European Journal of Forest Research
https://doi.org/10.1007/s10342-025-01789-y

RESEARCH

Can we monitor seedling stands using Landsat Time Series?

Parvez Rana[1] · Eero Muinonen[2,3] · Miika Tolvanen[3] · Anwarul Chowdhury[1] · Mikko Vastaranta[3] · Timo Tokola[3]

Received: 13 July 2024 / Revised: 7 March 2025 / Accepted: 7 April 2025

Abstract
Monitoring the status of seedling stands is crucial for determining whether seedlings have achieved the required density and height and assessing the need for weeding. We studied the potential of Landsat Time Series (LTS) data for monitoring young seedling stands in Liperi, eastern Finland. We assessed the ability of Normalized Burn Ratio (NBR) and Normalized Difference Vegetation Index (NDVI) to estimate stand-level attributes like stem density, height, and identify the need for weeding of both broadleaved and coniferous trees. According to the results, the variation within a stand is typically very high, and thus it is difficult to give a single prediction for an entire stand. In our study, the indices could not fully capture the structural variation in seedling stands, although some trends could be found between NDVI and the number of coniferous trees ($R^2 = 0.22–0.36$), height of the coniferous trees ($R^2 = 0.09–0.52$) and height difference between coniferous and deciduous trees ($R^2 = 0.25–0.38$). Our study also shows that the prediction of the need for weeding or tending using a binary decision-making process achieved an accuracy of 81% and a Cohen's kappa value of 0.55. Our study demonstrates that the LTS data can be used with a reasonable accuracy to monitor weeding or tending needs in seedling stands but its capability for predicting height and density is limited.

Keywords Regeneration · Seedling stands · Landsat · Monitoring · Seedling survival

Introduction

Forest management plans in Finland are based on airborne laser scanning (ALS) forest inventory techniques (Hämäläinen et al. 2014). Forest management practices have typically included clear-cutting followed by planting and intensive tending of the seedling and sapling stands (e.g., weeding and cleaning). In Finland, stands dominated by coniferous trees with an average height of < 7 m and those dominated by deciduous trees with an average height of < 9 m are referred to as seedling stands (White et al. 2019; Rana et al. 2023), and fairly similar classifications are used in other countries in the Boreal region as well. Seedling stands are the future of next mature forest stands (Momen et al. 2004). Moreover, seedling stands can define the future appearance and composition of the forest (Imangholiloo et al. 2019). Therefore, monitoring and managing seedling stands is crucial for achieving a sustainable forest. Seedling stands management involves different silvicultural operation such as tending, thinning of competitive vegetation which determines the future development of the forest stands (Huuskonen and Hynynen 2006; Imangholiloo et al. 2020). Scheduling of these silvicultural operations in seedling stands is essential for avoiding growth losses and impaired wood quality during later development. Early and intensive precommercial thinning led to the most significant increase in diameter growth. Conducting precommercial thinning when the dominant height reached 3 m and maintaining 2,000 trees per hectare resulted in a 15% increase in the mean diameter by the time of the first commercial thinning (Huuskonen and Hynynen 2006). In addition, precommercial thinning timing has a significant impact on diameter development (Simard et al. 2004; Varmola and Salminen 2004; Fahlvik et al. 2005; Ulvcrona et al. 2007). Assessing

Communicated by Peter Biber.

✉ Parvez Rana
parvez.rana@luke.fi

[1] Natural Resources Institute Finland (Luke), Latokartanonkaari 9, Helsinki 00790, Finland

[2] Bitcomp Ltd, Joensuu FIN-80100, Finland

[3] School of Forest Sciences, University of Eastern Finland, Joensuu FIN-80100, Finland

the time of tending needs of seedling stands is crucial for sustaining future timber supply (Korhonen et al. 2013).

Regular monitoring is crucial for ensuring the success of young seedling stands. This typically involves site visits to verify that the seeding stands meet minimum standards for height and density, as mandated by the Finnish Forest Act (Ministry of Agriculture and Forestry 2010). These standards are outlined by Nilsson et al. (2010) and Kankaanhuhta (2014) and require seeding stands to achieve a specific height (over 0.5 m) and minimum stem density within 10 years of planting. The Act further specifies these density requirements, varying by tree species: 1300 stems per hectare for Scots pine (*Pinus sylvestris*), 1200 for Norway spruce (*Picea abies* (L.) Karst.), and 1000 for deciduous trees, primarily birches (*Betula* sp.).

Precise and up-to-date information of different forest stand attributes including tree density, species composition is important for sustainable management (Stoffels et al. 2015). By analyzing that information, a forest owner can make the decision of implementing appropriate silvicultural practices, for instance, the extent of tending and the optimal time of tending. Traditionally, collecting those information through field inventories is time-consuming, laborious, and expensive (Wittke et al. 2019; Imangholiloo et al. 2020; Rana et al. 2023). Remote sensing (RS) technologies are becoming popular in recent years in forest inventories (Kangas et al. 2018; Rana and Vauhkonen 2023; Hou et al. 2019). By considering both spatial and temporal resolution, the most common and efficient RS technology is satellite imagery data in forest inventories (Nguyen et al. 2018; Rana et al. 2016).

The use of satellite or aerial imagery is usually grounded in the premise that different tree densities and species compositions lead to different spectral responses. Different indices calculated from the spectral response of satellite imagery can be used to predict the structure and species composition of the stand (Pesonen et al. 2007; Korhonen et al. 2013; White et al. 2018, 2019). Spectral and structural information are also combined in many studies. For instance, Goodbody et al. (2018) achieved a detection accuracy of 86–95% for seedlings 5–15 years after planting by employing RGB-derived spectral indices and photogrammetry across two different sites. Feduck et al. (2018) used RGB (red, green, and blue) data detecting replanted coniferous seedlings with an accuracy of 76%. Studies using point clouds from photogrammetry or airborne laser scanning (ALS) are often based on the assumption that three-dimensional information on height, density and height variation can be used to predict the height, height variation and density of the from seedling stands tree (Nivala 2012; Puliti et al. 2019; Imangholiloo et al. 2019).

In recent years, time series data from Landsat and other similar satellites have been used in temporal analysis (Banskota et al. 2014; Nguyen et al. 2018). Time series of RS data offer information for estimating a range of forest attributes, including characteristics of trees, saplings, and seedlings, over large geographical areas. For instance, Pesonen et al. (2007) combined Landsat Time Series (LTS) data with national forest inventory (NFI) field data can be used to predict seedling stand tending needs, achieving better accuracy on mineral soils than on peatlands. Korhonen et al. (2013) reported good results (overall accuracy of 86%) using support vector machine at the plot level, although the stand-level accuracy had decreased to 72%.

