Remote Sensing of Environment 216 (2018) 262–275

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Confirmation of post-harvest spectral recovery from Landsat time series using measures of forest cover and height derived from airborne laser scanning data

Joanne C. White[a,b,*], Ninni Saarinen[b], Ville Kankare[b], Michael A. Wulder[a], Txomin Hermosilla[c], Nicholas C. Coops[c], Paul D. Pickell[c], Markus Holopainen[b], Juha Hyyppä[d], Mikko Vastaranta[e]

[a] Canadian Forest Service, (Pacific Forestry Center), Natural Resources Canada, 506 West Burnside Road, Victoria, BC V8Z 1M5, Canada
[b] Department of Forest Sciences, P.O. Box 27, University of Helsinki, 00014, Finland
[c] Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada
[d] Department of Remote Sensing and Photogrammetry, Finnish Geospatial Research Institute FGI, National Land Survey of Finland, Masala 02431, Finland
[e] School of Forest Sciences, University of Eastern Finland, P.O. Box-111, Joensuu 80101, Finland

ARTICLE INFO

Keywords:
Landsat
Time series
Recovery
Regeneration
Harvest
ALS
Lidar
Boreal

ABSTRACT

Landsat time series (LTS) enable the characterization of forest recovery post-disturbance over large areas; however, there is a gap in our current knowledge concerning the linkage between spectral measures of recovery derived from LTS and actual manifestations of forest structure in regenerating stands. Airborne laser scanning (ALS) data provide useful measures of forest structure that can be used to corroborate spectral measures of forest recovery. The objective of this study was to evaluate the utility of a spectral index of recovery based on the Normalized Burn Ratio (NBR): the years to recovery, or Y2R metric, as an indicator of the return of forest vegetation following forest harvest (clearcutting). The Y2R metric has previously been defined as the number of years required for a pixel to return to 80% of its pre-disturbance NBR (NBR_{pre}) value. In this study, the Composite2Change (C2C) algorithm was used to generate a time series of gap-free, cloud-free Landsat surface reflectance composites (1985–2012), associated change metrics, and a spatially-explicit dataset of detected changes for an actively managed forest area in southern Finland (5.3 Mha). The overall accuracy of change detection, determined using independent validation data, was 89%. Areas of forest harvesting in 1991 were then used to evaluate the Y2R metric. Four alternative recovery scenarios were evaluated, representing variations in the spectral threshold used to define Y2R: 60%, 80%, and 100% of NBR_{pre}, and a critical value of z (i.e. the year in which the pixel's NBR value is no longer significantly different from NBR_{pre}). The Y2R for each scenario were classified into five groups: recovery within < 10 years, 10–13 years, 14–17 years, > 17 years, and not recovered. Measures of forest structure (canopy height and cover) were obtained from ALS data. Benchmarks for height (> 5 m) and canopy cover (> 10%) were applied to each recovery scenario, and the percentage of pixels that attained both of these benchmarks for each recovery group, was determined for each Y2R scenario. Our results indicated that the Y2R metric using the 80% threshold provided the most realistic assessment of forest recovery: all pixels considered in our analysis were spectrally recovered within the analysis period, with 88.88% of recovered pixels attaining the benchmarks for both cover and height. Moreover, false positives (pixels that had recovered spectrally, but not structurally) and false negatives (pixels that had recovered structurally, but not spectrally) were minimized with the 80% threshold. This research demonstrates the efficacy of LTS-derived assessments of recovery, which can be spatially exhaustive and retrospective, providing important baseline data for forest monitoring.

1. Introduction

Time series of remotely sensed data provide opportunities to characterize forest dynamics over large areas (Banskota et al., 2014). In particular, Landsat time series (LTS) support the characterization of long-term forest recovery (Chu et al., 2016; White et al., 2017); however much remains to be understood concerning the relationship between spectral measures and manifestations of recovery in forest

* Corresponding author at: Canadian Forest Service, (Pacific Forestry Center), Natural Resources Canada, 506 West Burnside Road, Victoria, BC V8Z 1M5, Canada.
E-mail address: joanne.white@canada.ca (J.C. White).

https://doi.org/10.1016/j.rse.2018.07.004
Received 9 February 2018; Received in revised form 14 June 2018; Accepted 2 July 2018

structural attributes. Definitions of forest recovery post-disturbance are not universal (Bartels et al., 2016) and often relate to the return of forest structural characteristics following a particular disturbance type (Frolking et al., 2009). Herein, we follow the approach of Frolking et al. (2009) and define recovery as the return of forest structure, quantified by measurable characteristics (e.g. canopy height and cover), against which target thresholds can be applied to indicate when recovery has occurred. In reality, forest recovery is a long-term ecological process, with different functions of a forest returning at different times through the successional process (Spake et al., 2015). Forest recovery post-disturbance is difficult to characterize using data from ground plots alone, particularly over large, remote areas with constraints to forest access (e.g. Canada; Bartels et al., 2016). In nations such as Finland, where intensive forest management practices prevail (Wulder et al., 2007), the capacity for synoptic, spatially-explicit monitoring of forest recovery through time, particularly in the context of a complex land use-land ownership mosaic, is of interest to resource managers and planners (Culotta et al., 2015). Remotely sensed assessments of forest recovery post-disturbance enable assessments of recovery over large spatial extents and different disturbance types (Frolking et al., 2009; Kennedy et al., 2012; Madoui et al., 2015), and provide a framework within which assessments of recovery from ground plot observations may be integrated (Bartels et al., 2016). Moreover remotely sensed assessments of recovery that take advantage of the Landsat archive enable retrospective studies, thereby providing baseline information for monitoring programs (White et al., 2017).

Airborne laser scanning (ALS) data have demonstrated capacity for accurately characterizing forest structure, but are typically limited either in spatial or temporal coverage. In contrast, Landsat data provide both large-area spatial coverage and a temporal archive that extends back to 1982 for 30 m spatial resolution data from the Landsat Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+), and Operational Land Imager (OLI) data. Landsat data have played an important role in the Finnish multi-source National Forest Inventory (MS-NFI) and since 1989, have been used as a means to cost-effectively obtain reliable forest information for areas smaller for which it is not possible to achieve target accuracies with the network of ground plots established for the NFI alone (e.g. a municipality) (Tomppo, 1990). Finland is now generating its 12th MS-NFI (Barrett et al., 2016). Tomppo et al. (2008) suggested that one potential option for enhancing the MS-NFI would be to incorporate historical satellite imagery as a source of additional information on the age and development of forests, citing that information on stand development would be particularly useful in Nordic countries because forest practices have typically been clearcutting (with some required number of retention trees/ha) followed by planting and intensive silviculture (e.g. weeding and cleaning of seedling and sapling stands). A nationwide acquisition of ALS data initiated by the National Land Survey of Finland (NLS) in 2008 has greatly expanded the coverage and availability of ALS data and related forest structural information across the country (Kotivuori et al., 2016).

Assessments of recovery via ground plots are valuable; however, these assessments are spatially and temporally constrained, (Bartels et al., 2016), precluding analyses that are both spatially explicit and spatially exhaustive. ALS data have been used to characterize post-fire forest structure and recovery (Bolton et al., 2015, 2017; Vogeler et al., 2015) and provide the requisite spatial detail and structural characterization; however, a single-date acquisition does not support retrospective assessments of forest structural development over time. Characterization of forest recovery with LTS has become increasingly common with the opening of the Landsat archive in 2008 (Woodcock et al., 2008). While post-disturbance recovery has been explored (Kennedy et al., 2012; Griffiths et al., 2014; Potapov et al., 2015; Frazier et al., 2015 and 2018; Senf et al., 2017), research has demonstrated that the disturbance agent (e.g. wildfire, harvest) influences recovery trajectories (Madoui et al., 2015; White et al., 2017). Characterizations of post-fire recovery with LTS are more common (e.g. Chu and Guo, 2014), with fewer studies focusing on post-harvest recovery (Schroeder et al., 2007; White et al., 2017). LTS metrics and ALS data can be combined to enhance large-area characterizations of forest structure (Pascual et al., 2010; Ahmed et al., 2014; Zald et al., 2014; Bolton et al., 2018). Moreover, spectral trends derived from LTS improve modeled estimates of forest structure (Pflugmacher et al., 2012) and biomass dynamics (Pflugmacher et al., 2014), and have been demonstrated to improve the characterizations of regenerating forests in temperate (Kennedy et al., 2007) and boreal forest environments (Olsson, 2009).

The temporal length and consistency of LTS are particularly well-suited to provide supporting information about forest regrowth trends. Schroeder et al. (2007) used LTS to examine the spatial and temporal variability in forest regrowth after clearcutting in western Oregon. To quantify forest regrowth, the authors used estimates of percent tree cover derived from ground plots and interpretation of aerial photographs, which were extrapolated to the LTS using date-invariant regression. The annual percent tree cover data were then grouped into four regrowth classes: little to no, slow, moderate, and fast, and different ecological regions were characterized by the prevalence of each of the regrowth classes. In addition, elevation and potential relative radiation were identified as the main drivers of the different regrowth classes. A similar approach was used by Chu et al. (2016) for assessing post-fire vegetation regrowth, whereby fractional vegetation cover was estimated to assess the return of vegetation. While these relative assessments of recovery can provide useful ecological insights regarding spatial and temporal variations in recovery, these approaches rely on the development of robust models of tree or vegetation cover, and the portability of those models through space and time. Other assessments have relied directly on the spectral metrics (e.g. Pickell et al., 2016; Frazier et al., 2015, 2018). Kennedy et al. (2012) defined an absolute and relative metric of short-term (5-year) recovery derived directly from Normalized Burn Ratio (NBR) values. Griffiths et al. (2014) assessed recovery following stand replacing disturbance in the Carpathians ecoregion using derivatives of the Disturbance Index (Healey et al., 2005). White et al. (2017) characterized both short- (5-year) and long-term (25-year) recovery from harvest and wildfire in a national assessment for Canada's forested ecosystems (~650 Mha) enabled by LTS, adapting the short-term metrics used by Kennedy et al. (2012) and a longer-term metric based on NBR (the Years to Recovery or Y2R metric) used by Pickell et al. (2016).

LTS offer new opportunities to characterize forest dynamics and in particular, provide for the characterization of recovery post-disturbance over large areas; however, there is a knowledge gap concerning how spectral measures of recovery relate to actual manifestations of forest structure (e.g. height and cover). The intensive forest management context in Finland provides a relatively controlled forest environment (i.e. even-aged, limited tree species) and a unique opportunity to explore the relationship between spectral measures of recovery derived from LTS, and actual manifestations of structure, as characterized with ALS data. The overarching goal of this research was therefore to improve our understanding of the linkages between spectral metrics of forest recovery post-harvest—as derived from LTS data—and manifestations of forest structure (height and cover) as measured from ALS data. The specific objectives of this study were threefold: (i) to apply an established image compositing and change detection approach (Composite2Change or C2C) to an area of managed forest in southern Finland and generate a spatially-explicit dataset characterizing forest change (1984–2012); (ii) to validate the detected changes using independent reference data; and (iii) to evaluate the utility and appropriateness of the Y2R spectral recovery metric for assessing the return of forest following harvest in a managed, boreal forest context. This last objective represents the unique contribution of this work: the use of ALS data to corroborate spectral metrics of forest recovery derived from LTS data.

