Remote Sensing of Environment 117 (2012) 366–380

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

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

Capabilities and limitations of Landsat and land cover data for aboveground woody biomass estimation of Uganda

Valerio Avitabile [a,*], Alessandro Baccini [b], Mark A. Friedl [c], Christiane Schmullius [a]

[a] Institute of Geography, Friedrich-Schiller-University, Grietgasse 6, 07743 Jena, Germany
[b] The Woods Hole Research Center, 149 Woods Hole Road, Falmouth, MA 02540, USA
[c] Department of Geography and Environment, Boston University, 675 Commonwealth Avenue, Boston, MA 02215, USA

ARTICLE INFO

Article history:
Received 4 December 2010
Received in revised form 5 October 2011
Accepted 8 October 2011
Available online 10 November 2011

Keywords:
Biomass
Carbon
Landsat
Uganda
Forest
Land cover
LiDAR
GLAS
Regression tree
Random forest
REDD+

ABSTRACT

Aboveground woody biomass for circa-2000 is mapped at national scale in Uganda at 30-m spatial resolution on the basis of Landsat ETM+ images, a National land cover dataset and field data using an object-oriented approach. A regression tree-based model (Random Forest) produces good results (cross-validated R^2 0.81, RMSE 13 T/ha) when trained with a sufficient number of field plots representative of the vegetation variability at national scale. The Random Forest model captures non-linear relationships between satellite data and biomass density, and is able to use categorical data (land cover) in the regression to improve the results. Biomass estimates were strongly correlated ($r=0.90$ and $r=0.83$) with independent LiDAR measurements. In this study, we demonstrate that in certain contexts Landsat data provide the capability to spatialize field biomass measurements and produce accurate and detailed estimates of biomass distribution at national scale. We also investigate limitations of this approach, which tend to provide conservative biomass estimates. Specific limitations are mainly related to saturation of the optical signal at high biomass density and cloud cover, which hinders the compilation of a radiometrically consistent multi-temporal dataset. As a result, a Landsat mosaic created for Uganda with images acquired in the dry season during 1999–2003 does not contain phenological information useful for discriminating some vegetation types, such as deciduous formations. The addition of land cover data increases the model performance because it provides information on vegetation phenology. We note that Landsat data present higher spatial and thematic resolution compared to land cover and allow detailed and spatially continuous biomass estimates to be mapped. Fusion of satellite and ancillary data may improve biomass predictions but, to avoid error propagation, accurate, detailed and up-to-date land cover or other ancillary data are necessary.

1. Introduction

Tropical forests are attracting increasing attention within the climate change debate because of their crucial role as carbon sinks (Stephens et al., 2007) and the large emissions associated with their disappearance (Houghton, 2007; Le Quéré et al., 2009). In Africa, deforestation and forest degradation are an especially large source of greenhouse gas emissions (Bombelli et al., 2009; Canadell, Raupach, & Houghton, 2009; Houghton & Hackler, 2006).

To address this, a new mechanism is currently under negotiation by the United Nations Framework Convention on Climate Change (UNFCCC) to provide financial incentives to tropical countries to reduce emissions from deforestation and forest degradation (REDD) below a historical reference levels (Gullison et al., 2007). The implementation of such a carbon market poses several challenges (e.g. political, economical, social) (Stickler et al., 2009), but first requires the accurate assessment of the forest carbon stocks and their dynamics at national level (UNFCCC, 2009). Spatial assessment of the amount and dynamics of woody aboveground biomass (AGB) is also crucial for national planning in many tropical countries, where local communities rely on woody vegetation as a primary source of products (e.g. timber, fodder) and energy (fuelwood). Since monitoring biomass resources is an expensive and time consuming task, few developing countries have in place efficient monitoring systems. As a result, current biomass estimates for these countries are highly variable (Gibbs, Brown, Niles, & Foley, 2007; Houghton, 2005; Houghton, Lawrence, Hackler, & Brown, 2001), and sub-Saharan Africa is one of the regions with the highest uncertainty (Houghton & Hackler, 2006).