Interestingly, mapping of forest recovery after disturbance with LTS has become popular and new methodologies have been developed for LTS data processing (Woodcock et al. 2008). For instance, Saarinen et al. (2018) studied the mean number of images per year (1984–2017) for LTS covering the whole of Finland and found that there were 3–9 images which had been acquired within ± 30 days of the target date (August 1) and with < 70% cloud cover. Even so, it is sometimes very difficult to find cloud-free coverage, and image mosaicking is needed to cover large inventory areas using data from a single year. LTS data can be used to characterize forest regeneration after various disturbances, such as certain harvesting practices (Chu et al. 2016; White et al. 2017). Concerning the needs for information on seedling stands, much remains to be understood regarding the relationship between changes in the spectral response and changes taking place in seedling stand canopies (White et al. 2018). Research into post-disturbance recovery (White et al. 2017), has shown that the follow-up of disturbances (e.g., fire, harvesting) can be monitored. Spectral variables and indices (e.g., Normalized Burn Ratio (NBR), Normalized Difference Vegetation Index (NDVI) derived from LTS have been shown to improve characterizations of seedling stands (Olsson 2009; Kennedy et al. 2012) and provide supporting information about biomass growth trends and species competition processes (White et al. 2018).

There are several studies which as focusing on seedling stands in predicting the tending requirement by using remote sensing data (Närhi et al. 2008; Korhonen et al. 2013; Miina et al. 2018). LTS can be an option in monitoring the seedling status post-disturbances such as harvesting and assess tending needs over time. However, there are few studies using LTS-derived NBR and NDVI data in this context. We had three research questions: (Q1) How do NBR and NDVI values vary in young seedling stands with different characteristics, such as site fertility and tree density? (Q2) How accurately can seedling stand-level attributes, such as tree density and height, be predicted for both broadleaved and coniferous trees? (Q3) What accuracy can be achieved in classifying the need for weeding and tending in seedling stands? We envisaged future developments in which RS data would reduce the costs of seedling stand monitoring activities and improve the planning of silvicultural operations.

Materials

Field data

A total of 371 circular field plots in the Liperi area (62° 31’ N, 29° 23’ E), eastern Finland (Fig. 1) were measured in summer 2016. The area concerned consists of managed boreal forests representing three site quality classes as defined by Cajander (1909) in terms of fertility and growth potential (Table 1). A specific field campaign was organized to collect data on young seedling stands (mean height of the trees lower than 1.3 m) and mature seedling stands (mean height higher than 1.3 m but below 7 m and mean diameter below 8 cm for both coniferous and broadleaved species). The stands to be measured were selected based on the existing stand register information. We then chose five circular sample plots of radius 2.66 m per stand, located in an ‘L’ pattern near the middle of the stand, their centres being 15 m apart. The mean heights of the conifers (CH) and broadleaf trees (BH) in each plot were determined in the field, and their stem densities were calculated by counting all the CH and BH trees within the plot that were more than 1.3 m in height. It is important to note that the study area was harvested (all 71 stands) in year 2010. We used a high-accuracy Trimble GeoXH 6000series GNSS receiver (Trimble, Sunnyvale, CA, USA) to pinpoint the exact locations of our sample plots. These locations were then post-processed using a differential GNSS algorithm, which relies on data from a nearby Trimble virtual reference station for even greater accuracy.

Remote sensing data

The study area spanned two overlapping Landsat scenes within the Worldwide Referencing System-2 (WRS-2). We searched for cloud-free image candidates in the Enhanced Thematic Mapper Plus (ETM+) and L1T formats, with cloud cover less than 70%. We then processed these images to create annual cloud-free mosaics representing the ground surface reflectance around July 15th (with a tolerance of 45 days) for each year between 2006 and 2017. The resulting mosaics had a spatial resolution of 30 m. In this way a 10-year history up to the reference year could be investigated and a total of 32 Landsat ETM LT05/LE07/LC08 sensor images could be selected representing the following years and date ranges: 2006 (4 images, June 19–August 13), 2007 (3 images, June 5–August 9), 2008 (1 image, July 25), 2009 (2 images, July 25–August 22), 2010 (3 images, June 29–August 17), 2011 (2 images, June 8–June 9), 2012 (1 image, August 14), 2013 (3 images, June 13–August 25), 2014 (2 images, June 13–August 25), 2015 (3 images, August 15–August 23), 2016 (3 images, June 21–August 8), and 2017 (5 images, June 8–August 11).

To minimize the influence of seasonal changes, we restricted image selection to those captured in June, July, and August. We then employed the *Fmask* method (Zhu and Woodcock 2012) to meticulously identify and exclude clouds and their shadows. This method ensured less than 10% cloud cover and that no major shadows fell on our sample plots. Finally, we used the Landsat Ecosystem Disturbance Adaptive System (LEDAPS) (Schmidt et al. 2013) to calibrate the remaining images, ensuring accurate surface reflectance values.

Methods

NBR and NDVI

The annual pixel-level series of NBR values has been identified as the most effective indicator of forest disturbance (including clear-cutting) and recovery among the spectral indices (Cohen et al. 2018; White et al. 2018, 2019). NBR, a spectral index (Key and Benson 2006) to detect burn severity, is calculated using Landsat ETM+bands 4 (NIR; near-infrared) and 7 (SWIR; short-wave infrared), as follows:

$$NBR = \frac{NIR - SWIR}{NIR + SWIR} \tag{1}$$

NBR was designed to identify the differences in the responses from disturbed (bare soil) and undisturbed areas (vegetation) visible in the near-infrared (NIR) and short-wave infrared (SWIR) spectral regions (see Cohen and Goward 2004; White et al. 2018). NBR will also be used with the NDVI, which uses the NIR (LT05/LE07:B4, LC08:B5) and red channels (LT05/LE07:B3, LC08:B4) in its formula:

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{2}$$

After the NBR and NDVI values had been calculated for each pixel and each year between 2006 and 2017, the changes in these indices were used to detect the year of regeneration and to analyze early stand development and its linkage to the structure and species composition of the stand. NBR and NDVI values for the stands were calculated