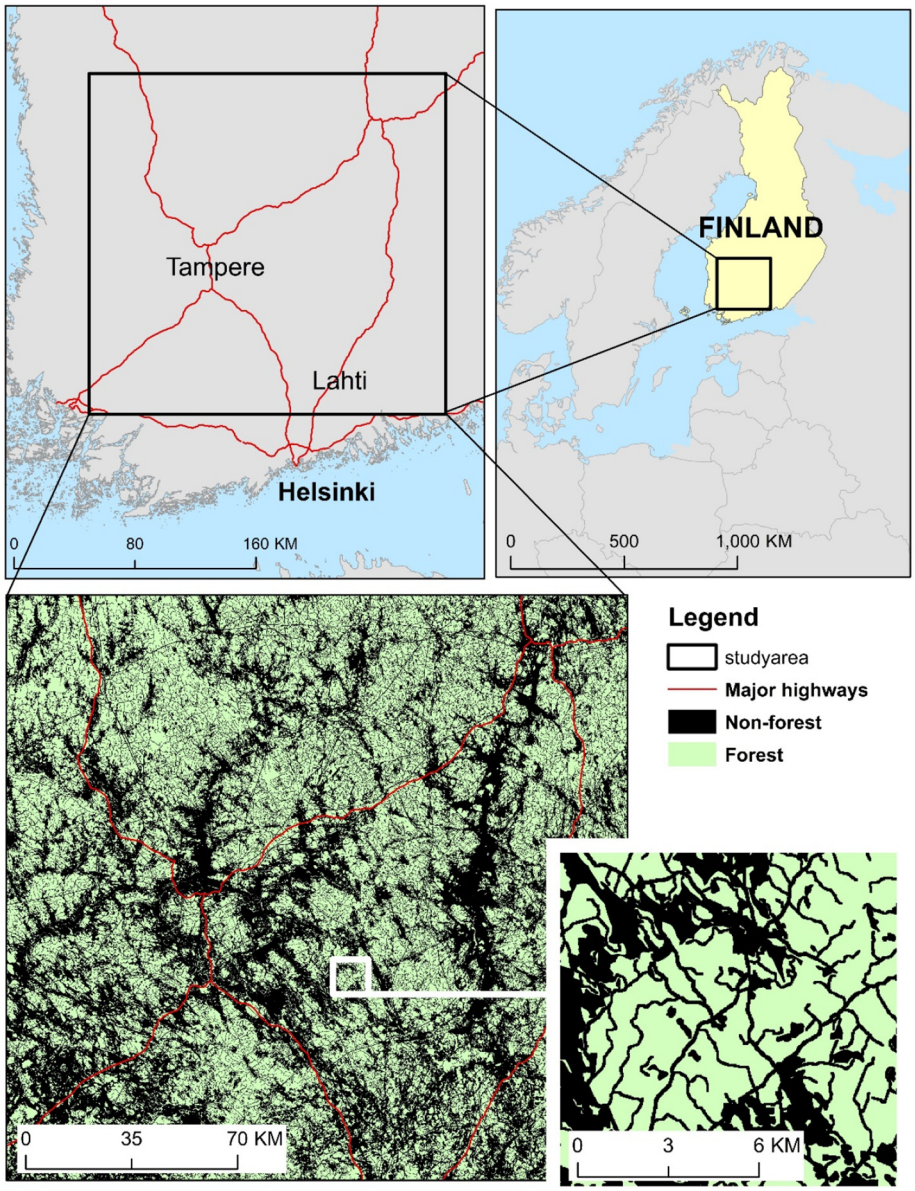

Fig. 1. Location of study area in southern Finland, with derived forest mask shown.

2. Study area

The study area, approximately 5.3 Mha in size, represents an intensively managed forest area in southern Finland and a complex landscape mosaic of agricultural, forest, and urban land use (Fig. 1). Approximately 86% of the study area belongs to southern boreal vegetation zone and the majority is considered forest (65% by area). Protected areas, such as national parks, represent 2.3% of the forested area whereas agricultural fields cover approximately 16% of the study area. Forests in this area have a mean stem volume of 146.4 $m^3\,ha^{-1}$ and the main tree species are Norway spruce (*Picea abies* L. (Karst.)) and Scots pine (*Pinus sylvestris* L.) contributing 40.2% and 38.5% of the stem volume, respectively. Approximately, 97.5% of the forest area is considered productive forest, with a growth increment of at least 1 $m^3\,ha^{-1}\,yr^{-1}$. Site type varies from herb-rich forest to barren heath forest, with the main site type being mesic heath forest covering 49.8% of the forest land within the study area (Natural Resources Institute of Finland, 2015).

3. Materials and methods

The primary objective of this work was to assess whether spectral measures of recovery were indicative of forest return, as determined by the FAO benchmarks of forest cover and height. To accomplish this required the generation and integration of numerous data sources and outputs (Fig. 2), following established methods and approaches from the literature.

3.1. Landsat data and best-available pixel composites

The study area covers eleven overlapping Landsat Worldwide Referencing System-2 (WRS-2) scenes (i.e. paths/rows). Candidate images for best-available pixel (BAP) compositing included Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) L1T format images with < 70% cloud cover. Using the C2C algorithm (Hermosilla et al., 2016), annual, cloud-free Landsat surface reflectance image composites with a 30-m spatial resolution for 1984 to 2012 were generated for the study area to represent August 1 (± 30 days). Preprocessing applied to the L1T images are described in detail in White

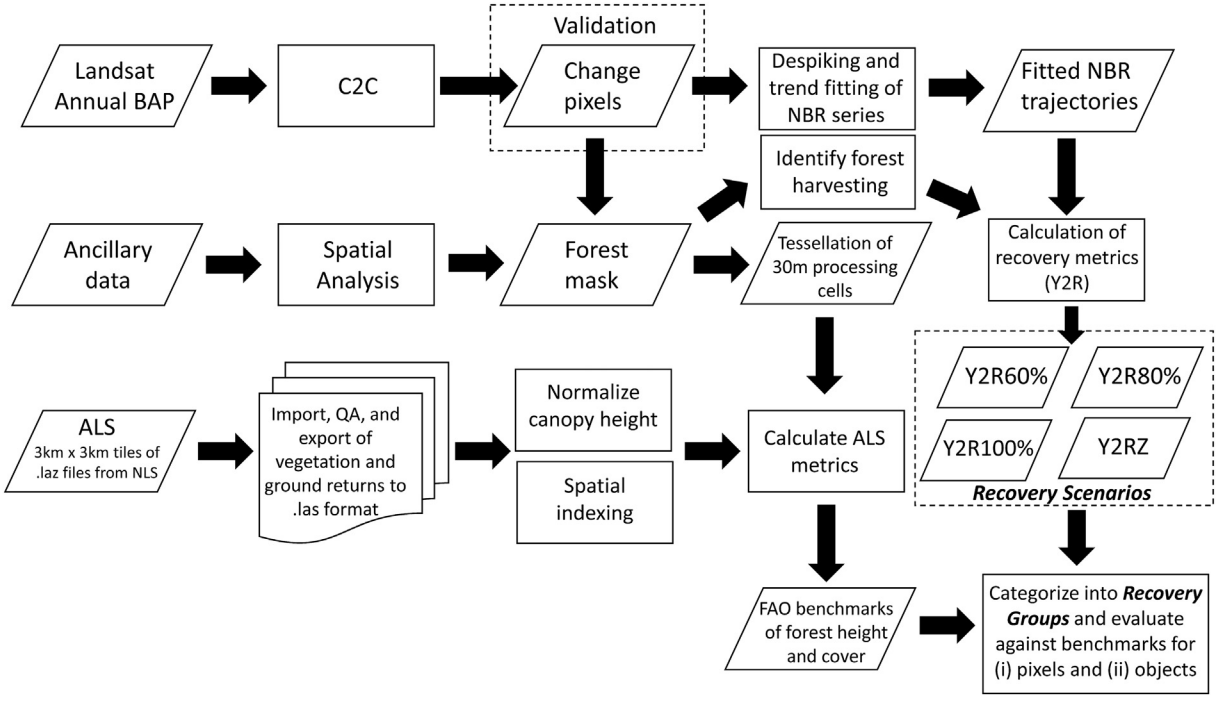

Fig. 2. An overview of the methodological workflow applied in this study.

et al. (2014). Briefly, prior to compositing, clouds and cloud shadows were detected using Fmask (Zhu and Woodcock, 2012), and the Landsat Ecosystem Disturbance Adaptive System (LEDAPS) was used to generate surface reflectance values (Schmidt et al., 2013). The compositing rules, as described in White et al. (2014) and Hermosilla et al. (2016), included distance to target day of year (August 1), distance to clouds and cloud shadows, sensor, and atmospheric opacity. For each year, pixels were scored against each of these rules, the scores for each pixel were summed, and the pixel with the highest score was considered as the best observation and was used in the annual composite. Gaps or "no data" areas exist in the composites where no best observation was available for a given pixel.

3.2. Detection of annual forest change via Landsat time series analysis

The annual BAP composites were further processed in C2C to remove noise and fill data gaps, as described by Hermosilla et al. (2015a). The annual pixel-level series of NBR values were used to identify and remove anomalous spectral values related to undetected clouds and cloud shadows or haze, thereby resulting in additional gaps in the image composites. Simultaneously in the temporal domain, the C2C algorithm identified spectral trends and detected changes in the NBR time series for each pixel using the breakpoint detection algorithm of Keogh et al. (2001). NBR is a spectral index that was first introduced by Key and Benson (2006) to map burn severity, and is calculated using Landsat TM/ETM+ bands 4 (B4; near-infrared) and 7 (B7; shortwave-infrared), as follows:

$$NBR = \frac{B4–B7}{B4 + B7} \quad (1)$$

NBR values range from −1 to 1, with positive values for pixels dominated by vegetation, and negative values for pixels dominated by bare soil (Escuin et al., 2008). As an index, the NBR was designed to take advantage of the different responses that disturbed and undisturbed areas will have in the near-infrared (NIR) and short-wave infrared (SWIR) spectral regions (Cohen and Goward, 2004). NBR is the spectral index used in the LandTrendr algorithm to characterize disturbance and recovery trends (Kennedy et al., 2010, 2012) and is among the most useful indices for disturbance detection in forests (Cohen et al., 2018).

In the C2C approach, analysis in the temporal domain was followed by contextual analysis of change pixels in the spatial domain, with the objective of identifying change events that are spatially cohesive and uniform (Hermosilla et al., 2015a). This spatial analysis also imposes a minimum mapping unit (MMU) of 0.5 ha, or approximately 5 pixels, for detected changes. As per Kennedy et al. (2012), this MMU is small enough to capture most forest management activities, while also being sufficiently large to enable validation. Following the contextual analysis, data gaps in the annual BAP composites are filled with a proxy surface reflectance value, using the spectral trend information derived from the aforementioned breakpoint detection process and piecewise linear interpolation, as described in Hermosilla et al. (2015a). The temporal breakpoints define the change intervals that bound the piecewise interpolation of the different temporal trends in a pixel's time series (e.g. monotonic, multiple breakpoint, single breakpoint).

The outputs from the combined spatial and temporal processing are gap-free, surface reflectance proxy image composites at a 30 m resolution, annual change detection information, and a series of change metrics characterizing the detected changes (including pre- and post-change conditions). A complete list of change metrics generated by C2C is provided in Hermosilla et al. (2016).

