International Journal of Applied Earth Observation and Geoinformation 14 (2012) 160–168

Contents lists available at SciVerse ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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

Estimation of forest above-ground biomass using multi-parameter remote sensing data over a cold and arid area

Xin Tian [a,b], Zhongbo Su [b], Erxue Chen [a], Zengyuan Li [a,*], Christiaan van der Tol [b], Jianping Guo [c], Qisheng He [a]

[a] Research Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Wanshoushanhou, 100091 Beijing, PR China
[b] Faculty of Geo-Information Science and Earth Observation, University of Twente, Hengelosestraat 99, 7500 AA Enschede, The Netherlands
[c] Centre for Atmosphere Watch and Services, Chinese Academy of Meteorological Sciences, No 46, Zhongguancun South Avenue, 100081 Beijing, PR China

ARTICLE INFO

Article history:
Received 1 December 2010
Accepted 9 September 2011

Keywords:
k-NN method
Regression method
Above-ground biomass
Configuration

ABSTRACT

Remote sensing is a valuable tool for estimating forest biomass in remote areas. This study explores retrieval of forest above-ground biomass (AGB) over a cold and arid region in Northwest China, using two different methods (non-parametric and parametric), field data, and three different remote sensing data: a SPOT-5 HRG image, multi-temporal dual-polarization ALOS PALSAR and airborne LiDAR data. The non-parametric method was applied in 300 different configurations, varying both the mathematical formulation and the data input (SPOT-5 and ALOS PALSAR), and the quality of the performance of each configuration was evaluated by Leave One Out (LOO) cross-validation against ground measurements. For the parametric method (the multivariate linear regression), the same remote sensing data were used, but in one additional configuration the airborne LiDAR data were used for stepwise multiple regression.

The result of the best performing non-parametric configuration was satisfactory (R = 0.69 and RMSE = 20.7 tons/ha). The results for the parametric method were notoriously inaccurate, except for the case where airborne LiDAR data were included. The regression method with airborne low density LiDAR point cloud data was the best of all tested methods (R = 0.84 and RMSE = 15.2 tons/ha). A cross comparison of the two best results showed that the non-parametric method performs nearly as well as the parametric method with LiDAR data, except for some areas where forests have a very heterogeneous structure. It is concluded that the non-parametric method with SPOT data is able to map forest AGB operatively over the cold and arid region as an alternative to the more expensive airborne LiDAR data.

1. Introduction

Describing and quantifying forest above-ground biomass (AGB) has become of importance to many scientific and societal tasks such as sustainable forest management, timber management, forest ecosystem productivity estimation, carbon sink evaluation, and studies of the role of forest in the global carbon cycle, and links between hydrology and ecology, etc. (Houghton et al., 1999, 2001; Foody, 2003; Zheng et al., 2004; Houghton, 2005; Palacios-Orueta et al., 2005; Lu, 2006). In conventional techniques on basis of statistical assessment (i.e., tree species, vertical structure, stand height, and stand density), the forest AGB information comes from expensive and time consuming field surveys by the high sampling intensity (Schroeder et al., 1997; Nelson et al., 1999; Parresol, 1999; Brown, 2002; Fournier et al., 2003; Woodbury et al., 2007). As an alternative, multi-parameter remote sensing techniques have been applied, in which remotely sensed data are used as proxy for quantitative forest AGB at various scales (Nelson et al., 1988; Le Toan et al., 1992; Ranson and Sun, 1994; Roy and Ravan, 1996; Ranson et al., 1997; Santos et al., 2002, 2003; De Jong et al., 2003; Coops et al., 2004; Luther et al., 2005; Thenkabail et al., 2004; Labrecque et al., 2006; Lu et al., 2005; Patenaude et al., 2005; Leboeuf et al., 2007; Tan et al., 2007; Lucas et al., 2008; Gonzalez et al., 2010).

