Int J Appl Earth Obs Geoinformation 72 (2018) 91–98

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

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

How much can airborne laser scanning based forest inventory by tree species benefit from auxiliary optical data?

Mikael Kukkonen*, Lauri Korhonen, Matti Maltamo, Aki Suvanto, Petteri Packalen

University of Eastern Finland, School of Forest Sciences, Finland

A R T I C L E I N F O

Keywords:
Tree species
LiDAR
Sentinel-2
Landsat 8
Aerial image
k nearest neighbor

A B S T R A C T

The objective of this study was to investigate the benefit of three different optical data sources (Sentinel-2, Landsat 8 and aerial images) to support airborne laser scanning (ALS) data in species-specific forest inventory. The data covered 633 sample plots in eastern Finland. We used nearest neighbor imputation for simultaneous prediction of Scots pine, Norway spruce and broadleaved species' volume by species group. The variable selection was performed by means of simulated annealing of different data combinations. The results showed that, on average, all optical data sources improved the species-specific plot volume predictions. The improvement was always greatest for broadleaves. The species-specific root mean square errors were 64.3%, 61.5%, 58.1% and 57.9% for ALS, ALS + Landsat 8, ALS + Sentinel-2 and ALS + aerial image data combinations, respectively, and 54.2% for ALS with the channels of both aerial images and Sentinel-2. Compared to using just ALS and aerial images, adding the Sentinel's second red edge and narrow near-infrared bands improved the separation of pine and spruce. Sentinel-2 outperformed Landsat 8 and was almost as good an option as aerial images. The results suggest that all optical data, from airborne or spaceborne sources, are useful when combined with ALS data in species-specific forest inventory.

1. Introduction

Forest resource inventories increasingly rely on airborne laser scanning (ALS) data. In Nordic countries, ALS has already replaced traditional labor-intensive fieldwork in practical stand level forest management inventories (Vauhkonen et al., 2014). In general, ALS performs well for estimation of height (Næsset and Økland, 2002), total volume (Holmgren et al., 2003; Popescu and Wynne, 2004) and aboveground biomass (Lefsky et al., 2002). However, although in most countries forest attributes are measured by tree species, the ALS inventory research does not particularly reflect this. In fact, due to the high number of tree species, separating them by remote sensing is nearly impossible in many biomes. An exception is the boreal forest in the Nordic countries, where three economically valuable tree species constitute a majority of the growing stock: Scots pine (*Pinus sylvestris* L.), Norway spruce (*Picea abies* (L.) Karst.) and birch (*Betula* spp.).

In Finland, the current stand level forest management inventory system is based on the combined use of low density (< 1 pulse m^{-2}) ALS data, aerial images and local field sample plots to estimate species-specific stand attributes (Maltamo and Packalen, 2014). The ALS-based canopy height and density features are complemented by spectral and texture features computed from aerial images (Packalén and Maltamo, 2006, 2007, 2008; Packalén et al., 2009; Niska et al., 2010; Vauhkonen et al., 2012; Maltamo et al., 2014). The companies that operate practical inventories usually apply a non-parametric k nearest neighbor (NN) model, which is trained based on 500–700 field sample plots (Maltamo and Packalen, 2014). The quality of remote sensing inventories is under constant monitoring because of the requirement for high-accuracy, species-specific forest information.

Optical satellite imagery can also be used to support ALS-based forest inventories. Medium resolution (10–30 m) images such as Landsat are freely available, and offer more spectral bands than typical false-color aerial images. However, temporal image coverage may be an issue in forest inventory applications, because cloud free images from the same season as other datasets may not always be available. However, the new generation of optical, medium resolution remote sensing satellites will considerably improve the data availability and open new options for species-specific forest inventories. The Landsat 8 (L8) satellite was launched in 2013, and contains an Operational Land Imager (OLI) sensor that has nine spectral bands and a spatial resolution of 15–30 m (Table 1). Landsat 9 is currently planned for launch in 2020. Sentinel-2 A (S2) satellite was launched in 2015, and its MSI (Multi-Spectral Instrument) sensor provides a 13-band spectrum, with 10–60 m resolution (Table 1). Its identical pair, Sentinel-2B, was

---

* Corresponding author.
E-mail address: mikael.kukkonen@uef.fi (M. Kukkonen).

https://doi.org/10.1016/j.jag.2018.06.017
Received 5 March 2018; Received in revised form 21 June 2018; Accepted 21 June 2018

Table 1
Spectral bands and resolutions of Landsat 8 OLI and Sentinel-2 MSI sensors.

Landsat 8 OLI			Sentinel-2 MSI		
Band	Wavelength range (nm)	Resolution (m)	Band	Wavelength range (nm)	Resolution (m)
1 Coastal aerosol	433 – 453	30	1 Coastal aerosol	433 – 453	60
2 Blue (B)	450 – 515	30	2 Blue (B)	458 – 523	10
3 Green (G)	525 – 600	30	3 Green (G)	543 – 578	10
4 Red (R)	630 – 680	30	4 Red (R)	650 – 680	10
			5 Red edge 1 (RE1)	698 – 713	20
			6 Red edge 2 (RE2)	733 – 748	20
			7 Red edge 3 (RE3)	773 – 793	20
5 Near infrared (NIR)	845 – 885	30	8 Near infrared (NIR)	785 – 900	10
			8a Near infrared narrow (NIRn)	855 – 875	20
			9 Water vapor	935 – 955	60
9 Shortwave infrared / Cirrus	1360 – 1390	30	10 Shortwave infrared / Cirrus	1360 – 1390	60
6 Shortwave infrared 1 (SWIR1)	1560 – 1660	30	11 Shortwave infrared 1 (SWIR1)	1565 – 1655	20
7 Shortwave infrared 2 (SWIR2)	2100 – 2300	30	12 Shortwave infrared 2 (SWIR2)	2100 – 2280	20
8 Panchromatic	500 – 680	15			

Table 2
Data from 633 sample plots. Note: sd, standard deviation; $V_{species}$, volume (m^3 ha^{-1}) of a particular species.