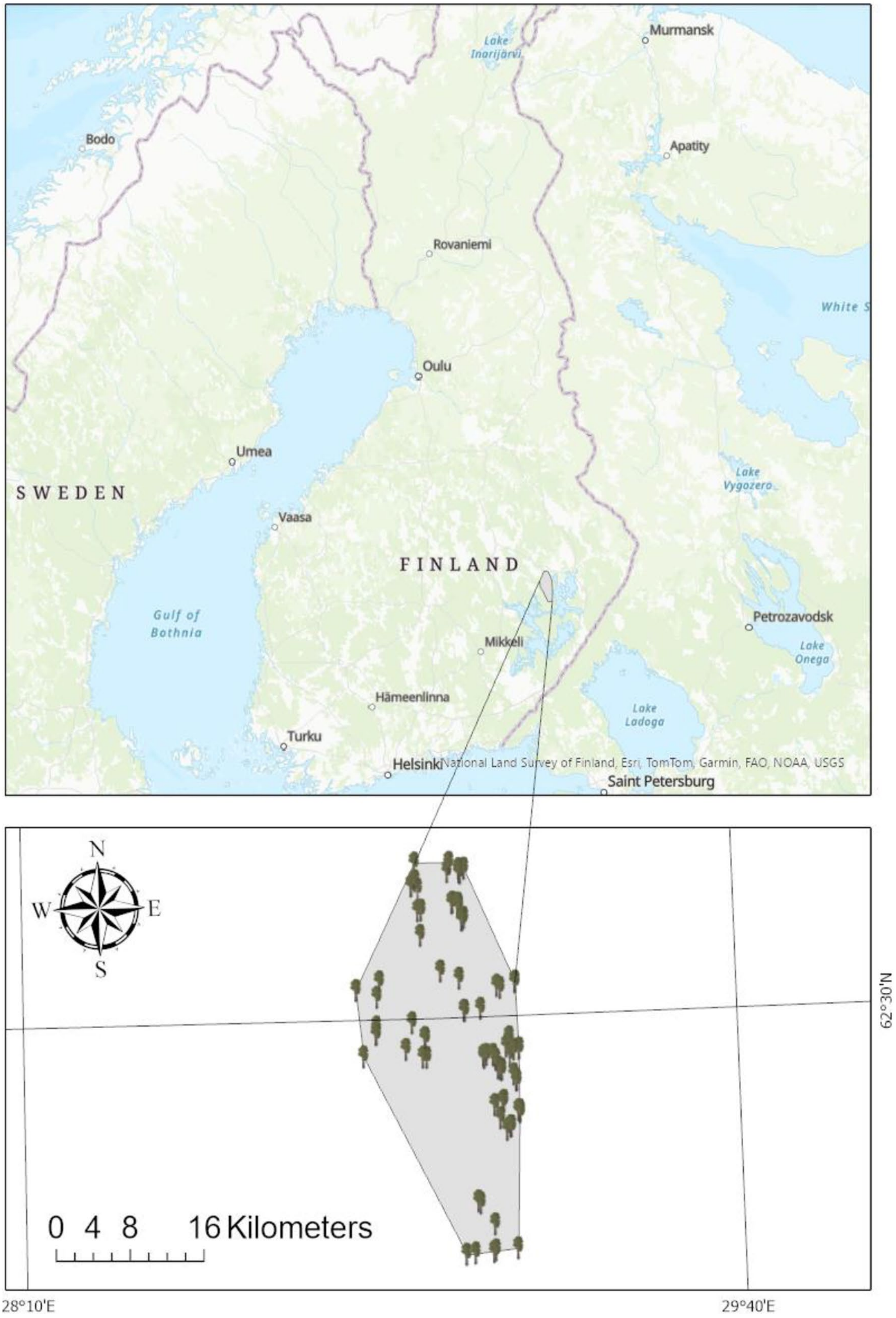

Fig. 1 Location of the area studied in Liperi, eastern Finland. Altogether 371 field plots belonging to 71 stands were measured (showed by tree icons). The coverage of the field information is indicated by the solid black lines

Table 1 Number of plots representing each site type measured in the field (n) and mean, standard deviation (Std), minimum (Min) and maximum (Max) values for the variables of interest

Site type (*n*)	Attributes of interest	Mean	Std	Min	Max
Fertile sites (78)					
	Stem density, conifers (count/ha), CD	2187	1073	0	8100
	Mean height, conifers (m), CH	1.69	1.04	0	5
	Stem density, broadleaf (count/ha), BD	8400	7753	0	40,950
	Mean height, broadleaf (m), BH	1.26	0.73	0	3.86
Semi-fertile sites (255)					
	Stem density, conifers (count/ha), CD	3730	2620	0	15,300
	Mean height, conifers (m), CH	1.82	1.23	0	7.1
	Stem density, broadleaf (count/ha), BD	9503	7932	0	42,750
	Mean height, broadleaf (m), BH	1.67	0.98	0	5.99
Poor sites (38)					
	Stem density, conifers (count/ha), CD	6095	2945	900	12,150
	Mean height, conifers (m), CH	1.46	0.75	0.3	3.4
	Stem density, broadleaf (count/ha), BD	4350	3625	0	13,500
	Mean height, broadleaf (m), BH	1.65	1.25	0	6
All sites (371)					
	Stem density, conifers (count/ha), CD	3598	2591	0	15,300
	Mean height, conifers (m), CH	1.76	1.17	0	7.1
	Stem density, broadleaf (count/ha), BD	8860	7840	0	42,750
	Mean height, broadleaf (m), BH	1.57	0.97	0	6

as the average of the pixels (one to three pixels per stand) within each stand. The stand-specific relative break year was defined according to the annual mean NBR or NDVI difference in the time series. An NBR or NDVI value greater than 0.25 indicates that the relative break year is marked by a drop of over 0.25 index units in the annual mean compared with the previous year. The break year was validated visually from the Landsat time series by investigating the forest structure before and after the detected break year. Developments in the NBR indices before and after the defined harvesting year (2010) were also analyzed by investigating the variation in the values using Box and Whiskers plots, which reveal changes in the distribution, median, and outliers, indicating potential impacts of the harvesting year.

Regeneration success and the need for tending were assessed using spectral indices derived from the fitted NBR and NDVI time series data, as in White et al. (2017). The indices were designed to characterize early stand development and were determined using trend-fitted index values from our time-series analysis. The sensitivity of the NBR and NDVI values to the numbers of stems in the tree density categories (number of stems: 0-3000, 3000–6000, 6000–9000, 9000–12000 and > 12000 stems per hectare) was also analyzed. This analysis is critical because silvicultural operations are typically required when tree density exceeds 3000 stems per hectare (Huuskonen and Hynynen 2006).

Density, height, and height differences prediction

The ordinary least squares regression (OLS) was designed to predict density, height and height differences in 2016. These can be predicted using historical NBR or NDVI index values as follows:

$$y = \alpha + \beta_1 VI_{2010} + \beta_2 \times VI_{2013} + \beta_3 \times VI_{2014} + \beta_4 \times VI_{2015} + \beta_5 \times VI_{2017} \quad (3)$$

where, y = density, height or height difference between coniferous and deciduous species in 2016,

VI = NRB or NDVI value. In addition to OLS, Gradient Boosting Machine (GBM) model was employed with 10-fold cross-validation to analyze the data. GBM is known for its robustness to class imbalance and noise, as well as its ability to handle multicollinearity among predictors (Dube et al. 2015; Lou et al. 2021; Qadeer et al. 2024). Hyperparameter tuning for GBM was conducted using a grid search strategy to identify the optimal combination of hyperparameters. The following hyperparameters were used:

- The number of boosting iterations (*n.trees*) was set to 100.
- The maximum depth of each tree (*interaction.depth*) was set to 2.
- The learning rate (*shrinkage*), which controls the contribution of each tree, was set to 0.01.
- The minimum number of observations in the terminal nodes (*n.minobsinnode*) was set to 8.

For validating tree density, height and height difference model, we employed root-mean-square errors (RMSEs) as

$$RMSE = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}(y_{ki} - \widehat{y}_{ki})^2} \quad (4)$$

and relative root-mean-square errors (RMSE%) as

$$RMSE\% = \left(\frac{RMSE}{\frac{1}{n}\sum\nolimits_{i=1}^{n} y_{ki}}\right) \times 100 \quad (5)$$

In this context, y_{ki} denotes the observed value for the sample stand, while $\widehat{y}_{ki}$ represents the predicted value for the same stand. Here, *i* indicates the specific sample stand, and *n* refers to the total number of stands. Predictive models are often evaluated using the coefficient of determination (R^2), a value between 0 and 1, shows how well the model explains the variation in the dependent variable. All modeling and cross-validation were conducted using the caret package (Kuhn 2008) in R (version 4.3.0).

