Environmental Modelling & Software 82 (2016) 142–151

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Deriving comprehensive forest structure information from mobile laser scanning observations using automated point cloud classification

Suzanne M. Marselis [a, b, *], Marta Yebra [b, c, d], Tom Jovanovic [d], Albert I.J.M. van Dijk [b, c, d]

[a] Department of Geographical Sciences, University of Maryland, United States
[b] Fenner School of Environment and Society, The Australian National University, ACT, Australia
[c] Bushfire & Natural Hazards Cooperative Research Centre, Melbourne, Australia
[d] CSIRO Land and Water, Canberra, ACT, Australia

ARTICLE INFO

Article history:
Received 20 November 2015
Received in revised form
22 March 2016
Accepted 25 April 2016

Keywords:
Ground-based LiDAR
Automatic classification
Vegetation components
Stem diameter

ABSTRACT

The advent of mobile laser scanning has enabled time efficient and cost effective collection of forest structure information. To make use of this technology in calibrating or evaluating models of forest and landscape dynamics, there is a need to systematically and reproducibly automate the processing of LiDAR point clouds into quantities of forest structural components. Here we propose a method to classify vegetation structural components of an open-understorey eucalyptus forest, scanned with a 'Zebedee' mobile laser scanner. It detected 98% of the tree stems (N = 50) and 80% of the elevated understorey components (N = 15). Automatically derived DBH values agreed with manual field measurements with $r^2 = 0.72$, RMSE = 3.8 cm, (N = 27), and total basal area agreed within 1.5%. Though this methodological study was restricted to one ecosystem, the results are promising for use in applications such as fuel load, habitat structure, and biomass estimations.

1. Introduction

Detailed information on three-dimensional forest structure is used in various fields of research, e.g., habitat characterisation and species diversity for ecological applications (Tews et al., 2004); fuel distributions for risk assessment (McKenzie et al., 2004); ecosystem dynamics (Kucharik et al., 2000), and CO_2 and climate change research (Cramer et al., 2001). Deriving such detailed information by conventional means requires extensive fieldwork, time, and expert knowledge. However, over the last decade, significant advancements have been made in the use of active and passive remote sensing data to measure forest structure (Lefsky et al., 2002; Wang et al., 2010; Xie et al., 2008).

Light Detection and Ranging (LiDAR), in particular, has shown potential for forest structure assessment, as it can provide detailed three-dimensional data on all reflecting elements within the forest (Lefsky et al., 2002; Vierling et al., 2008). Airborne LiDAR data has successfully been used to derive information on the canopy (e.g. canopy height, canopy cover, canopy base height, crown volume and foliage biomass) in a variety of forests including coniferous and deciduous forests, Mediterranean mixed forests and riparian forests (Andersen et al., 2005; García et al., 2015; Riaño et al., 2004; Wasser et al., 2015). However, these approaches have two limitations, (1) the point clouds are classified into different vegetation components only based on height thresholds (in this paper referred to as 'conventional classification') possibly creating artefacts (Hernández et al., 2013; Lefsky et al., 2002; Martinuzzi et al., 2009; Whitehurst et al., 2013; Wing et al., 2012) and (2) The data collected generally lacks comprehensive information on lower vegetation layers and trunks due to the point of view, scanning angle and footprint of airborne LiDAR systems (Dassot et al., 2011; Jakubowski et al., 2013; Lovell et al., 2003; Yebra et al., 2015).

The application of terrestrial LiDAR data collection for forest structure mapping has also been investigated and shown to provide information on stem and canopy characteristics (Beland et al., 2014; Côté et al., 2011; Hopkinson et al., 2004; Raumonen et al., 2013; Watt and Donoghue, 2005; Yao et al., 2011) and, to a lesser extent, understorey vegetation structures (Richardson et al., 2014). However, these studies use fixed-point terrestrial laser scanners which have a number of disadvantages: inflexibility, long scanning times and obscuring effects caused by tree stems (Strahler et al., 2008; Watt and Donoghue, 2005) The latter means that either

* Corresponding author. 1150 Lefrak Hall, University of Maryland, College Park, MD 20742, United States.
E-mail address: marselis@umd.edu (S.M. Marselis).

http://dx.doi.org/10.1016/j.envsoft.2016.04.025

information cannot be derived when trees are hidden behind others, or that multiple scans need to be done from different positions within the plot, which further increases scanning time and costs (Côté et al., 2011; Dassot et al., 2011; Strahler et al., 2008).

Potentially overcoming some of these issues, a mobile handheld LiDAR system, Zebedee, can be used. The instrument uses laser scanning to rapidly create a point cloud, by co-registration of points acquired from multiple locations (Bosse et al., 2012). Zebedee has been used in a small number of studies, including cave and mine mapping (Zlot and Bosse, 2014a,b), heritage mapping (Zlot et al., 2014), and forest inventory survey (Ryding et al., 2015). In comparison to fixed-point terrestrial LiDAR systems, Zebedee offers important benefits through the greater speed and flexibility of data collection (Ryding et al., 2015). Mobile laser scanners like Zebedee have been used successfully to extract single tree stems for forest assessments (Calders et al., 2015), and to calculate diameter at breast height (DBH) by fitting cylinders on the part of stems representing breast height (Ryding et al., 2015). However, so far, the data processing methods that have been applied are either not, or only partially, automated. This adds effort and uncertainty in the processing and requires expert knowledge. Furthermore, each of the applications on mobile laser scanning have so far focused on only one aspect of the vegetation in a plot, such as canopy cover fraction or stem basal area (Beland et al., 2014; Calders et al., 2015; Ryding et al., 2015). By contrast, many forest management applications require information about all vegetation layers, such as near-surface and elevated understorey vegetation (shrubs). Examples include fuel characterisations (Gould et al., 2011), forest succession characterisation (Falkowski et al., 2009) and habitat characterisation for bird diversity prediction (MacArthur and MacArthur, 1961).