Variable	mean	sd	min	max
V_{pine}	77.6	99.2	0.0	539.0
V_{spruce}	92.2	130.4	0.0	735.1
$V_{broadleaves}$	38.2	57.4	0.0	320.6
Stem number	1465	1635	60	25 700
Dominant height	20.1	5.7	3.3	33.4

launched on March 2017. The S2 MSI sensor has been especially designed for monitoring vegetation, and contains three new vegetation red edge (RE) bands that are not available in L8 or aerial images (Table 1). Already before the launch of S2, researchers developed several new spectral vegetation indices that utilized its novel RE bands (Frampton et al., 2013; Verrelst et al., 2015). These indices were especially designed for monitoring agricultural fields. Early empirical results showed that in forests, the gains of using S2 images in the estimation of growing stock volume (Chrysafis et al., 2017; Mura et al., 2018) and leaf area index (Korhonen et al., 2017) were relatively small compared to using L8. On the other hand, the new bands may add more value for estimating tree species.

The objective of this study was to compare ALS data supplemented with aerial images, Sentinel-2 images or Landsat 8 images in the prediction of forest stand volume by tree species. Combined use of ALS and aerial or satellite image data was also compared to an option where only ALS data is used.

2. Material and pre-processing

2.1. Study area and field data

The study area comprises around 400 000 ha of managed boreal forests located in Eastern Finland. The dominant tree species in the study area are Scots pine and Norway spruce. Broadleaved trees, such as downy birch (*Betula pubescens* Ehrh.) and silver birch (*Betula pendula* Roth.) make up a relatively minor proportion of the forest. All the broadleaved tree species were lumped together in the analyses; this species group is hence referred to as broadleaved trees.

The field measurements were carried out during the summer of 2016. Field data consisted of 633 circular plots with a radius of either 5.64 m (sapling stands where $h > 1.3$ m and dbh < 8 cm), 12.62 m (mature forests where stem number < 800 per hectare) or 9.00 m (other development classes). Seedling stands ($h < 1.3$ m) were excluded. The study area was stratified based on earlier inventory data, and the field plots were allocated to the strata as L-shaped clusters (stratified cluster sampling). If the predetermined plot location crossed a stand border, the plot was moved the shortest possible distance to relocate it inside the stand. Trimble ProXRT global navigation satellite system with an external L1/L2 antenna elevated to 5 m was used to determine the plot locations.

The diameter at breast height (dbh) and tree species were recorded for each tree with dbh ≥5 cm. In most plots the height of one sample tree of each species from each story class was measured using an electronic hypsometer, and the heights of other trees were predicted using locally calibrated models by Eerikäinen (2009). The volumes of individual trees were calculated as a function of dbh and height using the species-specific models by Laasasenaho (1982). Finally, tree volumes were aggregated to the plot level by tree species. The mean characteristics of plot attributes are presented in Table 2. A detailed description of plot placement and sample plot measurements is found from field guide of the Finnish Forestry Centre (Metsäkeskus, 2016).

2.2. ALS data

ALS data were collected between April 30 and May 4, 2016 using a Leica ALS60 laser scanning system under leaf-off conditions. In this device the divergence $(1/e^2)$ of the laser beam ($\lambda = 1064$ nm) is 0.22 mrad. The test site was measured from an altitude of 2400 m above ground level (AGL) using a pulse repetition frequency of 98 400 Hz, field of view of 40 degrees, and a side overlap of about 20%. This resulted in a sampling density of about 0.8 measurements per square meter. The scanner may record multiple range measurements for each pulse, but in this study, only the first and last echoes were used. First echoes contained original echo categories “first of many” and “only”, and last echoes “last of many” and “only”. A DTM was generated from the ALS data. First, laser echoes were classified as ground and non-ground echoes using the method reported by Axelsson (2000) and then a raster DTM with a pixel size of 2.0 m was interpolated by Delaunay triangulation using the ground echoes. Finally, the DTM was subtracted from the ellipsoidal heights of laser echoes to scale the ALS data to the AGL.

2.3. Satellite images

We used near-simultaneous Sentinel-2 A MSI (S2) and Landsat 8 OLI (L8) images acquired at 09:22 am UTC on 22 August 2015 and at 09:40 am UTC on 24 August 2015, respectively. Solar angles were similar, 51.30° for the L8 image and 51.49° for the S2 image. The S2 image consists of two tiles (T35VNK and T36VUQ) taken from the same orbit (A000891). Tile T36VUQ was warped to the same projected coordinate system as tile T35VNK (UTM zone 35 N). There were some clouds and their shadows in the S2 image, which were delineated manually, and