However, the application of remote sensing data for the forest AGB estimation is complicated in heterogeneous forests, where the tree species, stand density, canopy closure and crown overlap are variable. All these variables affect AGB, but it is difficult to differentiate them with remote sensing. In addition, topographic irregularities and various environmental conditions also affect the remote sensing signatures of the observed surfaces (Danson and Curran, 1993; Nelson et al., 2000a; Steininger, 2000; Foody et al., 2001; Saatchi et al., 2007). Both passive and active remote sensing measurements have been used for the estimation of forest AGB, each having different advantages and disadvantages.

Some passive remote sensing techniques make use of the sensitivity of the reflectance of forests to parameters like canopy cover,

* Corresponding author. Tel.: +86 1062889163; fax: +86 1062888315.
E-mail address: zy@caf.ac.cn (Z. Li).

0303-2434/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jag.2011.09.010

crown closure, leaf area index, which are in turn correlated with AGB. Many studies have investigated the relationship between forest AGB and satellite spectral values expressed as digital numbers, radiance, reflectance, or vegetation indices. Sensors with intermediate spatial resolution like Satellite Probatoire d'Observation de la Terre (SPOT), Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) data are commonly used for this purpose. Various correlations between forest AGB and the reflectance or vegetation indices were found and reported (Sader et al., 1989; Roy and Ravan, 1996; Nelson et al., 2000b; Steininger, 2000; Mutanga and Skidmore, 2004; Zheng et al., 2004; Hall et al., 2006; Labrecque et al., 2006; Rahman et al., 2008; Powell et al., 2008; Soenen et al., 2010).

Some active remote sensing techniques make use of the relation between Synthetic Aperture Radar (SAR) backscatter and forest AGB (Dobson et al., 1992; Le Toan et al., 1992; Ranson and Sun, 1994; Ranson et al., 1997; Luckman et al., 1997; Kasischke et al., 1997; Kurvonen et al., 1999; Kuplich et al., 2000; Yrjo, 2005). The advantage of SAR is that the signal is less prone to saturation with AGB than optical reflectance. Not all SAR bands are equally suitable: SAR backscatter in the P and L bands is has a much stronger correlation with forest AGB than the C and X bands. The L-band has proven particularly valuable for AGB estimation (Lu, 2006). Some studies (Moghaddam et al., 2002; Amini and Sumantyo, 2009; Chen et al., 2009) found that the accuracy of forest AGB estimation improved when optical reflectance data were added to the SAR data as an additional source of information.

The single most promising technique in active remote sensing is airborne laser scanning (ALS). ALS has recently led to a revolution in remote sensing of forest structure (Lefsky et al., 2002a; Popescu and Wynne, 2004; Maltamo et al., 2004; Reutebuch et al., 2005; Koukoulas and Blackburn, 2005; Salas et al., 2010), due to its unique ability to measure three-dimensional information (Nelson et al., 1988; Lefsky et al., 2002b; Lim and Treitz, 2004; Patenaude et al., 2004; Næsset, 2007; Lucas et al., 2008; Næsset and Gobakken, 2008; García et al., 2010).

In summary, there are several options for estimating forest AGB among the multi-parameter remote sensing data, differing in cost, spatial resolution, applicability and continuity. Optical remote sensing data is sensitive to the forest AGB, but the spectral signature saturates at a lower biomass level than the SAR data. SAR data are free of weather disturbance, but largely affected by topography and the quality depends on the waveform length (Sun et al., 2002). Airborne LiDAR is the most efficient technique to estimate forest AGB. It is also the most expensive technique, which makes it unsuited for routine applications.

Two methods exist for establishing (calibrating) the relation between remote sensing data and forest AGB with ancillary data: parametric and non-parametric. The parametric method employs multiple regressions, assuming that each forest attribute affects remote sensing signals at each wavelength in a unique and uncorrelated way (Lawrence and Ripple, 1998; Maselli and Chiesi, 2006; Rahman et al., 2008). The method is conceptually simple, but the success largely depends on the statistical robustness of the relationships. In reality a change in AGB rarely directly results in a change of remote sensing signatures. Consequently, the conventional parametric method applied to remote sensing data usually fails to map forest AGB satisfactorily (Maselli and Chiesi, 2006), independent on the univariate or multivariate type of regression that is used.