Remote sensing provides the key source of data for updated, consistent and spatially explicit assessment of biomass and its dynamics, especially in large countries with limited accessibility (DeFries et al., 2007; GOFC-GOLD, 2009; Herold & Johns, 2007; Penman et al., 2003; UNFCCC, 2007; UNFCCC, 2008). Application of remote sensing to tropical forests is particularly challenging because of complex and variable forest structure (Lu, 2005; Nelson, Kimes, Salas, & Routhier, 2000; Steininger, 2000) and difficulties obtaining high quality remote

* Corresponding author. Tel.: +49 3641 948881; fax: +49 3641 948882.
E-mail address: vale.avi@gmail.com (V. Avitabile).

0034-4257/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2011.10.012

sensing and corresponding ground data sets (Foody, Boyd, & Cutler, 2003). To this end, a number of approaches have been proposed to map AGB from satellite observations, relating the remotely sensed data directly (mapping biomass stock) or indirectly (mapping forest cover) to biomass density (reviewed by Boyd & Danson, 2005; Lu, 2006; Goetz et al., 2009).

Several recent studies have discussed the potential and limitations of optical sensors for direct biomass or volume estimation, demonstrating the sensitivity of visible and shortwave infrared wavelengths to vegetation density and structure, which in turn are related to AGB (Baccini, Friedl, Woodcock, & Warbington, 2004; Foody et al., 2003; Gemmell, 1995; Lu, Batistella, & Moran, 2005; Lu, Mausel, Brondizio, & Moran, 2004; Steininger, 2000).

Medium resolution data (e.g. Landsat, ASTER, SPOT) are usually employed for biomass analysis at local scales (Foody et al., 2001; Hall, Skakun, Arsenault, & Case, 2006; Labrecque, Fournier, Luther, & Piercey, 2006; Lu, 2005; Muukkonen & Heiskanen, 2005; Phua & Saito, 2003; Powell et al., 2010; Zheng, Chen, Tian, Ju, & Xia, 2007; Zheng et al., 2004) and moderate to coarse resolution data (e.g. MODIS) at regional scales (Baccini, Laporte, Goetz, Sun, & Dong, 2008; Baccini et al., 2004; Blackard et al., 2008; Saatchi, Houghton, Dos Santos Alvalá, Soares, & Yu, 2007; Zhang & Kondragunta, 2006). Some studies have integrated medium and coarse resolution data (Muukkonen & Heiskanen, 2007; Tomppo, Nilsson, Rosengren, Aalto, & Kennedy, 2002). While medium resolution data provide spatial detail compatible with the size of vegetation units and biomass field observations, compiling a temporally and radiometrically consistent cloud-free datasets over large areas is not always possible. This issue is addressed by the large swath and frequent repeat cycle of moderate resolution sensors, but the limited spatial detail misses the small-scale biomass variability, and it is often difficult to relate field data with satellite observations because of mismatch in measurement scale or resolution (Baccini, Friedl, Woodcock, & Zhu, 2007). The use of Landsat data for large-area analysis has become more practical with the opening and free distribution of its archive (Woodcock et al., 2008), which has global coverage and includes nearly 40 years of data (Goward et al., 2006; Williams, Goward, & Arvidson, 2006).

Recently, the use of LiDAR (Light Detection And Ranging) remote sensing in biomass estimation has increased (Koch, 2010). Because waveform LiDAR metrics are sensitive to vertical canopy structure and the interconnection of the latter with biomass (Drake et al., 2002; Lefsky et al., 2002), several studies have found a strong linear correlation between LiDAR data and AGB (Drake et al., 2002; Drake et al., 2003; Lefsky, Harding, Cohen, Parker, & Shugart, 1999; Nelson, Krabill, & Tonelli, 1988). So far, most of the large-scale investigations have focused on sub-boreal forests (Boudreau et al., 2008; Nelson et al., 2009; Ranson et al., 2007) and few studies have applied airborne (Asner, 2009; Drake et al., 2003) or spaceborne (Baccini et al., 2008; Lefsky et al., 2005) LiDAR data in tropical areas. Airborne sensors provide highly accurate biomass estimates (Ni-Meister et al., 2010), but the associated large data volume and high costs usually limit their application to local scales. As an alternative, the Geoscience Laser Altimeter System (GLAS) sensor on board the Ice, Cloud and Elevation satellite (ICESAT) satellite has proven to be valuable for biomass and canopy height estimation over large areas (Lefsky, 2010; Lefsky et al., 2005). Since the sampling strategy of the GLAS sensor does not allow spatially continuous representation of biomass or canopy variability, GLAS data are often integrated with imaging optical systems (Baccini et al., 2008; Goetz, Sun, Baccini, & Beck, 2010; Nelson et al., 2009).

