International Journal of Applied Earth Observation and Geoinformation 35 (2015) 284–293

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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

Sub-footprint analysis to uncover tree height variation using ICESat/GLAS

Jinyan Tian[a], Le Wang[b,*], Xiaojuan Li[a]

[a] Capital Normal University, Key Lab of 3D Information Acquisition and Application MOE, Beijing 100048, China
[b] Department of Geography, State University of New York at Buffalo, Buffalo, NY 14261, USA

ARTICLE INFO

Article history:
Received 25 June 2014
Accepted 23 September 2014
Available online 14 October 2014

Keywords:
GLAS
Tree height
Sub-footprint analysis
Levenberg–Marquardt
Gaussian decomposition

ABSTRACT

Detailed forest height data are an indispensable prerequisite for many forestry and earth science applications. Existing research of using Geoscience Laser Altimeter System (GLAS) data mainly focuses on deriving average or maximum tree heights within a GLAS footprint, i.e. an ellipse with a diameter of 65 m. However, in most forests, it is likely that the tree heights within such ellipse are heterogeneous. Therefore, it is desired to uncover detailed tree height variation within a GLAS footprint. To the best of our knowledge, no such methods have been reported as of now. In this study, we aim to characterize tree heights' variation within a GLAS footprint as different layers, each of which corresponds to trees with similar heights. As such, we developed a new method that embraces two steps: first, a refined Levenberg–Marquardt (LM) algorithm is proposed to decompose raw GLAS waveform into multiple Gaussian signals, within which it is hypothesized that each vegetation signal corresponds to a particular tree height layer. Second, for each layer, three parameters were first defined: Canopy Top Height (CTH), Crown Length (CL), and Cover Proportion (CP). Then we extracted the three parameters from each Gaussian signal through a defined model. In order to test our developed method, we set up a study site in Ejina, China where the dominant specie is *Populus euphratica*. Both simulated and field tree height data were adopted. With regard to the simulation data, results presented a very high agreement for the three predefined parameters between our results and simulation data. When our methods were applied to the field data, the respective R^2 become 0.78 (CTH), CL ($R^2 = 0.76$), CP ($R^2 = 0.74$). Overall, our studies revealed that large footprint GLAS waveform data have the potentials for obtaining detailed forest height variation.

1. Introduction

Reliable acquisition of forest structure information has been considered as an indispensable prerequisite for numerous critical environmental issues (NRC, 2007). Among all the forest structure variables, tree heights play a vital role in estimating tree biomass, forest volume and carbon stock over large regions (Nelson et al., 2009a). Light detection and ranging (Lidar) carried on both spaceborne and airborne platforms provides a great potential for direct measurements of forest heights (Dolan et al., 2011; Lee et al., 2011). Successes have been reported in many studies, in which airborne Lidar served as the core data for characterizing forest vertical information (Anderson et al., 2008; Harding and Carabajal, 2005; Lefsky et al., 2002a, 2002b; Nelson et al., 2008; Popescu and Zhao, 2008). However, when large area forest monitoring is needed, it is often logistically difficult to use airborne Lidar. To this end, Geoscience Laser Altimeter System (GLAS), the first spaceborne Lidar system, which was initially designed to measure ice sheet mass balance cloud and aerosol heights (Zwally et al., 2002), demonstrates its advantage in collecting repetitive and extensive forest data.

The history of forest height extraction with GLAS data can be traced back to Lefsky et al. (2005). Then a number of studies have been reported toward tree height extraction (Chen, 2010a, 2010b; Duncanson et al., 2010; Fatoyinbo and Simard, 2013; Hayashi et al., 2013; Iqbal et al., 2013; Wang et al., 2013a), and aboveground forest biomass (Ballhorn et al., 2011; Fatoyinbo and Simard, 2013; Boudreau et al., 2008; Helmer et al., 2009; Lefsky et al., 2005; Lefsky, 2009; Nelson et al., 2009a; Nelson, 2010; Popescu et al., 2011). In addition, forest volume (Nelson et al., 2009b) and carbon (Nelson, 2010) were also studied. These studies confirmed that GLAS waveform can be employed to uncover forest vertical structure. Chen (2010b) grouped all the methods to two categories:

* Corresponding author at: University at Buffalo, Geography Department, 105 Wilkeson Quad, Buffalo, NY 14261, USA. Tel.: +1 716 645 0474; fax: +1 716 645 2329.
E-mail address: lewang@buffalo.edu (L. Wang).

http://dx.doi.org/10.1016/j.jag.2014.09.016
0303-2434/

statistical and direct methods. The statistical methods acquire forest height by establishing a model between GLAS waveform metrics and ground truth data normally collected either in the field or through airborne Lidar. The most representative waveform metrics is waveform extent (Lefsky et al., 2005), which is defined as the distance between the start and end of the GLAS signal. Surrounding this idea, statistical methods have been applied in different types of forests. Rosette et al. (2008) and Pang et al. (2008) achieved good results with the statistical method to forest areas in dense canopy region respectively. On the other hand, Nelson (2010) compared the canopy height derived from airborne Lidar and GLAS in a study site with sparse and short canopy, and found the result of GLAS is hard to be in line with that from airborne Lidar. In addition, Chen (2010b) explored tree height extraction in both dense forest and woody savanna. In order to reduce the errors, he suggested that the footprint size should be less than 10 m. However, it is normally difficult to find a universally applied statistical model that can fit different study sites with various vegetation and terrain conditions.