The non-parametric k-nearest neighbors (k-NN) method is based on more flexible assumptions than the parametric method, and does not suffer from the same limitations. In the non-parametric method the AGB of a pixel is associated with that of the k 'nearest' pixels in terms of remote sensing observations for which field data of AGB are available. A training set of pixels for which both field measured AGB and remote sensing data are available is established first. This training set is further used to estimate the AGB of all other pixels by linking them to pixels with similar remote sensing data in the training set. A distance measure d is defined for the differences between the remote sensing data of a pixel and the data in the training set. The k elements in the training set with the smallest values of d are used to compute AGB of that pixel as a weighted average. The method requires a representative set of ground data, covering the complete range of values for AGB, forest structure and remote sensing signals in the study area. The advantage of this method is that no assumptions are made about the nature of the relationship between AGB and remote sensing. This makes the method more robust in complex environments, and reasonable results can be achieved (Maselli et al., 2005). The k-NN method is an attractive tool for multi-source forest inventory (Franco-Lopez et al., 2001; Katila and Tomppo, 2001; Holmström and Fransson, 2003; Tomppo and Halme, 2004; Tomppo et al., 2009), and it has been operationally applied for various scale forest inventories in many countries (Fazakas et al., 1999; Holmström et al., 2001).

However, most applications of k-NN used medium resolution remote sensing data such as Landsat TM/ETM+ images (Franco-Lopez et al., 2001; Katila and Tomppo, 2001; Finley et al., 2006). Only few studies have used fine resolution imagery such as QuickBird, aerial photographs (Muinonen et al., 2001; Tuominen and Pekkarinen, 2005; Fuchs et al., 2009) and airborne SAR data (Holmström and Fransson, 2003). Since SPOT-5 data, with intermediate spatial resolution, is widely available, it is our interest to investigate the applicability of SPOT-5 with combination of current operative L-band ALOS PALSAR data for mapping Forest AGB by k-NN algorithm.

This study investigated the performances of parametric and non-parametric methods for forest AGB by use of multi-parameter remote sensing data over a complex forest area in upper reaches of Heihe River Basin, a cold and arid region, where only limited forest inventory data are available. The objectives of this study are (1) to evaluate the predictive power of reflectance (SPOT-5) and SAR backscatter intensity (ALOS PALSAR) for the retrieval of AGB, (2) to evaluate different configurations of the non-parametric algorithm at pixel level, (3) to compare the parametric versus the non-parametric method, and (4) to identify the best k-NN algorithm for routine application of forest AGB mapping in a heterogeneous forest over a cold and arid region. For these objectives supporting ground data and data of an airborne LiDAR system have been used.

The paper is organized as follows. After describing the test area, the data and models are presented. This is followed by a description of the optimization procedure for various k-NN configurations by leave-one-out (LOO) cross-validation and the assessments of pixel level estimation accuracy for both k-NN and regression models. The best results from k-NN and regression model will compared, and the results presented and discussed.

2. Site observation

In this study we used the data collected in the framework of the Watershed Allied Telemetry Experimental Research (WATER) carried out at the Heihe River Basin in Northwest China, in 2008 (Li et al., 2009, 2011). Heihe River Basin, the second largest inland river basin in China, is located between 97°24′–102°10′E and 37°41′–42°42′N, with an area of about 130,000 km^2. It consists of three major geomorphic units: the southern Qilian Mountains, the middle Hexi Corridor, and the northern Alxa Highland and accordingly, the landscapes are various, including glacier, frozen soil, alpine meadow, forest, irrigated crops, riparian ecosystem, desert (Gobi) (Fig. 1). Aiming to improve the understanding of physical processes of the land surface-atmosphere interaction at