In this study, we explore capabilities and limitations of Landsat data and ancillary information (land cover) for producing high-resolution, accurate and spatially explicit AGB estimates at national scale in the tropics, to be used in support of national planning and REDD-related activities. We also use independent LiDAR measurements acquired by the GLAS sensor to assess model results.

2. Data and methods

Uganda is selected as a case-study because of the availability of an extensive biomass field dataset collected by the National Biomass Study program (Drichi, 2003), necessary for model training and validation. Because of the availability of field data for a wide range of vegetation types (e.g. woodland, shrubland, savanna), we did not limit our analysis to forests, but included all woody formations. This is especially important in the dry and semi-dry tropics where non-forest vegetation types store substantial amounts of biomass because their low biomass density is counterbalanced by coverage over large areas. In this paper, the terms biomass and AGB refer to live woody aboveground biomass.

2.1. Study area

Uganda is located along the equator in East Africa mainly between 900 and 1500 m above sea level and presents a tropical climate with two rainy seasons in the South and one in the North. It has a total area of 241,551 km2, of which subsistence cropland is the most widespread land cover type (35%), followed by grassland (21%) and woodland (16%). Water bodies cover 15% of total area (Drichi, 2003). Vegetation is mainly represented by shrubland in the north (yearly precipitation 900 mm) and grassland, woodland and forest in the south and west (yearly precipitation 1500 mm). The country biomass stock is estimated at 468 Tg of air-dry AGB, mainly located in tropical forests (29%) and woodlands (27%) (Drichi, 2003). Areas classified as cropland and grassland also store substantial amounts of biomass (24% and 10%, respectively) because of the abundance of scattered trees (Drichi, 2003; FAO, 2003).

The high economic and population growth experienced in the last 25 years has had a dramatic impact on forestry resources, which have been reduced by agricultural expansion and growing demand for charcoal, fuelwood and timber (Drichi, 2003; FAO, 2003). Woody vegetation (plantation, forest, woodland and shrubland) decreased from 26% to 20% of total area between 1990 and 2005, causing a reduction of 26% of Uganda's total biomass stock (FAO, 2006).

Uganda is expected to experience continued rapid population and economic growth in the coming years, with associated smallholder agricultural expansion into forested areas and increasing demand for forest products that will likely accelerate deforestation (FAO, 2003; Kanabahita, 2001). With 66% of the total AGB located outside of protected areas (Drichi, 2003) and insufficient incentives to pursue sustainable forest management (Kanabahita, 2001), Uganda is therefore highly relevant to REDD policies and biomass monitoring activities.

2.2. Satellite data

We compiled a mosaic of 17 Landsat L1T ETM+ images acquired in the period 1999–2003 during the dry season (December to March) when cloud coverage is at a minimum and spectral separability between trees and grass/shrub is at a maximum (Table 1, Fig. 1, Fig. 2). Landsat TM data for the period of interest were not available because of missing receiving stations for central Africa (Goward et al., 2006). Due to frequent cloud cover, it was not possible to compile a single phenologically consistent dataset for the whole country. Instead, the mosaic includes 2 phenologically consistent sets of images, hereafter referred to as image "blocks", used to develop two separate biomass models. The four images located in the South-Western part of the country form "block 2" while all the other images form "block 1" (Table 1).

We also used measurements from the GLAS instrument, a waveform sampling LiDAR sensor (Schutz et al., 2005) sensitive to vegetation structure (Lefsky et al., 1999, 2005; Sun, Ranson, Kimes, Blair, & Kovacs, 2008). The LiDAR metric Height of Median Energy (HOME) is highly correlated to AGB (Baccini et al., 2008; Drake et al., 2003) and can be used as an independent comparison with the model