Remote Sensing of Environment 216 (2018) 697–714

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Three decades of forest structural dynamics over Canada's forested ecosystems using Landsat time-series and lidar plots

Giona Matasci[a], Txomin Hermosilla[a], Michael A. Wulder[b], Joanne C. White[b], Nicholas C. Coops[a,*], Geordie W. Hobart[b], Douglas K. Bolton[a], Piotr Tompalski[a], Christopher W. Bater[c]

[a] *Integrated Remote Sensing Studio, Department of Forest Resources Management, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada*
[b] *Canadian Forest Service, Pacific Forestry Centre, Natural Resources Canada, 506 West Burnside Road, Victoria, BC V8Z 1M5, Canada*
[c] *Forest Management Branch, Forestry Division, Alberta Agriculture and Forestry, Edmonton, Alberta T5K 2M4, Canada*

ARTICLE INFO

Keywords:
Landsat
Lidar
Forest recovery
Imputation
Canopy cover
Biomass

ABSTRACT

Attributes that describe forest structure, such as height, canopy cover, volume, and biomass, are required to inform forest inventories and monitoring programs. Light Detection and Ranging (lidar) has been successfully demonstrated as a means to derive a suite of forest structure attributes at the plot level; however, these acquisitions are often constrained to limited spatial extents and to a given point in time. Sample based approaches for model development can accommodate the spatial limitations of lidar acquisitions when characterizing large areas. The combination of lidar plot data and time-series satellite imagery is well suited to provide spatially extensive, and temporally dense, information on forest structure and related dynamics over very large areas. In this research, we combine lidar plot-derived information with Landsat pixel-based composites to produce annual forest structure estimates from 1984 to 2016 over 650 million ha of Canada's forest ecosystems using a nearest neighbor imputation approach with a Random Forests-based distance metric. Imputed variables included lidar metrics of height (e.g., mean height, standard deviation of height) and cover, as well as area-based modelled inventory estimates of Lorey's height, basal area, stem volume, and biomass. Models were validated using reserved validation plots, with model R^2 ranging from 0.62 to 0.64 for lidar metrics of height and cover, and R^2 of 0.67, 0.68, 0.71, and 0.70 for Lorey's height, basal area, volume and biomass, respectively. Unique to this study was the assessment of model extension through time, with model performance for imputing lidar metrics evaluated at the forest stand-level using independent lidar data representing a latitudinal gradient of forest conditions and that was not used in model development. The period evaluated was 2006–2012, with R^2 values ranging from 0.36 to 0.66 for height metrics, and 0.47–0.77 for cover metrics. Ultimately, we show how deriving forest structural estimates on an annual basis enables the analysis of both the dynamics and regional trends of undisturbed forest, as well as regenerating stands following stand-replacing disturbances (i.e., fire, harvesting).

1. Introduction

Monitoring plays a foundational role in supporting sustainable forest management, and informing the development of policies aimed at preserving and maintaining ecosystem services and biodiversity in forests while concurrently accommodating human needs (Daily, 1997). Moreover, spatially-explicit estimates of forest attributes inform reporting activities by providing data for forest (White et al., 2014) and carbon (Boisvenue et al., 2016) monitoring programs. National forest inventory programs are typically designed to produce long-term data in support of forest monitoring (Kangas and Maltamo, 2006; MacDicken, 2015). Many of these programs, however, are sample-based and aspatial, and cannot provide spatially-explicit inputs for modelling unless they are combined with other forms of inventory data or remotely sensed data (e.g., Beaudoin et al., 2014; Tomppo et al., 2009). Thus, there is a need for spatially-explicit forest monitoring information collected at a resolution suitable for capturing anthropogenic impacts, and supporting a range of scientific and policy elements. Furthermore, the capacity to generate this information retrospectively can provide useful baseline information for understanding forest dynamics (White et al., 2017) and for modelling potential vulnerabilities to climate change (Price et al., 2013). In addition, a time-series of forest structure

* Corresponding author.
E-mail address: nicholas.coops@ubc.ca (N.C. Coops).

https://doi.org/10.1016/j.rse.2018.07.024
Received 9 May 2018; Received in revised form 17 July 2018; Accepted 21 July 2018

Table 1
Forest structural variables estimated in this study. Lidar returns elevation values are normalized to the ground surface.

Nature of variables	Forest structural variable	Variable name	Units	Description
Extracted directly from point cloud	Mean canopy height	elev_mean	m	Mean height of lidar first returns
	Standard deviation of canopy height	elev_sd	m	Standard deviation of first returns height
	Coefficient of variation of canopy height	elev_cv	–	Coefficient of variation of first returns height
	95th percentile of canopy height	elev_p95	m	95th percentile of first returns height
	Canopy cover	cover_2m	%	Percentage of first returns above 2 m
	Canopy cover above mean height	cover_mean	%	Percentage of first returns above the mean height
Modelled inventory attributes	Lorey's mean height	loreys_height	m	Average height of trees weighted by their basal area
	Basal area	basal_area	m^2/ ha	Cross-sectional area of tree stems at breast height. The sum of the cross-sectional area (i.e., basal area) of each tree in square metres in a plot, divided by the area of the plot.
	Gross stem volume	stem_volume	m^3/ ha	Individual tree gross volumes are calculated using species-specific allometric equations. Gross total volume per hectare is calculated by summing the gross total volume of all trees and dividing by the area of the plot.
	Total aboveground biomass	ag_biomass	t/ha	Individual tree total aboveground biomass is calculated using species-specific equations. Aboveground biomass per hectare is calculated by summing the values of all trees within a plot and dividing by the area of the plot. Aboveground biomass may be separated into various biomass components (e.g., stem, bark, branches, foliage).