Classifying weeding need

The height difference between coniferous and deciduous species taken as indicating the need for silvicultural operations i.e., weeding. Under Nordic Forest conditions it is typical for the dominant trees in seedling stands to be deciduous due to the rapid early development of these, so that the preferred or main tree species are suppressed. In this study, the height difference (<1 m) was used as a basis for a binary classification system (weeding needed vs. not needed) to remove unwanted deciduous trees (Miina et al. 2018; Rana et al. 2023). When broadleaves are taller than conifers, it indicates heightened competition, suggesting a greater need for tending. Conversely, if broadleaves are shorter than conifers, tending is generally unnecessary. Specifically, if the height difference exceeds 1 m (with conifers being taller), silvicultural operations are typically not required (Miina et al. 2018; Rana et al. 2023). The accuracy of classifying the need for weeding decisions was assessed using overall accuracy, and the kappa coefficient.

Results

NBR and NDVI values in young seedling stands (Q1)

The average variation in the NBR values was minimal (Fig. 2). The effect of regeneration was statically significant ($p<0.05$) compared to the values for the relative break year (Fig. 2). When multitemporal LTS data are available, clearcut stands can be detected using the NBR index, which was more sensitive in identifying a break-year change than the NDVI (Fig. 3). The trends in the NBR and NDVI indices after regeneration harvesting varied between the site types in the stands. For fertile forest site types (OMT), the mean NBR values changed from 0.28 to 0.66 and the NDVI values from 0.56 to 0.80 on average when the stand was harvested. In contrast, the values for semi-fertile sites (MT) changed from 0.33 to 0.50, and for poor sites (VT), the values changed from 0.65 to 0.80 on average. The differences between the NDVI curves were small on the poor sites (VT).

The variation between the tree density categories (number of stems: 0-3000, 3000–6000, 6000–9000, 9000–12000 and >12000 stems per hectare) was high (Fig. 4). This is indicating that there might also be a need to prioritize field operations to the most urgent sites.

Predicting seedling stand forests attributes (Q2)

The species-specific structural attributes using the NBR and NDVI indices can be predicted with limited accuracy (Table 2). For the number of deciduous seedlings per hectare, the RMSE was 5088 with an R^2 of 0.06 for NDVI, and 5006 with an R^2 of 0.14 for NBR. GBM results show an improvement in accuracy, with an RMSE of 4403 and an

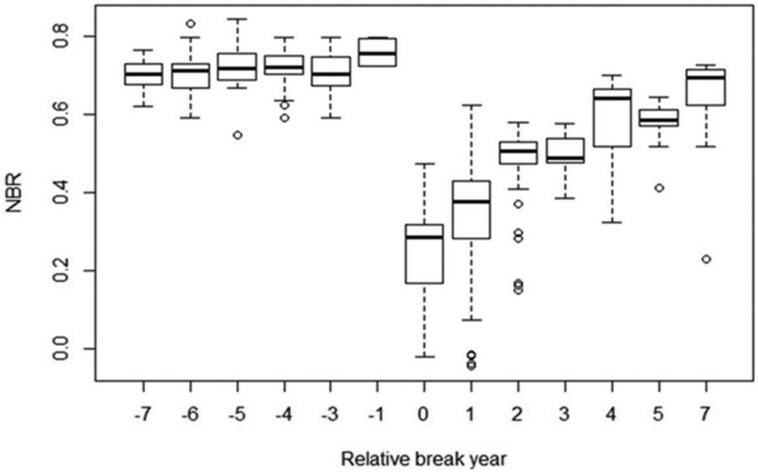

Fig. 2 Development of Normalized Burn Ratio (NBR) indices before and after defined harvesting years. The bold line in the box is the median and the upper and lower limits of the box indicate the variation of indices, while the circles indicate outliers

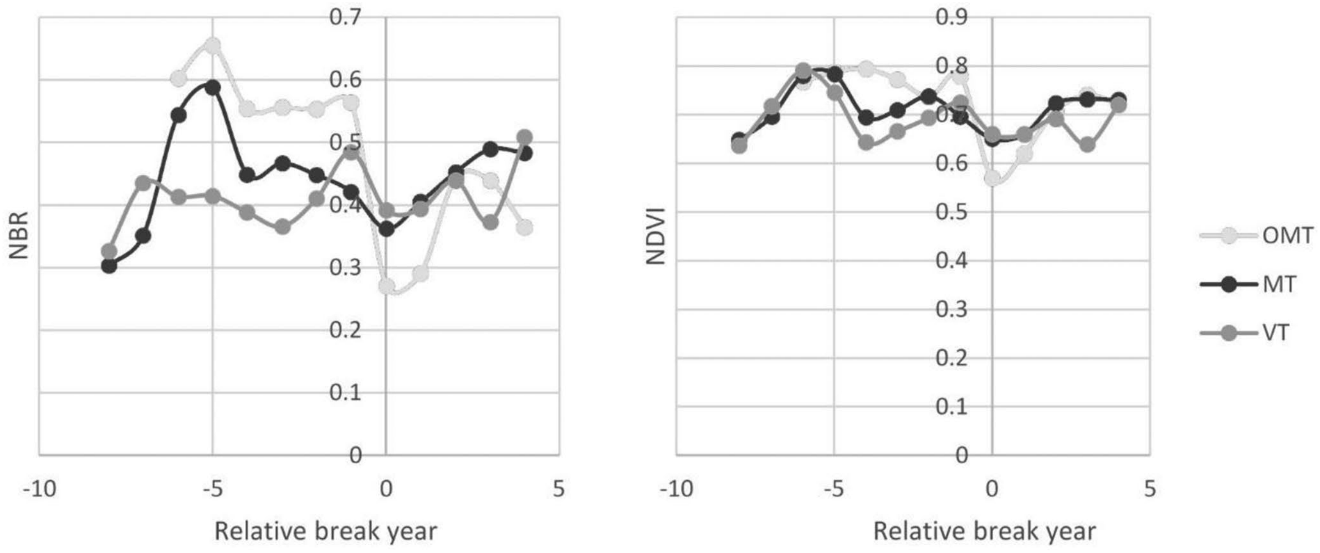

Fig. 3 Trends in the Normalized Burn Ratio (NBR) and Normalized Difference Vegetation Index (NDVI) by site types, fertile sites (OMT), semi-fertile sites (MT) and poor sites (VT), before and after the defined harvesting year (Relative break year)