The second group of methods for extracting tree heights is the direct method. The earliest work of the direct method aimed to estimate canopy height by the distances between signal start (canopy top) and ground peak (Neuenschwander et al., 2008). Subsequently, for flat area, Gaussian decomposition method was used to detect the Ground peak, assuming the lowest peak representing the ground elevation (Duong et al., 2008; Neuenschwander et al., 2008). Alternatively, when complex topography was encountered, Rosette et al. (2008) and Chen (2010b) have reported that the stronger one between the two lowest Gaussian peaks has a better correspondence to ground. Afterwards, Sun et al. (2008) compared the forest heights derived from the quartile waveform energy of GLAS and Laser Vegetation Imaging Sensor (LVIS) data, respectively, found that 50% and 75% waveform energy showed higher correlations than 100% waveform energy. Although the direct method tends to overestimate canopy height over complex terrain (Chen, 2010b), it is easier to implement than the statistical method.

It should be noted that both the statistical and the direct method attempts to tackle average or maximum tree heights within a GLAS footprint. Nevertheless, in most forests, it is likely that the tree heights within the GLAS footprintare heterogeneous. Therefore, in this study, our objectives are twofolds. First, we aim to propose a sub-footprint analysis (SFA) concept. Second, we aim to develop a new direct method for uncovering tree height variation over a flat region.

2. Study area and data preparation

2.1. Study area

The study site is located in the Ejina oasis in northwest of Inner Mongolia, China (97°10′ E to 103°7′ E, 39°52′ N to 42°47′ N). The Ejina oasis covers an overall area of 11.46 km^2 and is a detached island encompassed by desert. The elevation of land surface varies from 820 to 1127 m above sea level, with an average slope of less than 3°. Due to the north temperate continental arid climate, Ejina has an extremely harsh natural environment, and its climate is characterized by a large seasonal temperature range. The mean annual temperature is approximately 8 °C, with a minimum temperature of −36 °C in winter and a maximum temperature of 42 °C in summer. The annual natural precipitation of this region is only 42 mm and the mean annual pan evaporation is 3755 mm. The predominant species of the region mainly include *Populus euphratica*, *Tamarix ramosissima*, and *Sophora alopecuroides*, which all depends on groundwater derived from the Heihe River for sustenance. These vegetation functions as the first barriers protecting northwest China from sandstorms (Fig. 1).

2.2. GLAS data

The Ice, Cloud, and land Elevation Satellite (ICESat) was launched by NASA on January 12, 2003, which is a significant milestone in the history of Earth Observation as it is equipped with a large footprint full waveform laser radar sensor (GLAS) onboard satellite for the first time. In this study, 709 GLAS full waveform data were downloaded from the National Snow and Ice Data Center (NSIDC). The abnormal data influenced by cloud were excluded by the condition of the maximum intensity value of waveform under 0.1 (Lefsky et al., 2005). The original binary data GLA01 and GLA14 were converted into American Standard Code for Information Interchange (ASCII) with the IDL Readers routine provided by NASA (http://nsidc.org/data/icesat/tools.html). In addition, for validation purposes, we used one QuickBird image acquired on July 25, 2012, and collected data from 22 field observations plots (footprints), each of which is approximately treated as a circle with 65 m diameters, where *P. euphratica* is the dominant species. Of the 22 plots, nine spots are close to the water source and others not (Fig. 2). Water sources are important for the growth of *P. euphratica*. Generally, the *P. euphratica* close to the water grow straighter and have a more regular canopy. GLAS data were recorded in November 2003, 2005, 2006 and in March, 2004, 2005, 2007. The field tests were conducted in November 2013. Because this area is one of the world's three *P. euphratica* natural reserves, the *P. euphratica* here is well protected, and the height change per year is extremely small because they are all a few hundred years' old trees.

2.3. Field data

In order to prepare the ground truth data for validation, we collected field data by measuring relevant parameters at individual tree level, including the Canopy Top Height (CTH), Crown Length (CL), and Cover Proportion (CP). With the help of a QuickBird image and a hand-held high-precision GPS, we located the center of the GLAS footprint. In order to pinpoint each tree within the GLAS footprint, first, a plot is equally divided into eight sections with a 32.5 m long line radiating away from the center of the plot. Then all the trees were labeled following the ascending order of their distance to the plot center. For each of the individual tree, we measured CTH and CL with a laser altimeter (Criterion RD1000). The geographic coordinate was also recorded with a handheld GPS. On the other hand, CP is defined as the proportion of an individual tree to all the trees residing in one footprint with regards to canopy area. It was measured by first manually delineating all the individual trees on QuickBird image followed by a ratio calculation. We found that CTH ranges from 5 to 16 m and CL ranges from 3 to 13 m. A high correlation of R^2 value (0.89) was presented in Fig. 3 when a total of 79 tree measurements were randomly selected. Then all the rest of the trees were verified by the regression model, and the results confirmed that CTH and CL does satisfy a linear relationship.

3. Methodology

Three key steps were engaged in our developed method for uncovering tree height variations within a GLAS footprint. They are: GLAS data processing, a refined Levenberg–Marquardt (LM) algorithm for Gaussian decomposition, and extraction of three parameters (CTH, CL, and CP) from each Gaussian signal. Firstly, the original GLAS waveform was filtered by the Savitzky–Golay method (Wang et al., 2013b), signal location was detected, and then the peaks were identified between signal beginning and end of the filtered waveform by using second-order derivative. Secondly, we developed a refined LM (RLM) algorithm to decompose the original