The aim of this study is to demonstrate the possibility to derive comprehensive and internally consistent information on different vegetation components of an open eucalypt forest with sparse understorey from ground-based LiDAR point cloud data. To this end a single and automated data processing method was developed to simultaneously achieve three outcomes:

1) Classification of a Zebedee derived point cloud into ground returns, and four different vegetation components: near-surface vegetation, elevated understorey vegetation, tree trunks and tree canopy.
2) Locating individual tree stems and elevated vegetation objects based on the classified point cloud.
3) Derive relevant summary statistics for the classified vegetation components.

2. Materials and methods

2.1. Study area and data collection

The research was performed at Mulligans Flat Nature Reserve (Australian Capital Territory, Australia), in a Yellow Box – Red Gum, Grassy Woodland environment *(http://www.mulligansflat.org.au/history.html)*. These vegetation communities are commonly characterised by an open canopy, where either Yellow Box *(Eucalyptus melliodora)* or Blakely's Red Gum *(Eucalyptus blakelyi)* are usually present, dominant or co-dominant, together with an understorey of native tussock grasses, herbs and shrubs (Boland et al., 2006). The study area was chosen to well-represent the sparse understorey eucalypt vegetation in the nature reserve. The study area contained many small gum trees (DBH 5–35 cm) providing an open canopy and a number of elevated understorey vegetation (shrubs and short, young trees, centre-right in Fig. 1) and tussock grasses (right bottom corner, Fig. 1) between the trees.

On 2nd May 2014, ground-based LiDAR data was collected with the Zebedee at the study site. Zebedee is a light-weight mobile laser scanner (0.7 kg) with a 270° field of view and a maximum range of 30 m, collecting LiDAR points with sub-centimetre accuracy. The data collection is dependent on the movement of the scanning head of the system, mounted on a spring and requiring consistent motion for optimal data collection (Bosse et al., 2012; Ryding et al., 2015). Data collection took approximately 3 min, using Zebedee in a closed-loop scanning pattern with multiple passes through the plot, to collect information from all sides of the trees (starting at −35.164007°S, 149.181635°E). This scanning strategy ensured that most of the trees were scanned from different angles and points reflected on all sides of the stems were obtained. The primary scan data were uploaded to a designated web service and automatically pre-processed in 15 min, returning a point cloud file, data collection trajectory and diagnostic scan information (e.g. rate of oscillation) that can assist data quality control and analysis. The point density of the Zebedee point cloud is lower near the edges of a vegetation patch, because of the occlusion by trees and the range of the laser. A 14 × 25 m horizontal area was clipped from the point cloud data, resulting in a study area with consistent point density, witha total of 1,379,173 points. The average point density in the dataset was 3937 points per square meter but varied widely, depending on the presence of reflecting objects.

The locations of all the trees within the 14 × 25 m area were noted in the field. DBH values were measured in the field for a subset of 28 out of 50 trees in the scanned area. The measured trees were randomly selected from the subset of trees that were completely represented in the LiDAR point cloud, i.e. they were not on the edge of the scanned area. 15 shrubs and small trees of which the canopy started below 2 m were also located in the field plot as elevated understorey vegetation objects.

2.2. Data pre-processing

Ground returns were classified with the default version of the *lasground* function in software package LAStools (Isenburg, 2014). Subsequently the height above ground surface was computed for the remaining “non-ground” points. Each non-ground point was assigned a vegetation class based on height above ground surface (conventional classification, Table 1). The classes assigned were: low vegetation (0–0.3 m), medium vegetation (0.3–2 m) and high vegetation (>2 m), according to the conventional vegetation classification used for airborne LiDAR data as defined by the Intergovernmental Committee on Surveying and Mapping (ICSM, 2010).

2.3. New classification algorithm

All non-ground returns were reclassified by the new classification algorithm based on the structural characteristics of the vegetation objects. The classes in the new classification were based on structural forest layers generally distinguished in forest research: near-surface vegetation (understorey), elevated understorey vegetation (shrubs/midstorey) and tree canopy (overstorey) (Gould et al., 2011; Whitehurst et al., 2013) and tree trunks. Table 1 lists the structural characteristics defining these four different classes. The new classification algorithm consisted of five different steps which are outlined in the following subsections and summarised in the workflow in Fig. 7 and Table 2.

2.3.1. Step I – classification of trunks and elevated vegetation objects

All points with a height above ground between 1.25 and 1.35 m (resulting in a set of points named *layer 1*), and 1.9 and 2.0 m