3.2.1. Identification of forest harvesting

C2C identifies a wide range of forest change, including stand replacing and non-stand replacing changes (Hermosilla et al., 2015b). As we were primarily interested in evaluating post-harvest recovery, we needed to identify changes related to forest harvesting, specifically clearcutting. We define clearcutting as an even-aged silvicultural system that removes an entire stand of trees from an area in a single harvesting operation. In Finland, forest certification currently requires that at least 10 retention trees are left for every hectare of clearcut (Forest Stewardship Council Finland, 2010, PEFC Finland, 2014). To ensure that the changes we analyzed were predominantly clearcuts, we first generated a forest mask (Figs. 1 and 2) to constrain the area included in our analysis. Information from the national base map of Finland with a scale of 1:25,000 was used to exclude agricultural fields,

non-forest land, lakes and rivers wider than 125 m, as well as urban areas. In addition, highways, railways, and power lines were excluded (with a 60-m buffer), as well as main roads (30-m buffer) and other roads (15-m buffer), with buffer widths selected to represent the typical widths of these features (which are represented as single-line features in the base map data). For changes located within this forest mask, we used the size and magnitude of the changes to aid in identifying clearcuts. Given the complex mosaic of land ownership and land use in the study area, we selected a minimum size threshold of 2 ha to distinguish those change events corresponding to forest harvesting activities as opposed to changes related to other land uses. We acquired change validation data (see Section 3.3) through the visual interpretation of high resolution imagery available from Google Earth™ and we used this visual interpretation to identify a change magnitude threshold corresponding to clearcutting (change magnitude < -0.4). C2C is capable of capturing multiple changes within the time series for a given pixel; however, change metrics will only be generated for the change with the greatest magnitude in the time series (Hermosilla et al., 2016). In stratifying by change magnitude, we ensured that other management activities, such as pre-commercial thinnings were not included as areas identified as harvest. Lastly, we applied a 30-m (equivalent to 1 Landsat pixel) buffer to the interior of these change events to account for stand edge effects, particularly with the ALS data, and excluded these pixels from our analysis. Our 2-ha minimum change event size ensured that we had sufficient pixels for analysis remaining within each event after the 1-pixel internal buffer is applied.

3.3. Validation of change and change year

A stratified random sample of points was selected to evaluate the change detection outputs following the approach described in Olofsson et al. (2014) and similar to that implemented in Hermosilla et al. (2015b, 2016). We allocated a sample size of 400 points equally to our change and no change strata. As we were also interested in characterizing the frequency with which change events were attributed to the correct year, we distributed the change samples ($n = 200$) approximately equally to each year in which changes were detected. Similar to methods outlined in Hermosilla et al. (2016), each sample was manually interpreted from the LTS as per Cohen et al. (2010), augmented by interpretation of high resolution Google Earth™ imagery, when available. The spatial support region for interpretation of each validation point was considered as the area corresponding to a 30 m Landsat pixel surrounding the point. An interpreter visually examined each sample and identified whether the pixel at the sample location was considered "changed" or "not changed", and in what year the change occurred. Results were summarized using a confusion matrix, with associated measures (e.g. producer's, user's, and overall accuracy) calculated using estimated area proportions of change and no change, as per Olofsson et al. (2014).

3.4. Airborne laser scanning (ALS) data and processing of ALS metrics

The NLS began collecting ALS data throughout Finland in 2008 to provide a new national-level digital terrain model (DTM) with a 2-m resolution. The NLS and the Finnish Forest Centre have outlined a plan to cover the entire area of Finland by the end of 2019; free and open access to the ALS data are provided by the NLS data services.[1] The ALS data have been acquired for production areas of varying sizes. The ALS data used in this study were acquired between 2008 and 2016 for production areas that ranged from 43,200 ha and 417,600 ha in size. Target parameters set by the NLS for ALS data acquisitions were a minimum point density of 0.5 points/m^2; and a point height error $<$ 15 cm. Flying altitude of all acquisition campaigns was approximately 2000 m above sea level, a scan angle of ± 20° with a footprint size of $<$ 60 cm. The utility of these data for characterizing the vertical distribution of vegetation has been demonstrated (e.g. Kankare et al., 2015; Kotivuori et al., 2016).

An overview of the ALS processing is included in Fig. 2. A suite of ALS metrics was calculated for the ALS data using LAStools (Isenburg, 2017). ALS data were downloaded from the NLS data services in 3 × 3 km tiles, with ground and non-ground points classified by the NLS. Data processing was done independently for each production area to ensure metadata concerning data origin (i.e. data provider, sensor, year of acquisition) was maintained. A rough "skypoint" classification (single points above the canopy height level) was conducted using Terrascan (Terrasolid Ltd) with a threshold value of 40 m to remove outliers. Only ground and vegetation points were then exported into new (.las) files for further processing. Using *LasHeight* and *LasIndex* –tools (Isenburg, 2017), ALS elevation data were normalized to height from the ground and a spatial index for each 3 × 3 km tile was created to facilitate processing. A tessellation of 30 × 30 m grid cells identical to the footprint and orientation of Landsat pixels was then created and used in *LasCanopy* –tool to clip and generate metrics from the normalized ALS data. Calculated metrics used in the analysis included the mean and standard deviation of ALS heights, ALS height percentiles (1%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, 99%; referenced as p01, p05, p10, etcetera), and the percentage of ALS returns within specified height intervals relative to the total number of returns (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, and 25 m; referenced as d00 for 0–1 m, d01 for 1–2 m, etcetera). Following on the recommendations of Ørka et al. (2016) and Korhonen et al. (2013), who studied the use of ALS data to characterize regenerating forests, we applied no minimum height threshold (e.g. 2 m) and used only first returns when calculating ALS metrics.

3.5. Landsat time series assessment of forest recovery

3.5.1. The Y2R metric

Vegetation recovery post-disturbance was assessed using a spectral metric of recovery derived from the fitted NBR time series data as per White et al. (2017). The Y2R metric is designed to characterize the longer-term, sustained regeneration of forests at a site. Y2R is determined using trend-fitted NBR values from our time-series analysis. Prior to trend fitting, we applied a de-spiking approach similar to that of Kennedy et al. (2010) and Bolton et al. (2015), where noisy observations are detected by examining them in relation to their previous and subsequent spectral values in the time series (Hermosilla et al., 2015a). As noted by Schroeder et al. (2007), year-to-year differences, resulting from phenology or atmospheric effects such as haze, will be minimized by a fitted trajectory curve. In previous research (Pickell et al., 2016; White et al., 2017), Y2R was defined as the number of years it takes for a pixel to return to 80% of its pre-disturbance value, with the latter defined as the average NBR value of the 2 years prior to disturbance (y-2 and y-1):

$$NBR_{pre} = \frac{NBR_{y-2} + NBR_{y-1}}{2} \quad (2)$$

As noted earlier, the NBR captures different spectral responses in the NIR and SWIR spectral regions. Harvesting will cause a marked increase in reflectance in both the SWIR and NIR bands. The initial 20–30 years of stand development post-harvest are then characterized by an increase in the proportion of sunlit crowns and a decrease in the proportion of sunlit and shaded ground (Li and Strahler, 1985; Nilson and Peterson, 1994). SWIR, in particular, is sensitive to shadowing and vegetation density (Horler and Ahern, 1986) and thus as canopy complexity and shadowing increases with age, there is a decrease in reflectance in the SWIR (Peterson and Nilson, 1993). Uncertainty in monitoring the successional trajectories of forest post-disturbance is complicated by topography, atmosphere, phenology, and sun and view

[1] https://tiedostopalvelu.maanmittauslaitos.fi/tp/kartta?lang=en

angles (Song and Woodcock, 2003). While it is possible to model successional reflectance trajectories (e.g. Peterson and Nilson, 1993), real successional trajectories are noisy and often non-linear, being strongly influenced by background or understory conditions and topography (Song et al., 2002). The influence of understory on Landsat reflectance in Finnish boreal forests was found to vary by stand development and site fertility classes (Kuusinen et al., 2015), and were greatest in the NIR, similar to the findings of Miller et al. (1997) in Canadian boreal forests. As noted by Song et al. (2002), the impacts of understory on reflectance successional trajectories can be reduced by narrowing the image acquisition window, as was done in this study for the BAP composites, and by making use of multi-temporal imagery. Cohen et al. (2018) published a Disturbance Signal to Noise Ratio (DSNR) metric. NBR was found to have the highest median DSNR of all the spectral features tested, while the NIR band had the lowest median DSNR value. Moreover, the NIR was found to "be the single most important complementary spectral band, in spite of its tendency to exhibit low DSNR values."

As one of our objectives in this study was to determine the suitability of the Y2R metric for managed forests relative to the manifestation of forest structure, we evaluated alternative definitions of the Y2R metric. Specifically, four spectral recovery scenarios were considered (Fig. 3): recovery when NBR value was: 60% of NBR_{pre} (Y2R60%); 80% of NBR_{pre} (Y2R80%); 100% of NBR_{pre} (Y2R100%); and > the one-sided critical z-value ($\alpha = 0.05$, -1.645, Y2RZ). For the Y2RZ scenario, we standardized the NBR values for each pixel's time series to z-scores using the following equation:

$$Z_{NBR_i} = \frac{NBR_i - \mu_{NBRpre}}{\sigma_{NBRpre}} \quad (3)$$

For this scenario, Y2R is defined probabilistically and is the year in which the NBR value is no longer significantly different from NBR_{pre} (i.e. when the z-value exceeds the one-sided critical value of -1.645). The 80% threshold is based on the work of Pickell et al. (2016), who compared the performance of several spectral indicators to measure years to recovery. In turn, Pickell et al. (2016) had based their approach on earlier work that had similarly used relative metrics of spectral recovery (e.g. Schroeder et al., 2011; Kennedy et al., 2012, and Baumann et al., 2012). This 80% threshold ensures that most of the pre-disturbance spectral condition is met, but allows some latitude in spectral development, because attaining the 80% threshold does not necessarily indicate a return to similar pre-disturbance forest conditions. However, the suitability of this threshold in different forest environments with different disturbance characteristics is not known, so other threshold values were tested in this study. In addition, the use of a data driven, probabilistic assessment of recovery, which in practical applications would negate the requirement to establish a threshold, was also applied.

3.5.2. Pixel-level analysis of Y2R

The Y2R metrics were generated for each pixel. The number of years required to satisfy each of these aforementioned recovery scenarios were then grouped into five possible recovery groups: (1) < 10 years, (2) 10–13 years, (3) 14–17 years, (4) > 17 years, and (5) not recovered. In a previous study using ground plot data, we found that on average, the benchmarks of cover and height used herein could be achieved at the plot level in boreal forests within 10 years, but that this varied by location, tree species composition, site characteristics, and disturbance type (Bartels et al., 2016). Given the more northern latitude of the study area described herein, any sub-groupings < 10 years would not be meaningful in characterizing a return of forest structure at the pixel level. The remaining Y2R (10–21 years) were divided into equal 4-year intervals. We compared mean values of our derived ALS and LTS metrics using a one-way ANOVA to identify statistically significant differences between these five groups for each of the recovery scenarios. For a subset of pixels, the year of ALS measurement (which equates to time since disturbance) corresponded to the year in which the pixel was considered recovered by the Y2R metric; however, the sample size was relatively small ($n < 300$ pixels) for all recovery scenarios except for Y2R100% ($n = 2064$ pixels). Therefore, as we had a direct temporal correspondence between a spectral measure of recovery (from LTS) and a structural measure of recovery (from the ALS), for the Y2R100% scenario we also compared ALS and LTS metrics using a one-way ANOVA and tested for significant differences between the different Y2R.