attributes including height, canopy cover, volume, and biomass, can also inform on relative trends in forest growth and condition, as well as post-disturbance forest recovery (Bartels et al., 2016; Frolking et al., 2009; Masek et al., 2011). Further, such a time-series recording forest structure can fill critical information gaps for unmanaged forests, where there exists a paucity of spatially exhaustive forest inventory information (Gillis et al., 2005).

Satellite programs with medium spatial resolution (10–100 m) sensors (Belward and Skøien, 2014), such as those of the Landsat mission, provide data for capturing and characterizing both status and change over terrestrial ecosystems at human scales (Wulder et al., 2008b). Image acquisitions from Landsat sensors including Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+), and Operational Land Imager (OLI) have the spatial grain (30 m spatial resolution), spectral bands (visible to short-wave infrared), and revisit time (single sensor, 16 days) required to study vegetation trends with an annual/seasonal frequency (Kovalskyy and Roy, 2013). Since the launch of Landsat-7 in 1999, Landsat has effectively had an 8-day revisit based upon having two satellites in orbit at any given time. The opening of the multi-decadal Landsat archive (Woodcock et al., 2008), combined with the systematic production of science-supported, analysis-ready data products (e.g., surface reflectance, Vermote et al., 2016) has accelerated a number of methodological developments that have advanced satellite-based monitoring activities (Hansen and Loveland, 2012; Wulder et al., 2012a). The process of using analysis ready data, high performance computing, and robust automated algorithms to characterize large areas over time is reviewed in Wulder et al. (2018).

Previously, image compositing methods were more commonly applied to coarse spatial resolution data sources (Cihlar, 2000; Holben, 1986), which were freely available and had a frequent revisit rate. Free and open access to analysis-ready data led to the application of image compositing approaches to Landsat data (Roy et al., 2010). Image compositing allows clear observations for a given pixel to be selectively used from otherwise cloudy images, resulting in the generation of seasonal or annual, gap-free, composites (Griffiths et al., 2013; Hermosilla et al., 2015a; White et al., 2014). These best-available pixel (BAP) composites (White et al., 2014) result in a data space where the spectral bands can be considered representative of a given point in time (e.g., year, season). Furthermore, using surface reflectance derived from a radiometrically calibrated image data source (Markham and Helder, 2012) results in pixel-level values for a given land cover or forest structural condition that can be considered as temporally invariant (Fekety et al., 2014), enabling the application of models through time and space (Song et al., 2001). Thus, Landsat data have enabled the generation of wall-to-wall estimates of forest structure based on the temporal analysis of the spectral trends and/or the change information provided by Landsat time series data (Bolton et al., 2018; Matasci et al., 2018; Pflugmacher et al., 2014, 2012), and to extend these estimates through time (Deo et al., 2017).

Nearest neighbor (NN) imputation is a demonstrated methodological framework to relate environmental-based predictors and inventory-related attributes (Eskelson et al., 2009; Ohmann and Gregory, 2002), as well as Landsat data and lidar-derived attributes (Zald et al., 2014). With a 1-NN structure, imputation has the advantage of assigning a set of measured attributes that actually occur in a forest stand (at a given donor plot location), ensuring prediction of realistic canopy conditions (Hudak et al., 2008). Imputation has been the primary methodological building block of prior studies that investigated single-year forest structure mapping (Tomppo et al., 2009; Zald et al., 2016). A number of studies have shown promising results in extending imputation models to predict forest structure through time, demonstrating the opportunities offered by the generated outputs to inform the study of forest growth and post-disturbance recovery (Deo et al., 2017; Fekety et al., 2014).

In previous work, we applied an imputation approach using lidar plots and Landsat data and generated spatially explicit, wall-to-wall estimates of ten key forest structural attributes (see Table 1) across Canada's boreal forest for a single year (2010) (Matasci et al., 2018). In this current study, we extend the large-area forest attribute imputation model presented in Matasci et al. (2018) through both time (1984–2016) and space (integrating data from the hemi-boreal zone, see Fig. 1), thereby generating annual estimates of the same set of lidar-based metrics and forest structural attributes for the entire treed extent of Canada's forested ecosystems over 33 years. Our objectives were three-fold: (i) to demonstrate the temporal and spatial extension of the imputation model using a time-series of annual surface reflectance image composites and samples of airborne lidar; (ii) to demonstrate the robustness of the outputs by validating the resulting forest structural estimates using a decade of independent lidar data acquisitions across a latitudinal range of forest conditions; and (iii) to highlight the potential for scientific insights related to growth and recovery over large areas, which are enabled through the use of the time-series of forest structure developed herein.