Int J Appl Earth Obs Geoinformation 66 (2018) 174–183

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

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

Updating stand-level forest inventories using airborne laser scanning and Landsat time series data

Douglas K. Bolton[a,*], Joanne C. White[b], Michael A. Wulder[b], Nicholas C. Coops[a], Txomin Hermosilla[a], Xiaoping Yuan[c]

[a] Integrated Remote Sensing Studio, Department of Forest Resources Management, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, British Columbia, V6T 1Z4, Canada

[b] Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, 506 West Burnside Road, Victoria, British Columbia, V8Z 1M5, Canada

[c] Forest Analysis and Inventory Branch,Forest Stewardship Division, Ministry of Forests, Lands, Natural Resource Operations and Rural Development, PO Box 9512, Station Provincial Government, Victoria, BC, V8W 9C2, Canada

ARTICLE INFO

Keywords:
Remote sensing
Multispectral imagery
Forest attributes
Imputation
Time series
Lidar

ABSTRACT

Vertical forest structure can be mapped over large areas by combining samples of airborne laser scanning (ALS) data with wall-to-wall spatial data, such as Landsat imagery. Here, we use samples of ALS data and Landsat time-series metrics to produce estimates of top height, basal area, and net stem volume for two timber supply areas near Kamloops, British Columbia, Canada, using an imputation approach. Both single-year and time series metrics were calculated from annual, gap-free Landsat reflectance composites representing 1984–2014. Metrics included long-term means of vegetation indices, as well as measures of the variance and slope of the indices through time. Terrain metrics, generated from a 30 m digital elevation model, were also included as predictors. We found that imputation models improved with the inclusion of Landsat time series metrics when compared to single-year Landsat metrics (relative RMSE decreased from 22.8% to 16.5% for top height, from 32.1% to 23.3% for basal area, and from 45.6% to 34.1% for net stem volume). Landsat metrics that characterized 30-years of stand history resulted in more accurate models (for all three structural attributes) than Landsat metrics that characterized only the most recent 10 or 20 years of stand history. To test model transferability, we compared imputed attributes against ALS-based estimates in nearby forest blocks (> 150,000 ha) that were not included in model training or testing. Landsat-imputed attributes correlated strongly to ALS-based estimates in these blocks (R^2 = 0.62 and relative RMSE = 13.1% for top height, R^2 = 0.75 and relative RMSE = 17.8% for basal area, and R^2 = 0.67 and relative RMSE = 26.5% for net stem volume), indicating model transferability. These findings suggest that in areas containing spatially-limited ALS data acquisitions, imputation models, and Landsat time series and terrain metrics can be effectively used to produce wall-to-wall estimates of key inventory attributes, providing an opportunity to update estimates of forest attributes in areas where inventory information is either out of date or non-existent.

1. Introduction

Forest inventories are generated to support a variety of information needs ranging from operational to strategic. Data required to support these information needs likewise vary in terms of spatial and temporal resolution. In Canada, extensive forest management practices (Wulder et al., 2007) have led to the dominance of strategic-level forest inventories generated from air photo interpretation and ground sampling (Leckie and Gillis, 1995). However, these inventories are preferentially generated for managed forest areas, which represent only 65% of Canada's total forest area (Bernier et al., 2012). The remaining unmanaged forest areas have no systematic inventory information. Moreover, many of the existing forest inventories in the managed forest area are more than 20 years old, creating a need for up-to-date, cost-effective, strategic forest information to support applications such as projections of timber supply.

Airborne laser scanning (ALS) is an active remote sensing technology that enables three-dimensional forest structure to be characterized over larger spatial scales than is possible with conventional field methods (Lim et al., 2003). The use of ALS data has transformed forest inventory practices in many jurisdictions (White et al., 2016), and when combined with high quality ground plot data—in what is

* Corresponding author.
E-mail address: douglas.bolton@ubc.ca (D.K. Bolton).

https://doi.org/10.1016/j.jag.2017.11.016
Received 17 September 2017; Received in revised form 29 November 2017; Accepted 30 November 2017
Available online 08 December 2017

commonly referred to as an area-based approach (Næsset, 2002)—can produce attribute estimates that typically meet or exceed inventory accuracy requirements (Magnussen et al., 2012). Acquisitions of ALS data in Canada to support forest inventories have been increasing steadily over the past decade (D'Eon and Macafee, 2016), although the acquisitions have not been systematic and often target specific management areas of interest. Despite the pace of change in many regions, there is a need to generate and/or update strategic-level forest inventory information more rapidly over large areas, particularly in jurisdictions that have experienced widespread disturbance, such as via wildfire or as caused by the mountain pine beetle in British Columbia.

To address information needs in forests that lack ALS or inventory data, or where these data are out-of-date, optical remote sensing data can be used to estimate forest attributes from nearby ALS data collections to generate up-to-date and spatially extensive estimates of forest attributes (Frazier et al., 2014; Pflugmacher et al., 2014, 2012; Zald et al., 2016). Specifically, the sensors onboard the Landsat series of satellites have been collecting multispectral data of the planet continuously from 1982 to the present at 30-m spatial resolution, and back to 1972 at 60-m resolution, although prior to 1984, the acquisition of 30 m data was sparse (Goward et al., 2006). Following the opening of the Landsat archive in 2008 (Wulder et al., 2012a), along with advances in computing (hardware and software; Wulder and Coops, 2014), cloud masking (Zhu and Woodcock, 2014) and surface reflectance generation (Masek et al., 2006), approaches have been developed that utilize all available Landsat imagery to characterize forest change (Zhu, 2017). This Landsat time series information has been used in past research to extrapolate ALS-derived estimates of forest attributes (Frazier et al., 2014; Pflugmacher et al., 2012; Zald et al., 2016), with Pflugmacher et al. (2012) demonstrating that the inclusion of Landsat time series data improves model results over models that rely on a single date of Landsat imagery alone.