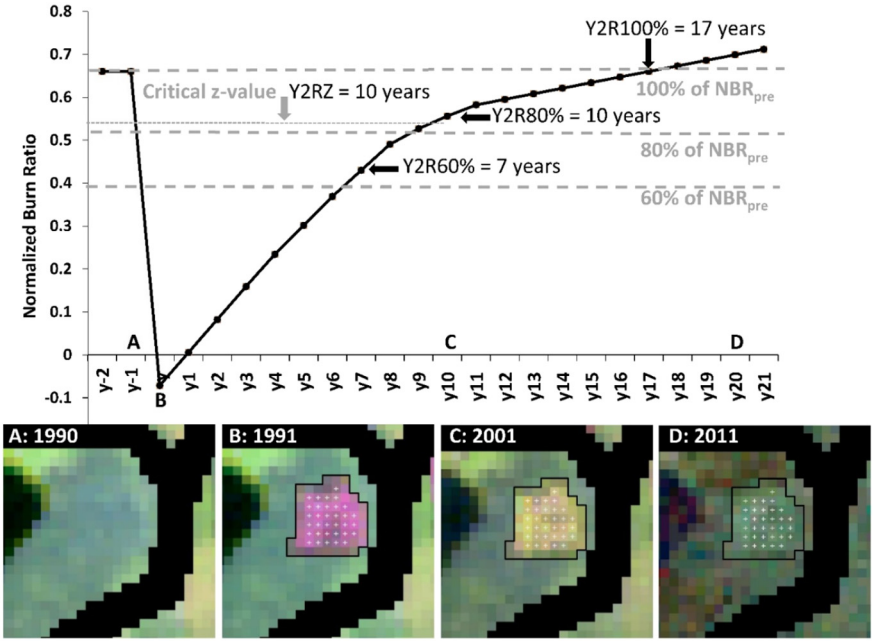

Fig. 3. An NBR trajectory for a sample change event from 1991. This trajectory represents the average of the pixels within the object (source pixels indicated with white cross). Note the 30-m buffer applied to the interior of the change event to reduce edge effects. Image in panels A–D is Landsat bands 5 (R), 4(G), 3(B).

3.5.3. Forest structural benchmarks of recovery derived from ALS data

Bartels et al. (2016) suggested the use of benchmark thresholds of canopy cover (> 10%) and height (> 5 m) that could be used to assess forest recovery, and these thresholds correspond to the minimum values required to satisfy the FAO's definition of forest (FAO, 2012). The cover benchmark was evaluated against the sum of the ALS return densities > 2 m (i.e. d02-d12), and the height benchmark was assessed using 99th percentile of height. We applied the minimum height and cover thresholds to each recovery scenario and analyzed the percentage of pixels that met these thresholds overall and within each of our recovery groups.

3.5.4. Event-level analysis of Y2R

As the information needs of forest monitoring associated with large-area indicators of forest recovery are primarily at the stand level, we summarized the Y2R and ALS metrics at the change event level (Hermosilla et al., 2016). We calculated the arithmetic mean of the Y2R and ALS metrics of the pixels found within each change event, for each recovery scenario. We then applied the same benchmarks of recovery for cover and height and analyzed the percentage of change events that met the benchmarks, by recovery group and scenario.

4. Results

4.1. Landsat data compositing, change detection and validation

A total of 554 images were used to generate annual BAP composites (1984–2012), with an average of 20 images used per year. Some years had no available imagery with < 70% cloud (Fig. 4). The availability of Landsat data impacts the detection of change. The lack of data availability prior to 1990 precluded change detection for this period. Moreover, as no suitable Landsat data were available in 1991, 1992, or 1993, changes that occurred in these years were assigned a greatest change year (GCY) of 1991. Likewise, changes that occurred in 1995, 1996, and 1997 were assigned a GCY of 1995. This accounts for the greater amount of change attributed to 1991 and 1995, relative to other years (Fig. 4). A total of 117,901 change events were detected within our forest mask (Fig. 1) between 1991 and 2011, with a mean event size of 1.85 ha. The overall accuracy of the change detection was 89%, with errors of omission for the change class greater than errors of commission (Table 1). Of the 177 reference samples that were correctly identified as change by C2C, 162 or 85.88% were detected in the correct year and ~94% were within ± 1 year (Table 2).

Of the total 117,901 change events detected, 33,164 change events were identified as forest harvesting (with this subset having a mean event size = 3.83 ha). To ensure the maximum recovery period, we only included change events that occurred in 1991 (n = 3697) in our analyses. These events yielded 40,365 pixels for our analysis of the Y2R metrics.

4.2. Landsat time series assessment of forest recovery

The relative distributions of change pixels within the four recovery groups, by recovery scenario, are summarized in Table 3. For the Y2R60% and Y2R80% scenarios, all pixels included in our analysis were considered recovered by the end of the time series (i.e. 2012). The Y2R80% and Y2RZ scenarios had similar distributions of pixels among the five recovery groups, with Y2RZ having 1.67% pixels identified as not recovered. In contrast, distributions for the Y2R60% and Y2R100% scenarios were markedly different: whereas 93.96% of pixels recovered in < 10 years for the Y2R60% scenario, only 4.80% of the pixels in the Y2R100% scenario recovered in < 10 years. Moreover, 30.99% of pixels were considered as *not* recovered in the Y2R100% scenario (Table 3).

We found significant differences for all ALS metrics among the recovery groups, with the exceptions of the density metrics associated with the upper canopy and the lowest height percentiles (d12 and p01; results not shown). Given our large sample sizes, the significance of these differences was expected. The metrics with the maximum absolute difference between recovery groups for mean metric values were d00 and p99 (Table 4). Generally, differences in ALS metrics between groups decreased with an increase in the spectral threshold used to define recovery. For example, the maximum difference among recovery groups for p99 was 4.68 m for the Y2R60% scenario, compared to only 1.03 m for the Y2R100% scenario (Table 4). Similarly, the maximum difference in the mean value of d00 was 28.61% for Y2R60%, compared

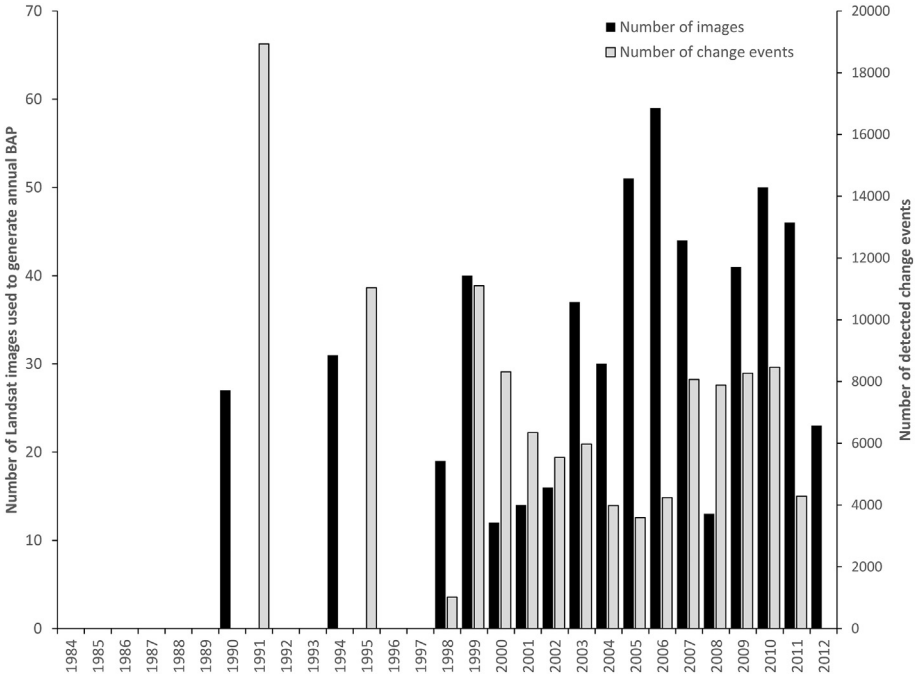

Fig. 4. Temporal distribution of Landsat TM and ETM+ images used to generate annual best-available-pixel (BAP) composites (1984–2012), and number of change events detected within the area of the forest mask (n = 117,091).

Table 1
Validation results of C2C change/no change detection. Accuracy measures were calculated using estimated area proportions of change and no change as per Olofsson et al. (2014).

	Reference data					
	Change	No change	Total	User's accuracy	Producer's accuracy	Overall accuracy
Change	0.152	0.013	0.164	0.92 ± 0.03	0.62 ± 0.06	0.89 ± 0.03
No change	0.092	0.743	0.836	0.89 ± 0.04	0.98 ± 0.02	
Total	0.244	0.756	1.000			

Table 2
Attribution of changes to the correct year.

Year of change	Percentage of samples
Changes detected in the correct year	85.88
Changes detected within ± 1 year	7.91
Changes detected within ± 2 years	3.95
Changes detected within ± > 2 years	2.26

Table 3
Percentage of pixels within each recovery group for each recovery scenario (n = 40,365 pixels). Recovery scenarios are defined according to the threshold applied to the fitted Normalized Burn Ratio (NBR) time series: 60%, 80%, or 100% from the pre-disturbance value of NBR, and Y2RZ, which is defined as the year in which the NBR value is no longer significantly different from NBR_{pre}. Note a very small sample size for Y2R60% recovery group 4 (n = 2 pixels).

		Recovery scenarios			
Recovery group	Group #	Y2R60%	Y2R80%	Y2R100%	Y2RZ
< 10 years	1	93.96	55.32	4.80	50.14
10–13 years	2	5.72	31.60	12.27	28.41
14–17 years	3	0.32	10.63	20.03	14.62
> 17 years	4	0.00	2.45	31.91	5.17
Not recovered	5	0.00	0.00	30.99	1.67

to 4.275% for Y2R100%.

Generally, pixels that were considered as spectrally recovered within 10 or fewer years (recovery group 1) had larger median values for ALS height percentiles p75, p90, p95, p99, relative to pixels that took longer to recover (Fig. 5). Hence pixels that recovered rapidly were taller on average at the time of ALS measurement, than pixels that took longer to recover. Fig. 5 indicates a decreasing trend in the median values of these height percentiles for the different recovery groups considered. For example, for the Y2R60% scenario, p99 had a median value of 10.03 m for pixels recovered in < 10 years, compared to a median of 7.54 m for pixels that recovered in 14–17 years (recovery group 3). Similar to the trend shown in Table 4, the difference between recovery groups decreases with an increase in the spectral threshold used to define recovery; the Y2R100% scenario, pixels in recovery group 1 had a median p99 of 10.98 m compared to a median of 9.74 m for pixels in recovery group 3 (Fig. 5).

Trends were similar for the ALS density metrics (Fig. 6). As indicated in Table 4, the largest differences between recovery groups was found for density in the 0–1 m stratum (d00), and this difference is greatest for Y2R60%, and smallest for Y2R100%. Pixels that had longer recovery times also had larger median values for d00. For example, in the Y2R60% scenario, the median d00 value for pixels that recovered in < 10 years was 53.00%, compared to 68.95% in pixels that recovered in 14–17 years. By comparison, in the Y2R100% scenario, the median value for pixels in recovery group 1 was 50.71% compared to 53.7% for recovery group 4 (Fig. 6).

The differences between the Y2R60% and Y2R100% scenarios, and the similarities between the Y2R80% and Y2RZ scenarios are indicated in Fig. 7. Generally, as the spectral threshold for recovery increased (i.e.

Table 4
Maximum differences between recovery groups for the mean values of ALS metrics, by recovery scenario.