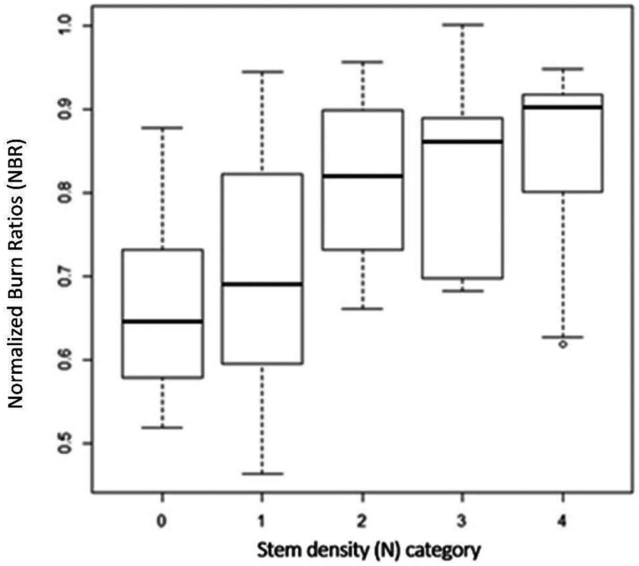

Fig. 4 Normalized Burn Ratios (NBR) across different stem density categories: 0 (0-<3000), 1 (3000-<6000), 2 (6000-<9000), 3 (9000-<12000), and 4 (≥12000). The bold line in the box is the median and the upper and lower limits of the box indicate the standard deviation, while the circles indicate outliers

R^2 of 0.35 for NDVI, and an RMSE of 4766 with an R^2 of 0.15 for NBR. For the number of coniferous seedlings per hectare, OLS results yielded an RMSE of 2794 with an R^2 of 0.22 for NDVI, and 3147 with an R^2 of 0.08 for NBR. GBM results demonstrated better predictive performance, with an RMSE of 2181 and an R^2 of 0.36 for NDVI, and an RMSE of 2864 with an R^2 of 0.27 for NBR. Similarly, for the total number of seedlings per hectare, OLS results indicated an RMSE of 6126 with an R^2 of 0.06 for NDVI, and 6276 with an R^2 of 0.06 for NBR, while GBM results showed a lower RMSE of 4746 with an improved R^2 of 0.38 for NDVI, and an RMSE of 5735 with an R^2 of 0.15 for NBR. Regarding the height of deciduous seedlings, OLS results showed an RMSE of 0.86 with an R^2 of 0.26 for NDVI, and 1.19 with an R^2 of 0.05 for NBR. In contrast, GBM results showed an RMSE of 0.95 with an R^2 of 0.39 for NDVI, and an RMSE of 1.08 with an R^2 of 0.17 for NBR. For the height of coniferous seedlings, OLS results showed an RMSE of 0.95 with an R^2 of 0.09 for NDVI, and 0.82 with an R^2 of 0.25 for NBR. GBM results demonstrated an RMSE of 1.20 with an

Table 2 Reliability of regression estimates for individual variables (in 2016) using historical the normalized difference vegetation index (NDVI) and normalized burn ratio (NBR) (applied only to 2010 as a break year). Root mean square error (RMSE) and coefficients of determination (R^2) are presented for the fitted prediction models. OLS = Ordinary least squares regression, gbm = gradient boosting machines

Variable in 2016	RMSE, NDVI	R^2, NDVI	RMSE, NBR	R^2, NBR
	OLS (GBM)	OLS (GBM)	OLS (GBM)	OLS (GBM)
Number of deciduous seedlings/ha	5088 (4403)	0.06 (0.35)	5006 (4766)	0.14 (0.15)
Number of coniferous seedlings /ha	2794 (2181)	0.22 (0.36)	3147 (2864)	0.08 (0.27)
Total number of seedlings/ha	6126 (4746)	0.06 (0.38)	6276 (5735)	0.06 (0.15)
Height deciduous (m)	0.86 (0.95)	0.26 (0.39)	1.19 (1.08)	0.05 (0.17)
Height conifers (m)	0.95 (1.20)	0.09 (0.52)	0.82 (0.77)	0.25 (0.34)
Height difference (conifer-deciduous) (m)	1.05 (1.45)	0.25 (0.38)	1.22 (1.16)	0.19 (0.24)

Table 3 Accuracy and Cohen's kappa value for classifying weeding need

Accuracy = 81%

Kappa = 0.55

		Prediction		Sum
		No need for weeding (h > 1)	Clear need (h < 1)	
Field	No need for weeding (h > 1)	10 (21.3%)	3 (6.4%)	**13 (27.7%)**
	Clear need (h < 1)	6 (12.7%)	28 (59.6%)	**34 (72.3%)**
	Sum	**16 (34.0%)**	**31 (66.0%)**	**47 (100%)**

R^2 of 0.52 for NDVI, and an RMSE of 0.77 with an R^2 of 0.34 for NBR.

Classifying weeding need (Q3)

When the predictions were divided into two categories (no need for weeding and clear need) and compared with the field observations, the predictions based on NBR showed an overall accuracy of 81% with a Cohen's kappa of 0.55 (Table 3).

Discussion

Monitoring the needs for weeding and tending operation in seedling stands are urgently required in Nordic forestry. An improved understanding of early stand development and its spatial and temporal characteristics is a prerequisite for achieving this goal. To meet these scientific objectives and the practical needs of decision-makers, we demonstrated the use of Landsat Time Series (LTS) data for monitoring and predicting the timing of regeneration in the forest stands. We evaluated how NBR and NDVI can be used to assess vegetation development in seedling stands with varying site types and to estimate the young seedling stands attributes.

NBR and NDVI values in young seedling stands

Our analysis focused on seedling stands with typical sizes of 1–3 hectares. In Landsat imagery, this corresponds to approximately 1–3 pixels, a resolution that is relatively coarse for capturing fine-scale variations within individual stands. This limitation makes it challenging to accurately represent the high spatial heterogeneity often present in these stands, such as differences in tree density, species composition, or growth conditions. Consequently, relying on a single prediction for the entire stand is not ideal, as it may overlook critical variations that could influence management decisions. Stand-level decisions should be based on the distribution of predictions within the stand boundaries. By analyzing the spatial patterns and variability of these stands, end-users can gain a more nuanced understanding of the stand's conditions. For example, areas with higher competition can be identified for the tending needs. This approach not only improves the accuracy of management decisions but also enhances the efficiency of resource allocation. Ultimately, the end-user needs to determine the feasible operational area based on this information.

Changes in NBR and NDVI indices provide valuable insights into the development of young trees after harvesting. Our findings, along with previous research (Schroeder et al. 2011; White et al. 2019; Chirici et al. 2020; Giannetti et al. 2020) indicate that clear-cut areas can be effectively identified using cloud-free imagery captured during suitable times of the year. For stands with excessive density (more than 3000 stems per hectare), weeding and tending are recommended treatments (Huuskonen and Hynynen 2006). These interventions are primarily necessary when unwanted vegetation, such as deciduous trees (often *Betula* sp.), hinders the growth of the desired conifer species. However, we acknowledge the challenges associated with using LTS data, such as the difficulty of acquiring cloud-free imagery and its limitations for stand-level predictions due to its lower spatial resolution.