K-Nearest Neighbor (kNN) imputation has been widely used to produce wall-to-wall estimates of forest attributes using remotely sensed data (Andersen et al., 2011; Beaudoin et al., 2014; Bright et al., 2014; Hudak et al., 2008; Makela and Pekkarinen, 2004; Mora et al., 2013), given that the technique is non-parametric and can support multi-variate analysis (Chirici et al., 2016). In most applications, forest attributes are imputed directly from field plots or inventory data to spatially extensive remote sensing layers (e.g., Makela and Pekkarinen, 2004; Tomppo et al., 2008; Beaudoin et al., 2014). In Alaskan boreal forests, however, Andersen et al. (2011) found that incorporating ALS data in a two-stage approach led to improved estimates of biomass across the landscape when compared to imputing directly from field plots. Specifically, estimates of forest attributes were first derived along ALS transects by developing models between ground plots and ALS metrics. Second, these ALS-derived forest attributes were used as input data to impute forest attributes across the area of interest using Landsat imagery and synthetic aperture radar (SAR) data. Given the large area covered by the ALS transects, a wider range of structural variability could be supplied to the imputation models, leading to increased estimation accuracy. Similarly, Wilkes et al. (2015) implemented a two-stage approach that used ALS data in combination with a range of satellite data products to estimate canopy height across 2.9 million ha of forest in Victoria, Australia with a relative Root Mean Square Error (RMSE) of $< 31\%$ when compared to independent field plots. In addition to incorporating optical imagery from Landsat and the Moderate Resolution Imaging Spectroradiometer (MODIS), Wilkes et al. (2015) also included climate data, topographic information, and soil maps as predictors to produce wall-to-wall maps of canopy height.

In order to implement a two-stage approach to estimating forest attributes, highly accurate estimates of forest attributes from ALS data are first required. The area-based approach has become a standard and accepted method for generating high accuracy estimates of forest attributes over large areas with ALS data (Næsset, 2014; Wulder et al., 2013). First, discrete point clouds of ALS data are summarized using a suite of metrics that describe vegetation cover, stand height, and the vertical distribution of ALS returns (Bouvier et al., 2015; Lefsky et al., 2005; Tompalski et al., 2015). Second, through parametric (Næsset et al., 2004; Woods et al., 2011; Wulder et al., 2012b) or non-parametric (Hudak et al., 2008; Penner et al., 2013) approaches, these ALS metrics are related against field measured forest attributes, such stem volume or basal area. Finally, the developed models are then applied to predict forest attributes wall-to-wall across the ALS data collection.

Given the incrementally developing coverage of ALS data that exist across many jurisdictions, an opportunity exists to explore how these high-quality datasets can be leveraged to provide wall-to-wall estimates of key forest attributes in areas where ALS data have not been collected and where existing inventory data is out of date. In this analysis, we develop a methodology to produce models that impute forest attributes using a suite of Landsat single-year and time series predictors near Kamloops, British Columbia, Canada, where ALS data was collected and forest attributes were predicted in 2014 for several spatially disjointed areas covering approximately 350,000 ha. Through this analysis, we ask the following specific questions.

1.1. How does the predictive capability of imputation models for top height, basal area, and stem volume change when single-year Landsat metrics are replaced with Landsat time series metrics?

Most attempts to predict forest attributes with Landsat time series information have relied on metrics that describe disturbance and recovery dynamics (e.g., Pflugmacher et al., 2012; Frazier et al., 2014). However, as many sampled stands within our study area have not undergone a major disturbance during the interval of the Landsat recorded we assessed (1984–2014), information on disturbance and recovery from 1984 to 2014 is not likely informative for many stands. Alternatively, time series metrics such as long-term spectral means, variability of spectral indices through time, and the slope of indices through time, can describe long-term stand conditions and development, regardless of disturbance history. Here, we assess the relative importance of Landsat predictor variables that describe single-year spectral conditions and disturbance history versus metrics that describe long-term spectral conditions, with a special interest on undisturbed stands.

1.2. For LANDSAT time series predictors, what length of time yields the best results for making attribute estimates: 10-year, 20-year, or 30-year descriptors?

Landsat time series metrics can be calculated across a number of time periods. Longer periods of time will capture more of a stand's history (e.g., 30 year spectral averages), while shorter time periods will more accurately describe the current state of a forest stand (e.g., 10 year spectral averages). To determine how the length of time described impacts model accuracy, we compare Landsat time-series metrics calculated at 10, 20, and 30 years prior to the year of the ALS data collection (2014).

1.3. How well do pixel-level imputation models estimate forest structure at the stand-level?

Landscape-level decision making for forest management is not often made for individual grid cells (pixels) from remotely sensed data products, but rather for stand-level polygons. In British Columbia, stand-level polygons are derived for managed forests through the Vegetation Resources Inventory (VRI), which is a forest inventory program developed by British Columbia's Ministry of Forests, Lands Natural Resource Operations (MFLNRO). Stand-level boundaries are derived in the VRI through photo interpretation, and are combined with a sample of ground plots to produce inventory data across British Columbia (Sandvoss et al., 2005). To determine if imputed forest attributes can accurately capture variability at the scale at which management decisions are made, we compare imputed estimates against ALS estimates at