	Maximum difference between mean values for recovery groups			
	Y2R60%	Y2R80%	Y2R100%	Y2RZ
ALS metric	4 groups	4 groups	5 groups	5 groups
Density (%)				
d00	28.61	11.43	4.28	8.40
d01	6.49	1.07	0.31	0.32
d02	2.20	1.10	0.69	0.69
d03	3.37	1.27	1.14	0.88
d04	4.62	1.58	1.24	1.05
d05	5.29	2.15	1.03	1.44
d06	4.20	2.41	1.03	1.71
d07	3.28	2.16	1.03	1.53
d08	2.45	1.69	0.85	1.23
d09	1.77	1.25	0.66	0.94
d10	3.83	2.50	1.69	1.90
d15	1.51	0.51	0.45	0.29
d20	1.24	0.11	0.09	0.04
Height (m)				
p01	0.00	0.00	0.00	0.00
p05	0.01	0.02	0.01	0.01
p10	0.05	0.06	0.03	0.03
p25	1.91	1.07	0.46	0.79
p50	4.01	1.79	0.72	1.29
p75	0.42	0.31	0.15	0.25
p90	4.29	2.04	0.96	1.46
p95	3.73	1.58	0.93	1.25
p99	4.69	1.92	1.03	1.41
avg	1.53	1.04	0.40	0.76
std	1.02	0.62	0.36	0.45

from 60% of NBR_{pre} to 100% of NBR_{pre}), the spread of the distribution of Y2R values also increased, indicative of a greater number of years required to attain the spectral threshold value.

For the Y2R100% scenario, we analyzed a subset of pixels where the year of spectral recovery was the same as the year of ALS acquisition. Within this subset, the Y2R ranged from 17 to 22 years. Significant differences were found between the mean values for these Y2R groups for a subset of ALS metrics; however, the magnitude of the differences in mean values for ALS metrics between years, within this single recovery group (recovery group 4), was markedly lower (Table 5) than between recovery groups (Table 4). Overall, and in contrast to Figs. 5 and 6, no trends are evident in the median values for these ALS metrics between years (Fig. 8).

We applied the FAO thresholds of > 5 m for height (Fig. 5) and > 10% for cover (Fig. 9) and determined the percentage of pixels that met or exceeded the thresholds for *both* cover and height, by recovery group, for each recovery scenario (Table 6).

The relative distribution of pixels among benchmark categories *within* recovery groups was relatively consistent across all four scenarios: benchmarks of height were more commonly achieved than benchmarks of cover, while achieving benchmarks of *both* cover and height was most common overall. Recalling that all pixels were considered recovered in the Y2R60% and Y2R80% scenarios, there was no

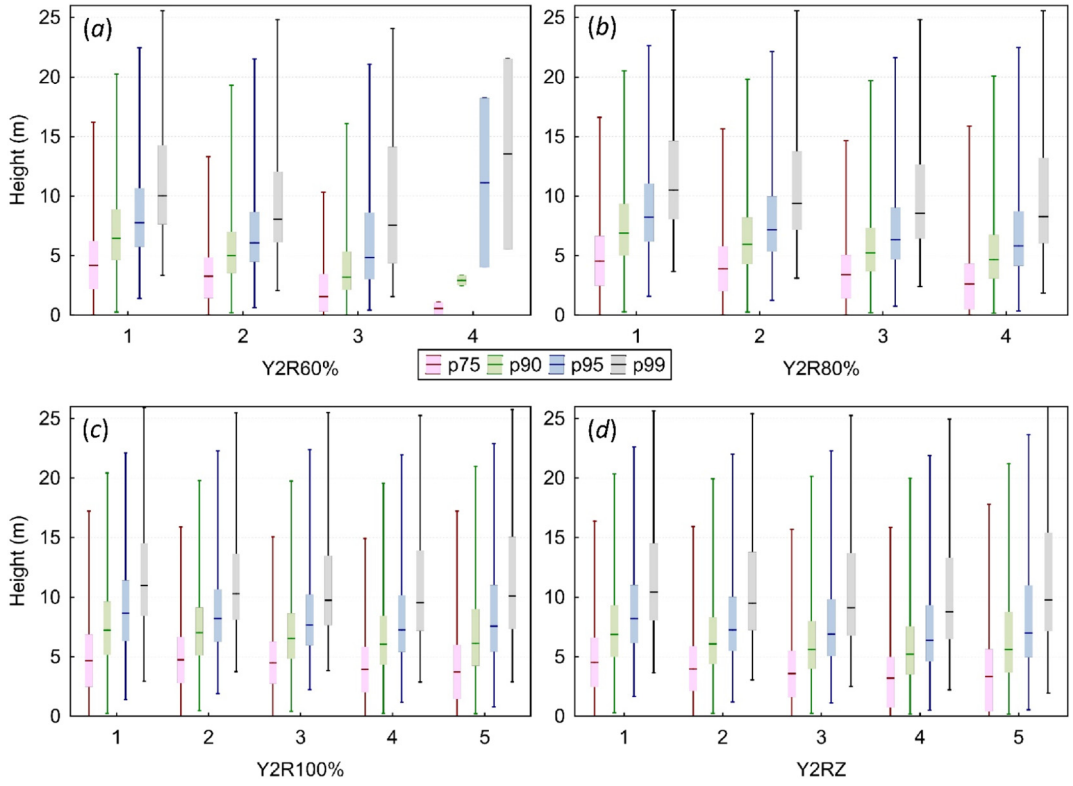

Fig. 5. ALS height percentiles p75, p90, p95, and p99 by recovery groups on x-axis (as defined in Table 3) for each recovery scenario: Y2R60% (a), Y2R80% (b), Y2R100% (c), and Y2RZ (d). Box plots show the median and first and third quartiles; whiskers represent the 1st and 99th percentiles of the ALS height percentiles.

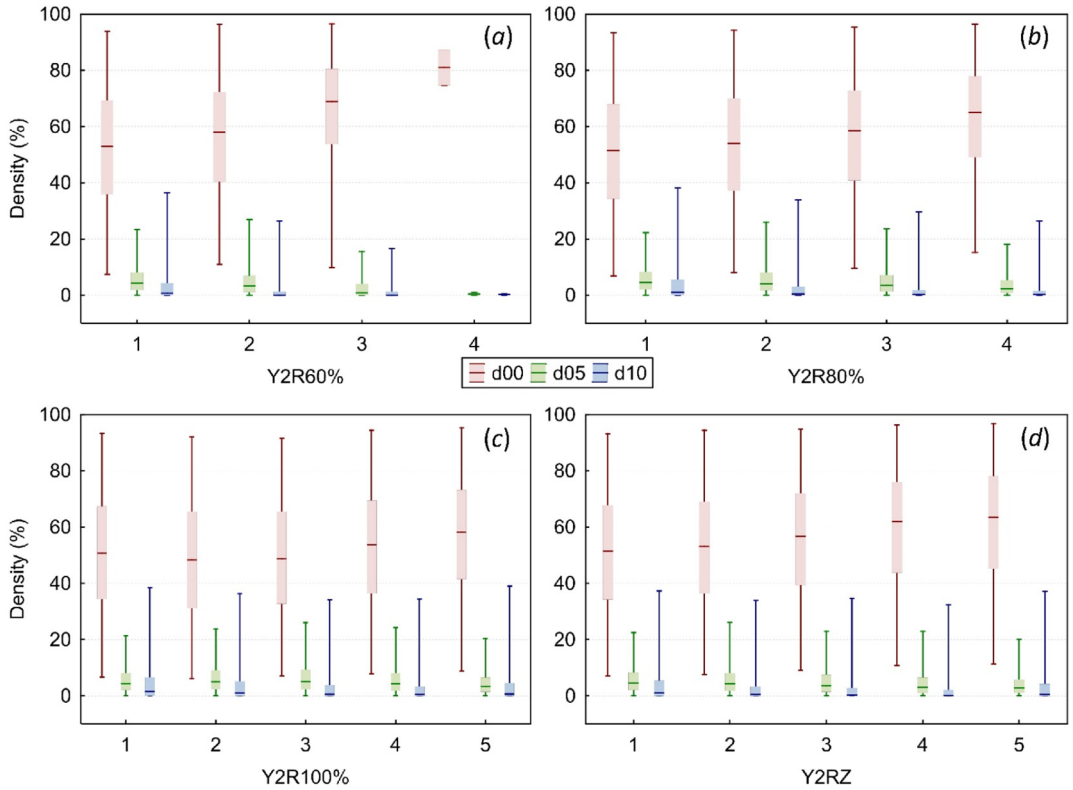

Fig. 6. ALS density metrics d00, d05, and d10 by recovery groups on the x-axis (as defined in Table 3) for each recovery scenario: Y2R60% (a), Y2R80% (b), Y2R100% (c), and Y2RZ (d). Box plots show the median and first and third quartiles; whiskers represent the 1st and 99th percentiles of the ALS density metrics.

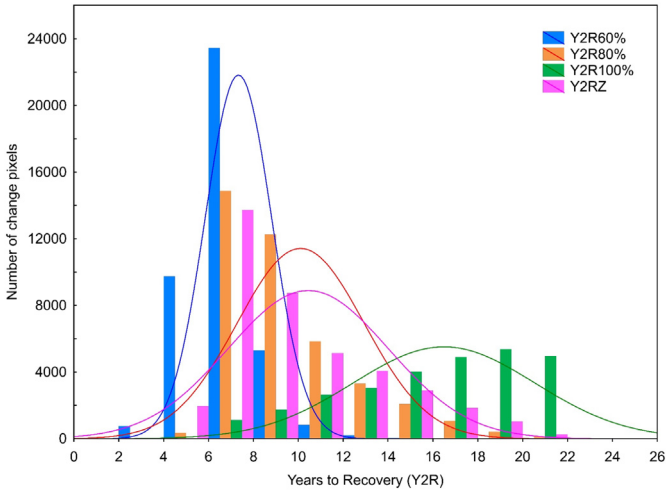

Fig. 7. Pixel-level distribution of years to recovery (Y2R), by recovery scenario (bars), with fitted normal distribution (lines).

Table 5
Maximum between-year differences (17–22 years) for selected ALS metrics.

Metric	Maximum difference between years
Density (%)	
d00	2.77
Height (m)	
p75	0.37
p90	0.58
p95	0.69
p99	0.81

difference between these two scenarios in terms of the total percentage of recovered pixels that had achieved both benchmarks of cover and height (88.88%). However, there were differences for these two scenarios in terms of the percentage of pixels within each of the recovery groups that had achieved the benchmark targets (Fig. 10). For example, in recovery groups 2 and 3, the Y2R60% scenario had a greater percentage of pixels that were indicated as spectrally recovered but that had not achieved the benchmarks of cover and height (Table 6). For the Y2R100% threshold scenario, the percentage of pixels that required > 17 years to recover and that attained the benchmarks of both cover and height was 27.92%. However, 26.84% of pixels that were considered as not recovered using the Y2R100% threshold had also achieved both benchmarks for height and canopy cover. The results for the Y2RZ scenario were very similar to those of Y2R80% (Fig. 10) with only a small percentage of non-recovered pixels (1.34%) that had achieved the benchmarks of both cover and height. Pixel-level trends observed for benchmarks of cover and height were also evident when Y2R metrics were averaged to the level of the change event (Fig. 10).

5. Discussion

Forest recovery is a long-term process, and it can be challenging to define the point at which a stand can be considered recovered (Frolking et al., 2009; Bartels et al., 2016). Depending on the information need, the application of benchmark targets of forest structure provides an objective assessment of recovery that can be readily measured with ALS data. In our analysis, the application of the FAO benchmarks of forest cover (> 10%) and height (> 5 m) provided insights on the appropriateness of the various spectral recovery thresholds considered in our analysis. For example, when only one benchmark was achieved, the height benchmark was more frequently attained relative to the cover benchmark (Table 6). A similar phenomenon was noted by Bolton et al. (2017) in their analysis of post-fire regeneration with ALS data: height growth post-fire was more rapid than the return of canopy cover, and this was thought to be a result of residual structures (i.e. snags or surviving trees) remaining at the site. In a meta-analysis of ground plot data, Bartels et al. (2016) similarly found that the FAO height benchmarks were attained more rapidly than cover following stand replacing disturbances (i.e. wildfire, clearcut) in Canada's boreal forests.

