Remote Sensing of Environment 147 (2014) 156–172

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Regional-scale mapping of tree cover, height and main phenological tree types using airborne laser scanning data

Cici Alexander [a,b,*], Peder Klith Bøcher [a,b], Lars Arge [b], Jens-Christian Svenning [a]

[a] Section for Ecoinformatics & Biodiversity, Department of Bioscience, Aarhus University, Ny Munkegade 114, DK-8000 Aarhus C, Denmark
[b] Center for MAssive Data ALGOrithmics (MADALGO), Department of Computer Science, Aarhus University, IT-parken, Aabogade 34, DK-8200 Aarhus N, Denmark

ARTICLE INFO

Article history:
Received 25 May 2013
Received in revised form 20 February 2014
Accepted 26 February 2014
Available online xxxx

Keywords:
Airborne laser scanning
LiDAR
Land cover classification
Tree cover
Tree height

ABSTRACT

Airborne Laser Scanning (ALS) data for generating Digital Terrain Models (DTM) at national level are often collected during leaf-off seasons. Leaf-off data are useful for classifying evergreen and deciduous trees since echoes at lower intensity are returned from deciduous trees when compared to evergreen trees. In addition to this, the proportion of echoes from the ground is higher for deciduous trees than for evergreen trees. In this study, we classified land cover, including evergreen and deciduous trees, using a Random Forest classifier based on LiDAR-metrics generated from leaf-off ALS data, and estimated tree cover and tree heights for the whole of Denmark. The results were compared with the CORINE Land Cover (CLC2006) data, percentage of tree cover from MODIS Vegetation Continuous Fields (VCF) and a global tree height map based on ICESat data. Considering tree class alone, deciduous and evergreen trees could be classified with an overall accuracy of 94% using validation data generated using aerial imagery from a 60-km strip across Central Jutland. The lower values of ALS-based percentage tree cover were overestimated and the higher values underestimated by MODIS VCF data, with a root–mean–square (RMS) deviation of 18.26%. The tree heights estimated using ALS data were generally lower than the global estimates of tree height with an RMS deviation of 5.1 m. The ALS intensity values were useful for classifying evergreen and deciduous trees. These findings show that ALS datasets collected for generating national DTMs can be used for tree cover and tree height mapping, as well as for regional classification of trees if data over the whole area are collected within a few months in the leaf-off season.

1. Introduction

Mapping of forests at the global level has gained importance in recent years due to the ability of trees to act as carbon sinks and thereby regulate climate change (Bonan, 2008). Trees also offer a range of other benefits including carbon sequestration, prevention of soil erosion and mitigation of floods as well as habitats for biodiversity. The expansion of urban areas and agricultural activities are often at the expense of tree cover, leading to fragmentation of wildlife habitats and loss of biodiversity (Phelps, Webb, & Adams, 2012; Sandel & Svenning, 2013; Strassburg et al., 2010). Mapping of tree cover and monitoring the changes in tree cover are therefore important for understanding ecological processes and devising effective conservation strategies. Remotely sensed images are a valuable source of data for the mapping of land cover and monitoring its change over time (Rogan & Chen, 2004).

Multi-spectral satellite images have been extensively used for mapping land cover (Hansen & Loveland, 2012). The vertical structure of forests at a global scale was mapped for the first time using data from the Geoscience Laser Altimeter System (GLAS) aboard the Ice, Cloud and land Elevation Satellite (ICESat) in 2010 (Lefsky, 2010), and improved in 2011 (Simard, Pinto, Fisher, & Baccini, 2011). GLAS footprints are approximately 65 m in diameter, and spaced at 170 m along track and several kilometres across tracks. Hence, the heights of forest patches where there are no observations were estimated using information from multi-spectral satellite imagery and globally available climate, elevation, and vegetation cover layers (Lefsky, 2010; Simard et al., 2011). Global vegetation height maps are therefore interpolated and inadequate for ecological studies at the local scale because of their low spatial resolution. Local tree cover and height maps could improve our understanding of ecological processes and patterns, including habitat fragmentation and species interactions, carbon sequestration in aboveground biomass and patterns of biodiversity (Vierling, Vierling, Gould, Martinuzzi, & Clawges, 2008).

Airborne Laser Scanning (ALS) uses the technique of Light Detection and Ranging (LiDAR) to estimate the elevations of points on and above the Earth's surface. The strength of the return signal, often referred to as

* Corresponding author at: AG Biologie/Ökologie, TU Bergakademie Freiberg, Leipziger Straße 29, 09599 Freiberg, Germany. Tel.: +49 3731 394154; fax: +49 3731 393012.
E-mail address: cici.alexander@gmail.com (C. Alexander).

http://dx.doi.org/10.1016/j.rse.2014.02.013

intensity, is also recorded providing an additional attribute of surface points. Generation of digital terrain models (DTMs) was one of the first applications of ALS, since laser beams can penetrate through gaps in vegetation to register points from below trees in forests (Flood, 2001). ALS data are increasingly being used for land cover mapping, often using elevation as an attribute in addition to spectral information from multi-spectral images (Arroyo, Johansen, Armston, & Phinn, 2010; Koetz, Morsdorf, Van Der Linden, Curt, & Allgöwer, 2008; Singh, Vogler, Shoemaker, & Meentemeyer, 2012; Yu, Cheng, Ge, & Lu, 2011).