Seedling stand forests attributes

While our study identified some trends between NDVI and NBR with seedling stand attributes (e.g., number of coniferous trees, height of deciduous trees, height difference), the indices could not fully capture the structural variation within these young stands. The R^2 values for all relationships were relatively low for OLS and medium for GBM (0.06 − 0.26 for OLS compared to 0.35 − 0.52 for GBM), and the RMSE values were medium to high, indicating limited accuracy in predicting specific attributes. These low R^2 values suggest a significant amount of unexplained variability, which may stem from the inherent heterogeneity within a pixel. Factors such as mixed species composition, varying seedling densities, age distribution, microsite conditions and spatial arrangement within a pixel likely contribute to this variability, highlighting the challenges in accurately characterizing stands inventory using spectral indices alone (White et al. 2016). For example, dense stands with overlapping canopies may lead to underestimation of individual tree counts, while variations in ground vegetation and shadow effects can introduce errors in height predictions (Kansanen et al. 2016). These challenges highlight the need for integrating multiple data sources and advanced analytical approaches and deep learning, to improve prediction accuracy and address spatial heterogeneity.

For comparison, a study by Rana et al. (2023) utilized multispectral airborne laser scanning data and reported similar findings for the number of broadleaved seedling (R^2: 0.14–0.25, RMSE: 5442–7782). In contrast, the regression model demonstrated higher accuracy for the number of coniferous seedling (R^2: 0.30–0.35, RMSE: 1292–2348). Additionally, study have shown strong predictive ability for tree height, particularly with broadleaved trees (R^2: 0.60–0.65, RMSE: 1.1–2.9) and even greater success for conifer heights (R^2: 0.74–0.82, RMSE: 0.9–2.5). This discrepancy in accuracy may be attributed to the centimeter-level resolution of the multispectral airborne laser scanning data used in Rana et al. (2023), as opposed to the coarser 30-meter resolution of Landsat data in this study. These findings suggest that alternative approaches (e.g., laser scanning data, Sentinel-2 and fine resolution UAV images) might be more suitable for detailed structural assessments in seedling stands. For instance, Feduck et al. (2018) found that UAV RGB imagery successfully detected 76% of coniferous seedlings, suggesting that UAV RGB imagery as a promising tool for identifying and monitoring coniferous seedlings. Furthermore, the integration of deep learning algorithms, such as convolutional neural networks, with UAV multispectral or RGB imagery has been shown to enhance species classification in seedling stand inventories (Imangholiloo et al. 2023). Similarly, Sentinel-2's multispectral capabilities and 10-meter spatial resolution offer opportunities for improved forest inventory parameter, especially when combined with other data sources and machine learning techniques (Wittke et al. 2019). In addition, incorporating individual tree features and correcting edge tree effects to the area-based approach (ABA) enhanced tree height and density prediction in seedling forest stands (Imangholiloo et al. 2022).

Classifying weeding need

The height difference between coniferous and deciduous trees serves as a valuable indicator for weeding or tending needs (Rana et al. 2023). Using this information for binary classification (weeding needed vs. not needed) resulted in a Cohen's kappa of 0.55. However, it is important to note that there is a great deal of variation within forest stands in the study area. For instance, a single Landsat pixel covers reflectance information from mixed seedling stands, introducing inherent inhomogeneity. This variability particularly affects model reliability in smaller (1–3 hectares) stands which are common in Finland. Despite this, the trend in the LTS-based prediction indicates that the need of silvicultural treatment in seedling stands can be effectively mapped. Additionally, low fertility sites are less sensitive to vegetation recovery and poses difficult challenges for LTS -based prediction.

Our study achieved an accuracy of 81% (kappa = 0.55) for predicting tending needs using the NBR. In comparison, a study by Korhonen et al. (2013) in Joutsa, Finland, reported lower performance (kappa = 0.37–0.38, accuracy = 71–72%) using airborne laser scanning data. Similar trends were observed in other Finnish studies, with accuracy ranging from 54 to 69% and kappa values from 0.27 to 0.34 (Närhi et al. 2008; Miina et al. 2018). These comparisons suggest that our approach leveraging NBR derived from LTS data offers a promising alternative for identifying stands requiring tending operations.

Implications for seedling stand monitoring method development

While temporal variations caused by phenology or atmospheric condition can be mitigated (e.g., Schroeder et al. 2007), challenges remain in monitoring attributes of young stands. NBR, which captures reflectance in both the NIR and SWIR regions, is particularly sensitive to this issue. As vegetation density and canopy complexity increase with stand age, SWIR reflectance decreases due to increased shadowing (Asner and Lobell 2000; Gonsamo and Pellikka 2012). This makes it difficult to accurately track early development in seedling stand, especially considering additional uncertainties from factors like atmospheric condition, topography, and sun/view angles (Song and Woodcock 2003; Chen et al. 2020). Incorporating forest stand data, particularly soil characteristics, could potentially improve the accuracy of results (Korhonen et al. 2013; Rana et al. 2023). Also, future research should incorporate a larger number of field sample plots encompassing a broader diversity of tree height, densities, and species composition (Imangholiloo 2024).

Seasonality also plays a significant role in Nordic seedling stands. Observations made in early summer differ greatly from those in late summer due to both vegetation phenology and variations in sun angle during image acquisition. For this reason, Sentinel-2 satellites, with their superior spatial resolution, hold promise for improved prediction compared to Landsat (Korhonen et al. 2017). Despite these challenges, the current Landsat-based approach demonstrates the potential for automated procedures to predict harvest years and identify stands in need of tending. Further development of reliable, large-scale methodologies utilizing satellite technology remains an urgent need for effective seedling stand monitoring.

Conclusion

We demonstrated LTS data for predicting seedling stand attributes and investigated the effect of NBR and NDVI on the predictive capacity. We also classified the need of weeding based on the height difference between coniferous and deciduous trees. Our results indicate that height and density of seedlings were less accurate in prediction, although there was a clear trend in the predictions. Consequently, relying on a single prediction for the entire stand is not ideal. Instead, stand-level decisions should be informed by the distribution of predictions within the stand boundaries. Ultimately, end-users must determine the feasible operational area based on this information. The height difference between coniferous tree species and the deciduous tree species is a valuable indicator for the need of weeding. When the predictions were used in a binary decision-making process regarding the need for weeding treatment, a reasonable classification accuracy was observed. However, we acknowledge the challenges associated with using LTS data, such as the difficulty of acquiring cloud-free imagery and its limitations for stand-level predictions due to its lower spatial resolution. Overall, the developed methodology shows promise for monitoring changes in seedling stand attributes and weeding needs. Future work should focus on improving prediction accuracy by integrating higher-resolution data (e.g., Sentinel-2) and addressing the limitations of LTS data to enhance its applicability in operational forestry management.

Author contributions E.M., P.R. and T.T. contributed to the study conception and design. E.M. did the analysis. E.M., P.R., M.T., A.C., M.V., and T.T. wrote the main manuscript draft. All authors reviewed the manuscript.

Funding Open access funding provided by Natural Resources Institute Finland.
This work was supported by the Ministry of Agriculture and Forestry of Finland [Dnro 2100/03.09.00/2015]. In addition, Parvez Rana received support through the RESONATE project (20366456), funded by the European Regional Development Fund (ERDF) under Interreg Aurora Program.