Our results confirm the utility and appropriateness of the Y2R metric as an indicator of forest recovery. Regardless of the Y2R threshold used herein, pixels that recovered rapidly had larger median values for the upper ALS height percentiles (Fig. 5). Conversely, pixels that took longer to recover had smaller median values for height percentiles. Differences between recovery groups decreased with an increasing spectral threshold for recovery, indicating that with the longer recovery times required to achieve a higher target threshold, forest structural conditions in terms of height and cover, begin to converge. Although the Y2R60% and Y2R80% scenarios were comparable in terms of the overall percentage of recovered pixels that had attained the benchmark values for both cover and height (88.88%; Table 6), the within-recovery group comparisons between these two scenarios revealed that only about half of pixels that the Y2R60% scenario indicated as recovered in 14–17 years had achieved the benchmarks for both cover and height. By comparison, approximately 80% of Y2R80% pixels recovered in 14–17 years had achieved both benchmarks for height and canopy cover (Table 6). Likewise, under the Y2R60% scenario, 2.5 times as many pixels that were spectrally recovered in < 10 years had not achieved the FAO benchmarks, compared to the Y2R80% scenario. This suggests that the Y2R60% may be an overly optimistic threshold for assessing the return of forests, with spectral recovery attained rapidly before forest structure, in terms of height and cover, have returned at a given location. Likewise, the Y2R100% may be an excessively pessimistic threshold, with > 86% of pixels that were considered as not yet recovered by the end of the time series having already attained the benchmark values of cover and height. As noted by Pickell et al. (2016), it may be unrealistic to assume that a pixel will return to 100% of its pre-disturbance value within the temporal window of the LTS, particularly if the forest was mature prior to disturbance (as in our study) or there was a change in dominant species or management practices post-disturbance that would alter the density and configuration of the canopy, and thereby also the reflectance properties of the stand. Real successional reflectance trajectories are noisy (Song et al., 2002) and are driven by a myriad of factors.

Our results indicated that of the recovery scenarios evaluated, the Y2R80% threshold provided the most realistic assessment of forest recovery post-harvest in the intensively managed forests of southern Finland: all pixels were considered recovered within the time period assessed, and sites that recovered rapidly (< 10 years) had forest structural properties at the time of ALS measurement that were distinct from sites that took longer (i.e. > 17 years) to recover. For Y2R80%, 88.88% of recovered pixels met the FAO benchmarks for *both* cover and height (Table 6). Moreover, under the Y2R80% scenario, there were fewer pixels that were spectrally recovered but not structurally recovered, and conversely, no pixels that were structurally recovered but not spectrally recovered. False positives were more common for the Y2R60% scenario, while false negatives accounted for 26.84% of pixels under the Y2R100% scenario.

The results of the Y2RZ scenario were similar to those of the Y2R80%, with the Y2RZ having a small percentage of pixels identified as not recovered that had actually attained the FAO benchmarks for cover and height. However, unlike the Y2RZ, the Y2R80% scenario is straightforward to apply and requires no further computation once the NBR trajectories are processed. However, as our results indicate, spectral thresholds can dramatically alter the characterization of recovery: a 60% threshold indicated that 93.96% of pixels recovered in < 10 years, whereas a 100% threshold indicated that only 4.80% of pixels

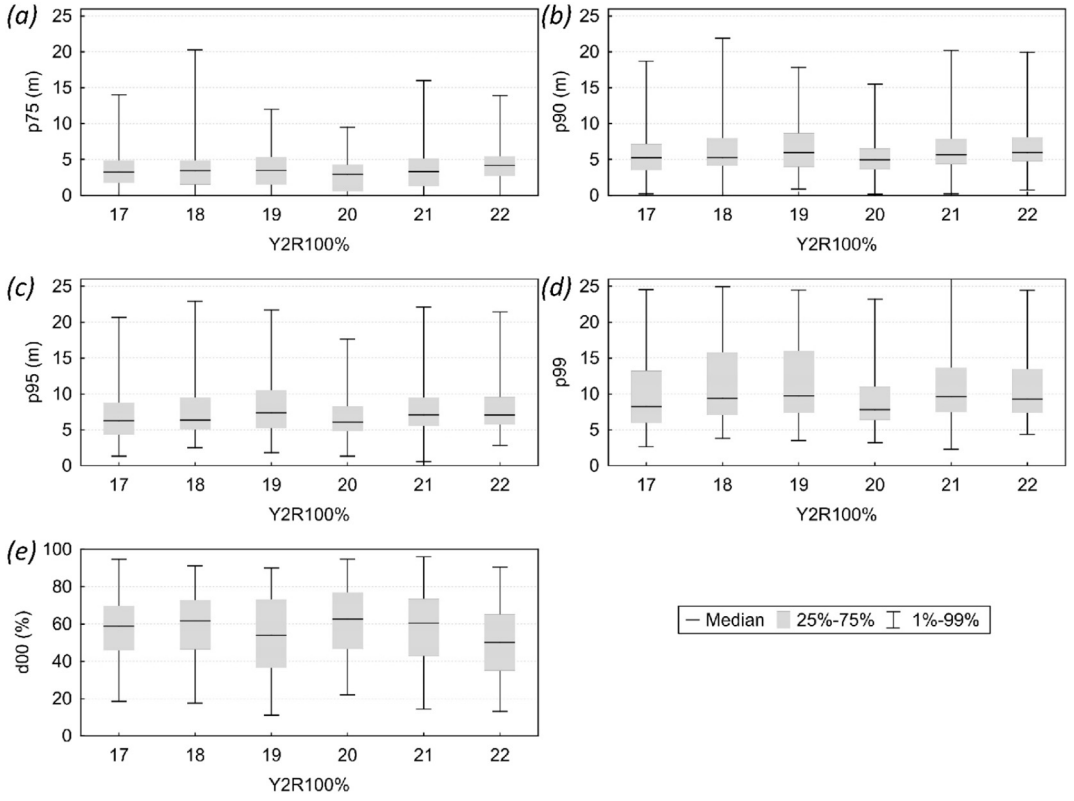

Fig. 8. LTS and ALS metrics by years to recovery Y2R (17–22; x-axis) for the Y2R100% scenario where the year of ALS acquisition equals the year of recovery (N = 2064 pixels). Box plots show the median, first and third quartile; whiskers represent the 1st and 99th percentiles for height percentiles p75 (a), p90 (b), p95 (c), p99 (d), d00 (e).

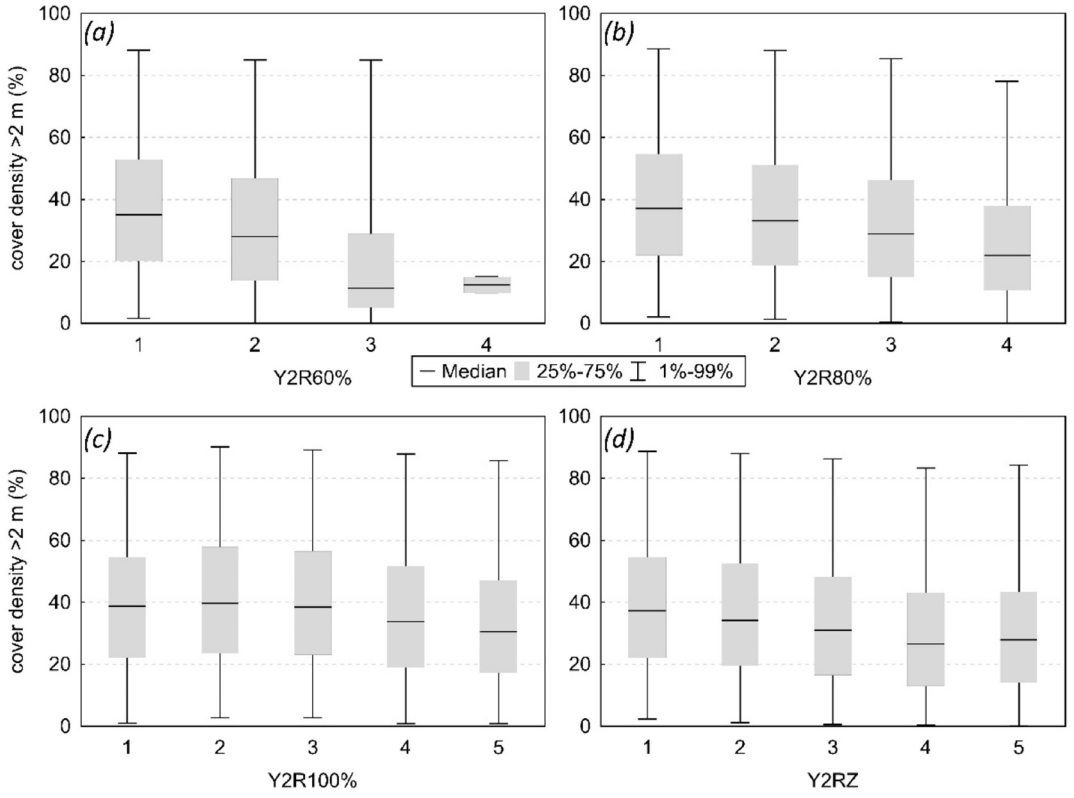

Fig. 9. Cover density > 2 m from the ALS metrics by recovery groups on the x-axis (as defined in Table 3) for each recovery scenario: Y2R60% (a), Y2R80% (b), Y2R100% (c), and Y2RZ (d). Box plots show the median and first and third quartiles; whiskers represent the 1st and 99th percentiles.

Table 6
Percentage of pixels within each of the recovery groups that achieved the benchmark values for canopy cover (> 10%), height (> 5 m), both cover and height, or neither benchmark, by recovery scenario.

Benchmarks of recovery	Recovered					Not recovered
	< 10 yrs	10–13 yrs	14–17 yrs	> 17 yrs	Total	
Y2R60%						
Cover	1.29	0.22	0.00	0.00	1.52	0.00
Height	5.08	0.41	0.06	0.00	5.55	0.00
Cover & height	84.20	4.51	0.17	0.00	88.88	0.00
None	3.38	0.58	0.09	0.00	4.05	0.00
Total (see Table 3)	93.96	5.72	0.32	0.00		0.00
Y2R80%						
Cover	0.47	0.69	0.31	0.05	1.52	0.00
Height	2.71	1.81	0.75	0.28	5.55	0.00
Cover & height	50.75	27.58	8.73	1.82	88.88	0.00
None	1.39	1.52	0.83	0.30	4.05	0.00
Total (see Table 3)	55.32	31.60	10.63	2.45		0.00
Y2R100%						
Cover	0.05	0.14	0.32	0.58	1.09	0.42
Height	0.20	0.50	0.82	1.79	3.30	2.25
Cover & height	4.43	11.34	18.35	27.92	62.04	26.84
None	0.13	0.29	0.54	1.62	2.58	1.48
Total (see Table 3)	4.80	12.27	20.03	31.91		30.99
Y2RZ						
Cover	0.47	0.57	0.34	0.09	1.48	0.04
Height	2.42	1.56	0.90	0.49	5.37	0.18
Cover & height	45.92	25.00	12.48	4.14	87.54	1.34
None	1.33	1.27	0.90	0.44	3.94	0.11
Total (see Table 3)	50.14	28.41	14.62	5.17		1.67

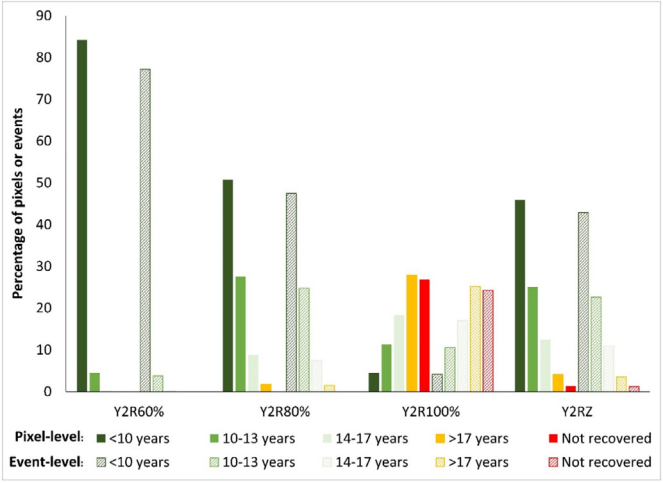

Fig. 10. Percentage of pixels and change events within each recovery group that achieved benchmarks of canopy cover and height, by recovery scenario.

recovered in < 10 years. Moreover, a 60% threshold indicated that all pixels had recovered, whereas a 100% threshold indicated that 30.99% of pixels had not recovered. Further research is necessary to determine the appropriateness of the probabilistic approach to defining spectral recovery across different forest environments; however, one of the key advantages of the Y2RZ scenario for operational programs would be removing the necessity to identify an optimal threshold for a given forest environment or disturbance context.