Charaniya, Manduchi, and Lodha (2004) classified ALS data into trees, grass and building roofs with a classification accuracy of 66%–84%. In their study, the ALS data were interpolated to a regular grid, and classification was based on normalised height, height variation, multiple echoes, luminance and intensity. The luminance values were obtained from an additional grey scale aerial image. Miliaresis and Kokkas (2007) employed parametric classification and k-means clustering for the extraction of building and vegetation classes from LiDAR DEMs based on elevation, roughness, mean slope and standard deviation of the slope of grid cells belonging to a region. Antonarakis, Richards, and Brasington (2008) used elevation and intensity data in a method based on the frequency of point distribution to classify forest types and ages in flood plains.

ALS data have been used for creating digital terrain models (DTMs) by many countries, including the Netherlands, Denmark and Switzerland, on a national level, and many states in Germany and the USA (Hyyppa et al., 2007). If ALS data are collected in leaf-off season, more points are collected from the ground below trees than would be possible in leaf-on season, thereby increasing the accuracy of the generated DTM. Although canopy height and canopy cover estimated using leaf-off data are less accurate than those estimated using leaf-on data, for deciduous trees, they are nevertheless useful for mapping land cover and estimating biomass (Antonarakis et al., 2008; Nord-Larsen & Schumacher, 2012; Wasser, Day, Chasmer, & Taylor, 2013). Since many of the estimates are based on normalised elevations, or heights of points from the ground, improved accuracy of DTMs from leaf-off data could even be an advantage.

At the individual tree level, and using ALS data at a high point density, it has been shown that leaf-off data perform better (98%) than leaf-on data (90%) for separating spruce from birch and aspen trees (Ørka, Næsset & Bollandsås, 2010). A Random Forest classifier from predictors based on height and intensity from points within individual tree crowns was used to classify spruce and the deciduous trees. Attributes from full-waveform ALS acquired during leaf-off season have been shown to separate coniferous and deciduous trees with an accuracy of 96% (Reitberger, Krzystek, & Stilla, 2008). Liang, Hyyppä, and Matikainen (2007) used height differences between first and last echoes to classify coniferous and deciduous trees.

If the information contained in the ALS data points are aggregated in raster format, they can be made accessible to researchers in other disciplines such as environmental modelling (Vierling et al., 2008). Using ALS data collected for generating DTM for other purposes would allow sharing costs and lead to more efficient use of resources. The ALS dataset collected for the whole of Denmark in 2006–2007 for generating a national DTM has been used for estimating forest basal area, volume, aboveground biomass and total biomass in forested plots, which are part of the Danish National Forest Inventory (NFI) data (Nord-Larsen & Schumacher, 2012). It was noted that model predictions could be improved by classification into deciduous and evergreen forests. Here, we (i) aggregate information contained in discrete-return ALS data points in 10-m cells and (ii) classify the cells based on height and intensity attributes into pre-defined classes to explore (iii) whether evergreen and deciduous trees can be classified using low-density discrete return ALS data alone, and (iv) to generate high-resolution tree cover and tree height maps covering the whole country and (v) compare them with global tree height and tree cover maps generated from satellite-based LiDAR and multispectral data.

2. Methodology

LiDAR metrics, based on height and intensity, were calculated from the ALS dataset with a pixel size of 10 m. Random forests, an ensemble classifier which consists of decision trees, was used to classify the cells into five broad classes: water, open areas, shrubs, trees and buildings. Trees were further classified into evergreen and deciduous trees, and the accuracies of the results were evaluated.

Table 1
Description of the LiDAR metrics computed within 10-m cells for the whole of Denmark.

No.	LiDAR Metric	Description
1	PTDENS	Number of points within a 10 m × 10 m cell
2	HMIN	Minimum height
3	HMAX	Maximum height
4	HMEAN	Mean height
5	HSTD	Standard deviation of heights
6	HCV	Coefficient of variation of heights
7	CAN1	Percentage of points above 1 m
8	H05PCT	5th percentile height
9–17	H10PCT, H20PCT, … H90PCT	10th, 20th, … 90th percentile height
18	H95PCT	95th percentile height
19–28	P10, P20, … P100	Percentage of points within each bin when the heights are divided into ten bins; P10 is the number of points between 0 and 10% of maximum height, P20 between 10% and 20% of maximum height, and so on.
29	STRAT0	Percentage of points on the ground
30	STRAT1	Percentage of points > 0 and ≤1 m
31	STRAT2	Percentage of points > 1 m and ≤2.5 m
32	STRAT3	Percentage of points > 2.5 m and ≤5 m
33	STRAT4	Percentage of points > 5 m and ≤10 m
34	STRAT5	Percentage of points > 10 m and ≤20 m
35	STRAT6	Percentage of points > 20 m and ≤30 m
36	STRAT7	Percentage of points > 30 m
37	IMEAN	Mean intensity
38	ISTD	Standard deviation of intensity
39	IMEAN1	Mean intensity of points above 1 m
40	ISTD1	Standard deviation of intensity of points above 1 m