Data availability No datasets were generated or analysed during the current study.

Code availability Code available from corresponding author upon request.

Declarations

Competing interests The authors declare no competing interests.

References

Asner GP, Lobell DB (2000) A biogeophysical approach for automated SWIR unmixing of soils and vegetation. Remote Sens Environ 74:99–112. https://doi.org/10.1016/S0034-4257(00)00126-7

Banskota A, Kayastha N, Falkowski MJ et al (2014) Forest monitoring using Landsat time series data: A review. Can J Remote Sens 40:362–384. https://doi.org/10.1080/07038992.2014.987376

Cajander AK (1909) Über Waldtypen. Acta Forestalia Fennica 1:1–175

Chen R, Yin G, Liu G et al (2020) Evaluation and normalization of topographic effects on vegetation indices. Remote Sens 12:2290. https://doi.org/10.3390/rs12142290

Chirici G, Giannetti F, Mazza E et al (2020) Monitoring clearcutting and subsequent rapid recovery in mediterranean coppice forests with Landsat time series. Ann for Sci 77:40. https://doi.org/10.1007/s13595-020-00936-2

Chu T, Guo X, Takeda K (2016) Remote sensing approach to detect post-fire vegetation regrowth in Siberian boreal larch forest. Ecol Ind 62:32–46. https://doi.org/10.1016/j.ecolind.2015.11.026

Cohen WB, Goward SN (2004) Landsat's role in ecological applications of remote sensing. Bioscience 54:535. https://doi.org/10.1641/0006-3568(2004)054[0535:LRIEAO]2.0.CO;2

Cohen WB, Yang Z, Healey SP et al (2018) A landtrendr multispectral ensemble for forest disturbance detection. Remote Sens Environ 205:131–140. https://doi.org/10.1016/j.rse.2017.11.015

Dube T, Mutanga O, Abdel-Rahman EM, Ismail R, Slotow R (2015) Predicting Eucalyptus spp. Stand volume in Zululand, South Africa: an analysis using a stochastic gradient boosting regression ensemble with multi-source data sets. Int J Remote Sens 36(14):3751–3772

Fahlvik N, Ekö P-M, Pettersson N (2005) Influence of precommercial thinning grade on branch diameter and crown ratio in *Pinus sylvestris* in Southern Sweden. Scand J for Res 20:243–251. https://doi.org/10.1080/02827580510008266

Feduck C, McDermid GJ, Castilla G (2018) Detection of coniferous seedlings in UAV imagery. Forests 9:432. https://doi.org/10.3390/f9070432

Giannetti F, Pegna R, Francini S et al (2020) A new method for automated clearcut disturbance detection in mediterranean coppice forests using Landsat time series. Remote Sens 12:3720. https://doi.org/10.3390/rs12223720

Gonsamo A, Pellikka P (2012) The sensitivity based Estimation of leaf area index from spectral vegetation indices. ISPRS J Photogrammetry Remote Sens 70:15–25. https://doi.org/10.1016/j.isprsjprs.2012.03.009

Goodbody TRH, Coops NC, Hermosilla T et al (2018) Assessing the status of forest regeneration using digital aerial photogrammetry and unmanned aerial systems. Int J Remote Sens 39:5246–5264. https://doi.org/10.1080/01431161.2017.1402387

Hämäläinen J, Holopainen M, Hynynen J et al (2014) Perusteita Seuraavan Sukupolven metsävarajärjestelmälle– Forest big data -hanke. Metsätieteen Aikakauskirja 2014. https://doi.org/10.14214/ma.5884

Hou Z, Mehtätalo L, McRoberts RE, Ståhl G, Tokola T, Rana P, Siipilehto J, Xu Q (2019) Remote sensing-assisted data assimilation and simultaneous inference for forest inventory. Remote Sens Environ 234:111431

Huuskonen S, Hynynen J (2006) Timing and intensity of precommercial thinning and their effects on the first commercial thinning in Scots pine stands. Silva Fenn 40:645–662. https://doi.org/10.14214/sf.320

Imangholiloo M (2024) Toward an enhanced characterization of seedling stands using remote sensing. Dissertationes Forestales 1017. https://doi.org/10.14214/df.355

Imangholiloo M, Saarinen N, Markelin L et al (2019) Characterizing seedling stands using Leaf-Off and Leaf-On photogrammetric point clouds and hyperspectral imagery acquired from unmanned aerial vehicle. Forests 10:415. https://doi.org/10.3390/f10050415

Imangholiloo M, Saarinen N, Holopainen M et al (2020) Using Leaf-Off and Leaf-On multispectral airborne laser scanning data to characterize seedling stands. Remote Sens 12:3328. https://doi.org/10.3390/rs12203328

Imangholiloo M, Yrttimaa T, Mattsson T, Junttila S, Holopainen M, Saarinen N, Savolainen P, Hyyppä J, Vastaranta M (2022) Adding single tree features and correcting edge tree effects enhance the characterization of seedling stands with single-photon airborne laser scanning. ISPRS J Photogrammetry Remote Sens 191:129–142

Imangholiloo M, Luoma V, Holopainen M, Vastaranta M, Mäkeläinen A, Koivumäki N, Honkavaara E, Khoramshahi E (2023) A new approach for feeding multispectral imagery into convolutional neural networks improved classification of seedlings. Remote Sens 15(21):5233

Kangas A, Astrup R, Breidenbach J et al (2018) Remote sensing and forest inventories in nordic countries– roadmap for the future. Scand J for Res 33:397–412. https://doi.org/10.1080/02827581.2017.1416666

Kankaanhuhta V (2014) Quality management of forest regeneration activities. Diss 2014. https://doi.org/10.14214/df.174

Kansanen K, Vauhkonen J, Lähivaara T, Mehtätalo L (2016) Stand density estimators based on individual tree detection and stochastic geometry. Can J for Res 46(11):1359–1366

Kennedy RE, Yang Z, Cohen WB et al (2012) Spatial and Temporal patterns of forest disturbance and regrowth within the area of the Northwest forest plan. Remote Sens Environ 122:117–133. https://doi.org/10.1016/j.rse.2011.09.024

Key CH, Benson NC (2006) Landscape assessment (LA). FIREMON: Fire effects monitoring and inventory system 164:LA-1

Korhonen L, Pippuri I, Packalén P et al (2013) Detection of the need for seedling stand tending using high-resolution remote sensing data. Silva Fenn 47. https://doi.org/10.14214/sf.952

Korhonen L, Hadi, Packalen P, Rautiainen M (2017) Comparison of Sentinel-2 and Landsat 8 in the Estimation of boreal forest canopy cover and leaf area index. Remote Sens Environ 195:259–274. https://doi.org/10.1016/j.rse.2017.03.021

Kuhn M (2008) Building predictive models in R using the caret package. J Stat Softw 28(5). https://doi.org/10.18637/jss.v028.i05

Lou XW, Weng YH, Fang LM, Gao HL, Grogan J, Hung IK, Oswald BP (2021) Predicting stand attributes of loblolly pine in West Gulf coastal plain using gradient boosting and random forests. Can J for Res 51(6):807–816