The application of the C2C approach in southern Finland demonstrates the portability of the approach to other regions. Of note, data availability was markedly lower for this study area in the earlier part of the time series in the pre-ETM+ era (i.e. pre 1999; Saarinen et al., 2018) compared to previous implementations of C2C (White and Wulder, 2013; Hermosilla et al., 2016). This reflects the less systematic acquisition of Landsat data globally prior to the launch of Landsat 7, which impacts regional Landsat archive data holdings differently (Wulder et al., 2012). The Landsat Global Archive Consolidation initiative seeks to incorporate all previously acquired Landsat data that is currently archived regionally by groups such as ESA, into the central global USGS archive (Wulder et al., 2016), which may result in additional data for our study site. However, logistical and programmatic challenges during the commercialization phase of the Landsat program (1985–2001) reveal insights into the decline in Landsat data acquisition during this time (Goward et al., 2017). The increased frequency of data gaps in the BAP composite influences the attribution of change to the exact year in which the change occurred, and likely would also impact the detection of non-stand replacing changes (Hermosilla et al., 2015b), as detection accuracy has been known to diminish over time for non-stand replacing changes such as insect damage (Wulder et al., 2005). In this study however, our targets were clearcuts in boreal forests, which were readily distinguishable on the landscape 3–4 years post-disturbance. Another notable difference for this study site is the intensity and mixture of land uses found. Compared to other pan-boreal environments in Russia, Canada, and Alaska where large wildfires are the dominant stand-replacing change (de Groot et al., 2013), and moreover, where changes in the forested areas are typically remote from areas of other land use (White et al., 2017), change events in our study site were relatively small (1.85 ha overall, 3.83 ha for harvest) and embedded within a complex mosaic of different land uses.

The overall accuracy of the change detection was 89% and is comparable to that reported by other studies using LTS data in similar mixed land use European contexts (e.g., Griffiths et al., 2014). Forest practices in Finland have evolved over time, with retention trees within harvested areas becoming increasingly common from the mid-1990s (Gustafsson et al., 2010). Retention trees may confound assessment of recovery at the event level (given the spatial resolution used in this analysis—30 m), thereby emphasizing the importance of the initial per-pixel analysis of recovery, and the use of Landsat pixels as a spatial reference for ALS metric generation, as demonstrated in this study. By considering only those changes that occurred in 1991, we examined those pixels that had the longest possible recovery times afforded by

our time series; however both short- and long-term recovery metrics can provide meaningful insights regarding vegetation return post-disturbance (Kennedy et al., 2012; White et al., 2017).

Landsat time series data allow for a more holistic assessment of forest dynamics: it is now possible to characterize both disturbance and recovery on an annual basis, which was previously not feasible with bi-temporal or multi-temporal epochal image data. In contrast to ground-based surveys of forest recovery, which are often spatially and temporally constrained (Bartels et al., 2016), the opening of the Landsat archive has enabled assessments of recovery that are spatially exhaustive and retrospective—providing important baseline data for forest monitoring in an era of climate change.

6. Conclusions

LTS-derived measures of spectral recovery are useful for understanding regional (Schroeder et al., 2007; Kennedy et al., 2012), temporal (Frazier et al., 2018), and agent-based (White et al., 2017) variations in forest recovery over large areas. These remotely-sensed assessments can augment ground-based surveys, providing improved understanding of variations in forest return following disturbance (Frolking et al., 2009); however, the linkages between spectral measures of recovery and manifestations of forest structure have not been well understood (Schroeder et al., 2011). In this research we demonstrated that LTS-derived spectral measures of recovery, specifically the number of years required for a pixel to return to a certain percentage of its pre-disturbance NBR value, relate to measures of forest cover and height derived from ALS data. We found that an 80% threshold provided the most realistic assessment of recovery, with the majority of pixels identified as spectrally recovered also attaining FAO benchmarks of cover and height for forest, while minimizing false positives that were more common with the lower 60% threshold (i.e. pixels that were spectrally recovered but which had not attained benchmarks of cover and height), as well as the false negatives associated with the 100% threshold (i.e. pixels identified as not spectrally recovered but which had attained the benchmarks for cover and height). In the forest environment of southern Finland, a probabilistic definition of recovery (Y2RZ: the year in which the pixel's NBR value is no longer significantly different from its pre-disturbance NBR value) provided similar results to that of the 80% threshold. Such a data driven approach may be desirable for operational assessments of recovery over large areas; however, the approach as presented herein, should be tested in other forest environments prior to its widespread application. Spectral recovery metrics derived from LTS data offer synoptic, spatially-explicit and retrospective assessments of forest recovery and can provide a useful and meaningful heuristic for large-area assessments of forest recovery. We have demonstrated that national ALS acquisitions can provide useful data to corroborate these thresholds and improve our understanding of the linkage between forest structural development and spectral measures of recovery.

Acknowledgements

This research was undertaken as part of the “Earth Observation to Inform Canada's Climate Change Agenda (EO3C)” project jointly funded by the Canadian Space Agency (CSA), Government Related Initiatives Program (GRIP), and the Canadian Forest Service (CFS) of Natural Resources Canada. This research was enabled in part by support provided by WestGrid (www.westgrid.ca) and Compute Canada (www.computecanada.ca) with additional support from a Natural Sciences and Engineering Research Council (NSERC) (RGPIN 311926-13). Discovery Grant to N. Coops and from the Academy of Finland through the Centre of Excellence in Laser Scanning Research (project number 272195). G. Hobart is thanked for his assistance in acquiring the Landsat data and preparing the annual best-available pixel image composites.

References

Ahmed, O.S., Franklin, S.E., Wulder, M.A., 2014. Interpretation of forest disturbance using a time series of Landsat imagery and canopy structure from airborne lidar. Can. J. Remote. Sens. 39 (6), 521–542.
Banskota, A., Kayastha, N., Falkowski, M.J., Wulder, M.A., Froese, R.E., White, J.C., 2014. Forest monitoring using Landsat time series data: a review. Can. J. Remote. Sens. 40 (5), 362–384.
Barrett, F., McRoberts, R.E., Tomppo, E., Cienciala, E., Waser, L.T., 2016. A questionnaire-based review of the operational use of remotely sensed data by national forest inventories. Remote Sens. Environ. 174, 279–289.
Bartels, S.F., Chen, H.Y.H., Wulder, M.A., White, J.C., 2016. Trends in post-disturbance recovery rates of Canada's forests following wildfire and harvest. For. Ecol. Manag. 361, 194–207.
Baumann, M., Ozdogan, M., Kuemmerle, T., Wendland, K.J., Esipova, E., Radeloff, V.C., 2012. Using the Landsat record to detect Forest-cover changes during and after the collapse of the Soviet Union in the temperate zone of European Russia. Remote Sens. Environ. 124, 174–184.
Bolton, D.K., Coops, N.C., Wulder, M.A., 2015. Characterizing residual structure and forest recovery following high-severity fire in the western boreal of Canada using Landsat time-series and airborne LiDAR data. Remote Sens. Environ. 163, 48–60.
Bolton, D.K., Coops, N.C., Hermosilla, T., Wulder, M.A., White, J.C., 2017. Assessing variability in post-fire forest structure along gradients of productivity in the Canadian boreal using multi-source remote sensing. J. Biogeogr. 44, 1294–1305.
Bolton, D.K., White, J.C., Wulder, M.A., Coops, N.C., Hermosilla, T., Yuan, X., 2018. Updating stand-level forest inventories using airborne laser scanning and Landsat time series data. Int. J. Appl. Earth Obs. Geoinf. 66, 174–183.
Chu, T., Guo, X., 2014. Remote sensing techniques in monitoring post-fire effects and patterns of forest recovery in boreal forest regions: a review. Remote Sens. 6, 470–520.
Chu, T., Guo, X., Takeda, K., 2016. Remote sensing approach to detect post-fire vegetation regrowth in Siberian boreal larch forest. Ecol. Indic. 62, 32–46.
Cohen, W.B., Goward, S.N., 2004. Landsat's role in ecological applications of remote sensing. Bioscience 54, 535–545.
Cohen, W.B., Yang, Z., Kennedy, R., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — tools for calibration and validation. Remote Sens. Environ. 114, 2911–2924.
Cohen, W.B., Yang, Z., Healey, S.P., Kennedy, R.E., Gorelick, N., 2018. A LandTrendr multispectral ensemble for forest disturbance detection. Remote Sens. Environ. 205, 131–140.
Culotta, S., Boncina, A., Carvalho-Ribeiro, S.M., Chauvin, C., Farcy, C., Kurttila, M., Maetzke, F.G., 2015. Forest planning across Europe: the spatial scale, tools, and intersectoral integration in land-use planning. J. Environ. Plan. Manag. 58 (8), 1384–1411.
de Groot, W.J., Cantin, A.S., Flannigan, M.D., Soja, A.J., Gowman, L.M., Newbery, A., 2013. A comparison of Canadian and Russian boreal forest fire regimes. For. Ecol. Manag. 294, 23–34.
Escuin, S., Navarro, R., Fernández, P., 2008. Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from Landsat TM/ETM images. Int. J. Remote Sens. 29, 1053–1073.
FAO, 2012. FRA 2015: terms and definitions. In: Forest Resource Assessment Working Paper 180. Food and Agriculture Organization of the United Nations, Rome, Italy Available online. http://www.fao.org/docrep/017/ap862e/ap862e00.pdf, Accessed date: 5 January 2018.
Forest Stewardship Council Finland, 2010. FSC standard for Finland. FSC-SECR-0160. Available online. https://fi.fsc.org/preview.fsc-standard-for-finland-v1-1-approved-210111.a-84.pdf, Accessed date: 8 May 2018.
Frazier, R.J., Coops, N.C., Wulder, M.A., 2015. Boreal shield forest disturbance and recovery trends using Landsat time series. Remote Sens. Environ. 170, 317–327.
Frazier, R., Coops, N.C., Wulder, M.A., Hermosilla, T., White, J.C., 2018. Analyzing spatial and temporal variability in short-term rates of post-fire vegetation return from Landsat time series. Remote Sens. Environ. 205, 32–45.
Frolking, S., Palace, M.W., Clark, D.B., Chambers, J.Q., Shugart, H.H., Hurtt, G.C., 2009. Forest disturbance and recovery: a general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure. J. Geophys. Res. 114, 1–27.
Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V.C., Abrudan, I.V., Lieskovsky, J., Munteanu, C., Ostapowicz, K., Hostert, P., 2014. Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites. Remote Sens. Environ. 151, 72–88.
Goward, S.M., Williams, D.L., Arvidson, T., Rochchio, L.E.P., Irons, J.R., Russel, C.A., Johnston, S.S., 2017. Landsat’s enduring legacy: pioneering global land observations from space. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland. 1-57083-101-7pp. 586.
Gustafsson, L., Kouki, J., Sverdrup-Thygeson, A., 2010. Tree retention as a conservation measure in clear-cut forests of northern Europe: a review of ecological consequences. Scand. J. For. Res. 25, 295–308.
Healey, S.P., Cohen, W.B., Zhiqiang, Y., Krankina, O.N., 2005. Comparison of tasseled cap-based Landsat data structures for use in forest disturbance detection. Remote Sens. Environ. 97, 301–310.
Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., 2015a. An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. Remote Sens. Environ. 158, 220–234.
Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., 2015b. Regional detection, characterization, and attribution of annual forest change from 1984 to

