coniferous-dominated forests *Can J For Res*. 2020;**50**:113–25. <https://doi.org/10.1139/cjfr-2019-0102>.
- Roberge C, Grafström A, Ståhl G. Forest damage inventory using the local pivotal sampling method *Can J For Res*. 2017;**47**:357–65. <https://doi.org/10.1139/cjfr-2016-0411>.
- Roussel J-R, Auty D. Airborne LiDAR Data Manipulation and Visualization for Forestry Applications, 2022. R package version 3.1.0. <https://cran.r-project.org/package=lidR>.
- Roussel J-R, Auty D, Coops NC. et al. lidar: an R package for analysis of airborne laser scanning (ALS) data *Remote Sens Environ*. 2020;**251**:112061. <https://doi.org/10.1016/j.rse.2020.112061>.
- Schroeder LM, Eidmann HH. Attacks of bark- and wood-boring coleoptera on snow-broken conifers over a two-year period *Scand J For Res*. 1993;**8**:257–65. <https://doi.org/10.1080/02827589309382775>.
- Schroeder TA, Healey SP, Moisen GG. et al. Improving estimates of forest disturbance by combining observations from Landsat time series with U.S. Forest Service Forest inventory and analysis

- data *Remote Sens Environ.* 2014;**154**:61–73. <https://doi.org/10.1016/j.rse.2014.08.005>.
- Siitonen J. Forest Management, Coarse Woody Debris and Saprophytic Organisms: Fennoscandian Boreal Forests as an Example. *Ecological Bulletins*, 2001, 11–41. <http://www.jstor.org/stable/20113262>.
- Suominen O. Susceptibility of stands to devastation by snow. Investigation into snow devastation in southern Finland in winter 1958–59 *Silva Fenn.* 1961;**112**:1–35. <https://doi.org/10.14214/sf.a14234>.
- Suvanto S, Lehtonen A, Nevalainen S. et al. Mapping the probability of forest snow disturbances in Finland *PLoS One.* 2021;**16**:1–20. <https://doi.org/10.1371/journal.pone.0254876>.
- Tomppo E, Antropov O, Praks J. Boreal Forest snow damage mapping using multi-temporal Sentinel-1 data *Remote Sens.* 2019;**11**:1–20. <https://doi.org/10.3390/rs11040384>.
- Torgo L. An infra-structure for performance estimation and experimental comparison of predictive models in R. 2014; **abs/1412.0436**:1–40. <http://arxiv.org/abs/1412.0436>.
- Valinger E, Fridman J. Models to assess the risk of snow and wind damage in pine, spruce, and birch forests in Sweden *Environ Manag.* 1999;**24**:209–17. <https://doi.org/10.1007/s002679900227>.
- Valinger E, Lundqvist L, Bondesson L. Assessing the risk of snow and wind damage from tree physical characteristics *Forestry.* 1993;**66**: 249–60. <https://doi.org/10.1093/forestry/66.3.249>.
- Vastaranta M, Korpela I, Uotila A. et al. Mapping of snow-damaged trees based on bitemporal airborne LiDAR data *Eur J Forest Res.* 2012;**131**:1217–28. <https://doi.org/10.1007/s10342-011-0593-2>.
- Vauhkonen J, Mehtätalo L. Matching remotely sensed and field-measured tree size distributions *Can J For Res.* 2015;**45**:353–63. <https://doi.org/10.1139/cjfr-2014-0285>.
- Wagner W, Ullrich A, Ducic V. et al. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner *ISPRS J Photogramm Remote Sens.* 2006;**60**: 100–12. <https://doi.org/10.1016/j.isprsjprs.2005.12.001>.
- Wright MN, Ziegler A. Ranger: a fast implementation of random forests for high dimensional data in C++ and R *J Stat Softw.* 2017;**77**:1–17. <https://doi.org/10.18637/jss.v077.i01>.
- Zubkov P, Gardiner B, Nygaard BE. et al. Predicting snow damage in conifer forests using a mechanistic snow damage model and high-resolution snow accumulation data *Scand J For Res.* 2024;**39**: 59–75. <https://doi.org/10.1080/02827581.2023.2289660>.