Miina J, Turunen J, Korhonen KT et al (2018) Predicting the need for tending of conifer seedling stands in Southern Finland. Scand J for Res 33:641–649. https://doi.org/10.1080/02827581.2018.1447681

Ministry of Agriculture and Forestry (2010) Forest Decree 1234/2010. https://www.finlex.fi/fi/laki/alkup/2010/20101234#Pidp446415648 [cited August 8, 2018] In Finnish

Momen B, Anderson PD, Sullivan JH, Helms JA (2004) A multivariate statistical approach to predicting mature tree performance based on seedling characteristics. New Forest 27:303–313. https://doi.org/10.1023/B:NEFO.0000022226.95493.d7

Närhi M, Maltamo M, Packalén P et al (2008) Kuusen Taimikoiden inventointi Ja Taimikonhoidon Kiireellisyyden Määrittäminen Laserkeilauksen Ja Metsäsuunnitelmatietojen avulla [Inventory ofspruce seedling stands and determination of tending need by meansof laser scanning and forestry plan information]. https://doi.org/10.14214/ma.6419. Metsätieteen aikakauskirja 2008:

Nguyen TH, Jones SD, Soto-Berelov M et al (2018) A Spatial and Temporal analysis of forest dynamics using Landsat time-series. Remote Sens Environ 217:461–475. https://doi.org/10.1016/j.rse.2018.08.028

Nilsson U, Luoranen J, Kolström T et al (2010) Reforestation with planting in Northern Europe. Scand J for Res 25:283–294. https://doi.org/10.1080/02827581.2010.498384

Nivala MT (2012) Laserkeilauksen käyttö metsänhoitotarpeen määrittämisessä taimikoissa ja nuorissa kasvatusmetsiköissä [Using Laser scanning to examine need for forest management in seedlings and young forest stands]. Master's Thesis, Itä-Suomen yliopisto

Olsson H (2009) A method for using Landsat time series for monitoring young plantations in boreal forests. Int J Remote Sens 30:5117–5131. https://doi.org/10.1080/01431160903022993

Pesonen A, Korhonen K, Tuominen S et al (2007) Taimikonhoitotarpeen arviointi Valtakunnan Metsien Inventoinnin Metsävarakartan Pohjalta [Assessment of seedling stand tending need based on national forest inventory forest resource maps]. https://doi.org/10.14214/ma.6695. Metsätieteen aikakauskirja 2007:

Puliti S, Solberg S, Granhus A (2019) Use of UAV photogrammetric data for Estimation of biophysical properties in forest stands under regeneration. Remote Sens 11:233. https://doi.org/10.3390/rs11030233

Qadeer A, Shakir M, Wang L, Talha SM (2024) Evaluating machine learning approaches for aboveground biomass prediction in fragmented high-elevated forests using multi-sensor satellite data. Remote Sens Applications: Soc Environ 36:101291

Rana P, Vauhkonen J (2023) Stochastic multicriteria acceptability analysis as a forest management priority mapping approach based on airborne laser scanning and field inventory data. Landsc Urban Plann 230:104637

Rana P, Gautam B, Tokola T (2016) Optimizing the number of training areas for modeling above-ground biomass with ALS and multispectral remote sensing in subtropical Nepal. Int J Appl Earth Obs Geoinf 49:52–62

Rana P, Mattila U, Mehtätalo L et al (2023) Monitoring seedling stands using National forest inventory and multispectral airborne laser scanning data. Can J Res 53:302–313. https://doi.org/10.1139/cjfr-2022-0135

Schmidt G, Jenkerson CB, Masek J et al (2013) Landsat ecosystem disturbance adaptive processing system (LEDAPS) algorithm description. US Geological Survey

Schroeder TA, Cohen WB, Yang Z (2007) Patterns of forest regrowth following clearcutting in Western Oregon as determined from a Landsat time-series. For Ecol Manag 243:259–273. https://doi.org/10.1016/j.foreco.2007.03.019

Schroeder TA, Wulder MA, Healey SP, Moisen GG (2011) Mapping wildfire and clearcut harvest disturbances in boreal forests with Landsat time series data. Remote Sens Environ 115:1421–1433. https://doi.org/10.1016/j.rse.2011.01.022

Simard SW, Blenner-Hassett T, Cameron IR (2004) Pre-commercial thinning effects on growth, yield and mortality in even-aged paper Birch stands in British Columbia. For Ecol Manag 190:163–178. https://doi.org/10.1016/j.foreco.2003.09.010

Song C, Woodcock CE (2003) Monitoring forest succession with multitemporal Landsat images: factors of uncertainty. IEEE Trans Geosci Remote Sens 41:2557–2567. https://doi.org/10.1109/TGRS.2003.818367

Stoffels J, Hill J, Sachtleber T et al (2015) Satellite-Based derivation of High-Resolution forest information layers for operational forest management. Forests 6:1982–2013. https://doi.org/10.3390/f6061982

Ulvcrona KA, Claesson S, Sahlen K, Lundmark T (2007) The effects of timing of pre-commercial thinning and stand density on stem form and branch characteristics of Pinus sylvestris. Forestry 80:323–335. https://doi.org/10.1093/forestry/cpm011

Varmola M, Salminen H (2004) Timing and intensity of precommercial thinning in *Pinus sylvestris* stands. Scand J for Res 19:142–151. https://doi.org/10.1080/02827580310019545

White JC, Coops NC, Wulder MA, Vastaranta M, Hilker T, Tompalski P (2016) Remote sensing technologies for enhancing forest inventories: A review. Can J Remote Sens 42(5):619–641. https://doi.org/10.1080/07038992.2016.1207484

White JC, Wulder MA, Hermosilla T et al (2017) A nationwide annual characterization of 25 years of forest disturbance and recovery for Canada using Landsat time series. Remote Sens Environ 194:303–321. https://doi.org/10.1016/j.rse.2017.03.035

White JC, Saarinen N, Kankare V et al (2018) Confirmation of post-harvest spectral recovery from Landsat time series using measures of forest cover and height derived from airborne laser scanning data. Remote Sens Environ 216:262–275. https://doi.org/10.1016/j.rse.2018.07.004

White JC, Saarinen N, Wulder MA et al (2019) Assessing spectral measures of post-harvest forest recovery with field plot data. Int J Appl Earth Obs Geoinf 80:102–114. https://doi.org/10.1016/j.jag.2019.04.010

Wittke S, Yu X, Karjalainen M et al (2019) Comparison of two-dimensional multitemporal Sentinel-2 data with three-dimensional remote sensing data sources for forest inventory parameter Estimation over a boreal forest. Int J Appl Earth Obs Geoinf 76:167–178. https://doi.org/10.1016/j.jag.2018.11.009

Woodcock CE, Allen R, Anderson M et al (2008) Free access to Landsat imagery. Science 320:1011–1011. https://doi.org/10.1126/science.320.5879.1011a

Zhu Z, Woodcock CE (2012) Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sens Environ 118:83–94. https://doi.org/10.1016/j.rse.2011.10.028