2012 using Landsat-derived time-series metrics. Remote Sens. Environ. 170, 121–132.
Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Campbell, L.B., 2016. Mass data processing of time series Landsat imagery: pixels to data products. Int. J. Digital Earth 9 (11), 1035–1054.
Horler, D.N.H., Ahern, F.J., 1986. Forestry information content of Thematic Mapper data. Int. J. Remote Sens. 7 (3) (405–128).
Isenburg, M., 2017. Efficient LiDAR processing software (version 160910, academic). obtained from. http://rapidlasso.com/LAStools.
Kankare, V., Vauhkonen, J., Holopainen, M., Vastaranta, M., Hyyppa, J., Hyyppa, H., Alho, P., 2015. Sparse density, leaf-off airborne laser scanning data in aboveground biomass component prediction. Forests 6 (6), 1839–1857.
Kennedy, R.E., Cohen, W.B., Schroeder, T.A., 2007. Trajectory-based change detection for automated characterization of forest disturbance dynamics. Remote Sens. Environ. 110, 370–386.
Kennedy, R.E., Yang, Z., Cohen, W.B., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — temporal segmentation algorithms. Remote Sens. Environ. 114, 2897–2910.
Kennedy, R.E., Yang, Z., Cohen, W.B., Pfaff, E., Braaten, J., Nelson, P., 2012. Spatial and temporal patterns of forest disturbance and regrowth within the area of the northwest forest plan. Remote Sens. Environ. 122, 117–133.
Keogh, E., Chu, S., Hart, D., Pazzani, M., 2001. An online algorithm for segmenting time series. In: Proceedings IEEE International Conference on Data Mining, 2001, pp. 289–296 (ICDM 2001).
Key, C.H., Benson, N.C., 2006. Landscape assessment (LA). FIREMON: fire effects monitoring and inventory system. In: Lutes, D.C., Keane, R.E., Carati, J.F., Key, C.H., Benson, N.C., Gangi, L.J. (Eds.), General Technical Report RMRS-GTR-164-CD. USDA Forest Service, Rocky Mountains Research Station, Fort Collins, CO (pp. LA − 1-55).
Korhonen, L., Pippuri, I., Packalén, P., Heikkinen, V., Maltamo, M., Heikkilä, J., 2013. Detection of the need for seedling stand tending using high-resolution remote sensing data. Silva Fennica 47 (2), 1–20.
Kotivuori, E., Korhonen, L., Packalen, P., 2016. Nationwide airborne laser scanning based models for volume, biomass and dominant height in Finland. Silva Fennica 50 (4), 1–28.
Kuusinen, N., Stenberg, P., Tomppo, E., Bernier, P., Berninger, F., 2015. Variation in understory and canopy reflectance during stand development in Finnish coniferous forests. Can. J. For. Res. 45, 1077–1085.
Li, X., Strahler, A., 1985. Geometric-optical modeling of a conifer forest canopy. IEEE Trans. Geosci. Remote Sens. GE-23 (5), 705–721.
Madoui, A., Gauthier, S., Leduc, A., Bergeron, Y., Valeria, O., 2015. Monitoring forest recovery following wildfire and harvest in boreal forests using satellite imagery. Forests 6, 4105–4134.
Miller, J.R., White, P.H., Chen, J.M., Peddle, D.R., McDermid, G., Fournier, R.A., Shepherd, P., Rubinstein, I., Freemantle, J., Soffer, R., Ledrew, E., 1997. Seasonal change in understory reflectance of boreal forests and influence on canopy vegetation indices. J. Geophys. Res. 102 (D24), 475–482.
Natural Resources Institute of Finland, 2015. The multi-source national forest inventory raster maps of 2013. Available at. http://kartta.luke.fi/index-en.html, Accessed date: 1 June 2017.
Nilson, T., Peterson, U., 1994. Age dependence of forest reflectance: analysis of main driving factors. Remote Sens. Environ. 48, 319–331.
Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57.
Olsson, H., 2009. A method for using Landsat time series for monitoring young plantations in boreal forests. Int. J. Remote Sens. 30 (19), 5117–5131.
Ørka, H.O., Gobakken, T., Næsset, E., 2016. Predicting attributes of regeneration forests using airborne laser scanning. Can. J. Remote. Sens. 42, 541–553.
Pascual, C., Garcia-Abril, A., Cohen, W.B., Martin-Fernandez, S., 2010. Relationship between LiDAR-derived forest canopy height and Landsat images. Int. J. Remote Sens. 31 (5), 1261–1280.
PEFC Finland, 2014. Criteria for PEFC forest certification. Available online. http://pefc.fi/wp-content/uploads/2016/09/PEFC_FI_1002_2014_Criteria_for_Forest_Certification_20141027.pdf, Accessed date: 8 May 2018.
Peterson, U., Nilson, T., 1993. Successional reflectance trajectories in northern temperate forests. Int. J. Remote Sens. 14 (3), 609–613.
Pflugmacher, D., Cohen, W.B., Kennedy, R.E., 2012. Using Landsat-derived disturbance history (1972–2010) to predict current forest structure. Remote Sens. Environ. 122, 146–165.
Pflugmacher, D., Cohen, W.B., Kennedy, R.E., Yang, Z., 2014. Using Landsat-derived disturbance and recovery history and lidar to map forest biomass dynamics. Remote Sens. Environ. 151, 124–137.
Pickell, P.D., Hermosilla, T., Frazier, R.J., Coops, N.C., Wulder, M.A., 2016. Forest recovery trends derived from Landsat time series for North American boreal forests. Int. J. Remote Sens. 37, 138–149.
Potapov, P.V., Turubanova, S.A., Tyukavina, A., Krylov, A.M., McCarty, J.L., Radeloff, V.C., Hansen, M.C., 2015. Eastern Europe's forest cover dynamics from 1985 to 2012 quantified from the full Landsat archive. Remote Sens. Environ. 159, 28–43.
Saarinen, N., White, J.C., Wulder, M.A., Kangas, A., Tuominen, S., Kankare, V., Holopainen, M., Hyyppä, J., Vastaranta, M., 2018. Landsat archive holdings for Finland: Opportunities for forest monitoring. Silva Fennica 52 (3), 9986. http://dx.doi.org/10.14214/sf.9986.
Schmidt, G.L., Jenkerson, C.B., Masek, J., Vermote, E., Gao, F., 2013. Landsat ecosystem disturbance adaptive processing system (LEDAPS) algorithm description. U.S. Geol. Surv. Open File Rep. 2013–1057, 17. Available online. http://pubs.usgs.gov/of/2013/1057/.
Schroeder, T.A., Cohen, W.B., Yang, Z., 2007. Patterns of forest regrowth following clearcutting in western Oregon as determined from Landsat time-series. For. Ecol. Manag. 243, 259–273.
Schroeder, T.A., Wulder, M.A., Healey, S.P., Moisen, G.G., 2011. Mapping wildfire and clearcut harvest disturbances in boreal forests with Landsat time series. Remote Sens. Environ. 115, 1421–1433.
Senf, C., Pflugmacher, D., Hostert, Pl, Seidl, R., 2017. Using Landsat time series for characterizing forest disturbance dynamics in the coupled human and natural systems of Central Europe. ISPRS J. Photogramm. Remote Sens. 130, 453–463.
Song, C., Woodcock, C.E., 2003. Monitoring forest succession with multitemporal Landsat images: factors of uncertainty. IEEE Trans. Geosci. Remote Sens. 41 (11), 2557–2567.
Song, C., Woodcock, C.E., Li, X., 2002. The spectral/temporal manifestation of forest succession in optical imagery: the potential of multitemporal imagery. Remote Sens. Environ. 82, 285–302.
Spake, R., Ezard, T.H.G., Martin, P.A., Newton, A.C., Doncaster, C.P., 2015. A meta-analysis of functional group responses to forest recovery outside of the tropics. Conserv. Biol. 29 (6), 1695–1703.
Tomppo, E., 1990. Satellite image-based National Forest Inventory of Finland. In: Proceedings of the Symposium on Global and Environmental Monitoring, Techniques and Impacts, Victoria, British Columbia, Canada, 17–21 September 1990. International Archives of Photogrammetry and Remote Sensing Vol. XXVIII. pp. 419–424 (part 1–7 (1991)).
Tomppo, E., Olsson, H., Ståhl, G., Katila, M., 2008. Creation of forest data bases by combining National Forest Inventory Field Plots and Remote Sensing. Remote Sens. Environ. 112, 1982–1999.
Vogeler, J.C., Yang, Z., Cohen, W.B., 2015. Mapping post-fire habitat characteristics through the fusion of remote sensing tools. Remote Sens. Environ. 173, 294–303.
White, J.C., Wulder, M.A., 2013. The Landsat observation record of Canada: 1972–2012. Can. J. Remote. Sens. 39 (6), 455–467.
White, J.C., Wulder, M.A., Hobart, G.W., Luther, J.E., Hermosilla, T., Griffiths, P., Coops, N.C., Hall, R.J., Hostert, P., Dyk, A., Guindon, L., 2014. Pixel-based image compositing for large-area dense time series applications and science. Can. J. Remote. Sens. 40 (3), 192–212.
White, J.C., Wulder, M.A., Hermosilla, T., Coops, N.C., Hobart, G., 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.
Woodcock, C.E., Allen, R.G., Anderson, M., Belward, A., Bindschadler, R., Cohen, W.B., ... Wynne, R., 2008. Free access to Landsat imagery. Science 320 (5879), 1011.
Wulder, M.A., Skakun, R.S., Dymond, C.C., Kurz, W.A., White, J.C., 2005. Characterization of the diminishing accuracy in detecting forest insect damage over time. Can. J. Remote. Sens. 31 (6), 421–431.
Wulder, M.A., Campbell, C., White, J.C., Flannigan, M., Campbell, I.D., 2007. National circumstances in the international circumboreal community. For. Chron. 83 (4), 539–556.
Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2012. Opening the archive: how free data has enabled the science and monitoring promise of Landsat. Remote Sens. Environ. 122, 2–10.
Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G., Roy, D.P., 2016. The Global Landsat Archive: status, consolidation, and direction. Remote Sens. Environ. 185, 271–283.
Zald, H.S.J., Ohmann, J.L., Roberts, H.M., Gregory, M.J., Henderson, E.B., McGaughey, R.J., Braaten, J., 2014. Influence of lidar, Landsat imagery, disturbance history, plot location accuracy, and plot size on accuracy of imputation maps of forest composition and structure. Remote Sens. Environ. 143, 26–38.
Zhu, Z., Woodcock, C.E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sens. Environ. 118, 83–94.