Remote Sensing of Environment 115 (2011) 1837–1849

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Modeling the height of young forests regenerating from recent disturbances in Mississippi using Landsat and ICESat data

Ainong Li [a,b], Chengquan Huang [a,*], Guoqing Sun [a], Hua Shi [c], Chris Toney [d], Zhiliang Zhu [e], Matthew G. Rollins [f], Samuel N. Goward [a], Jeffrey G. Masek [g]

[a] Department of Geography, University of Maryland, College park, MD 20742, USA
[b] Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu, Sichuan 610041, China
[c] ASRC Research and Technology Solutions (ARTS), Contractor to the USGS/EROS, Sioux Falls, SD 57198, USA
[d] USDA Forest Service, Rocky Mountain Research Station, Missoula, MT 59808, USA
[e] U.S. Geological Survey, 12201 Sunrise Valley Drive, Reston, VA 20771, USA
[f] Center for Earth Resources Observation and Science, US Geological Survey, Sioux Falls, SD 57198, USA
[g] Biospheric Science Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

ARTICLE INFO

Article history:
Received 30 December 2009
Received in revised form 2 March 2011
Accepted 2 March 2011
Available online 9 April 2011

Keywords:
Young forest
Disturbance
Height modeling
VCT
LTSS
GLAS

ABSTRACT

Many forestry and earth science applications require spatially detailed forest height data sets. Among the various remote sensing technologies, lidar offers the most potential for obtaining reliable height measurement. However, existing and planned spaceborne lidar systems do not have the capability to produce spatially contiguous, fine resolution forest height maps over large areas. This paper describes a Landsat–lidar fusion approach for modeling the height of young forests by integrating historical Landsat observations with lidar data acquired by the Geoscience Laser Altimeter System (GLAS) instrument onboard the Ice, Cloud, and land Elevation (ICESat) satellite. In this approach, "young" forests refer to forests reestablished following recent disturbances mapped using Landsat time-series stacks (LTSS) and a vegetation change tracker (VCT) algorithm. The GLAS lidar data is used to retrieve forest height at sample locations represented by the footprints of the lidar data. These samples are used to establish relationships between lidar-based forest height measurements and LTSS–VCT disturbance products. The height of "young" forest is then mapped based on the derived relationships and the LTSS–VCT disturbance products. This approach was developed and tested over the state of Mississippi. Of the various models evaluated, a regression tree model predicting forest height from age since disturbance and three cumulative indices produced by the LTSS–VCT method yielded the lowest cross validation error. The R^2 and root mean square difference (RMSD) between predicted and GLAS-based height measurements were 0.91 and 1.97 m, respectively. Predictions of this model had much higher errors than indicated by cross validation analysis when evaluated using field plot data collected through the Forest Inventory and Analysis Program of USDA Forest Service. Much of these errors were due to a lack of separation between stand clearing and non-stand clearing disturbances in current LTSS–VCT products and difficulty in deriving reliable forest height measurements using GLAS samples when terrain relief was present within their footprints. In addition, a systematic underestimation of about 5 m by the developed model was also observed, half of which could be explained by forest growth that occurred between field measurement year and model target year. The remaining difference suggests that tree height measurements derived using waveform lidar data could be significantly underestimated, especially for young pine forests. Options for improving the height modeling approach developed in this study were discussed.

1. Introduction

Obtaining reliable and up-to-date information on forest structure has been identified as one of the key needs for advancing studies on many pressing environmental issues (NRC, 2007). In particular, forest height is needed for biomass and carbon stock assessment (Brown & Schroeder, 1999; Houghton, 2005), fuel estimation and fire behavior modeling (Rollins, 2009; Scott & Burgan, 2005), habitat assessment (Hinsley et al., 2002; Zarnetske et al., 2007), and many other environmental and earth system studies. Although field survey or airborne methods can be used to create forest height maps, these methods are labor and resource intensive. Satellite remote sensing provides a cost effective alternative for large area applications.

Among the various remote sensing technologies, lidar offers the most potential for direct measurement of forest height. As a laser altimeter, lidar is used to determine the distance from the instrument

* Corresponding author.
E-mail address: cqhuang@umd.edu (C. Huang).

0034-4257/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2011.03.001

to a ground target by measuring the time elapsed between a laser pulse emission and a return signal from that target (Bachman, 1979). Tree height is calculated as the difference between the distance values from tree canopy and the ground to the instrument. Good agreements between height estimates derived using lidar data and ground measurements have been reported in many studies (Dubayah & Drake, 2000; Harding et al., 2001; Lefsky et al., 2002; Næsset & Økland, 2002; Nilsson, 1996). As of the writing of this paper, however, the Geoscience Laser Altimater System (GLAS) instrument onboard the Ice, Cloud, and land Elevation (ICESat) satellite is the only spaceborne lidar instrument capable of making measurements with a sub-ha footprint. Constrained by the nature of the instrument, these measurements are only available at sample locations distributed along the satellite track. While spatially contiguous data sets needed for creating spatially contiguous products can be acquired using airborne lidar, due to high cost of flights, acquiring airborne lidar data over large regions is often cost prohibitive. So far, statewide or near statewide lidar data sets have been acquired for only a few states in the United States (e.g., North Carolina and Maryland).

Unlike lidar data, spatially contiguous radar and optical remote sensing data sets are more readily available over very large areas. Radar signal has been found useful for deriving information on forest structure at low biomass levels (e.g. Balzter, 2001; Garestier & Dubois-Fernandez, 2008; Sun et al., 2003; Walker et al., 2007). Its sensitivity to biomass change, however, saturates when biomass reaches certain levels, and the location of the saturation point is wavelength dependent (Balzter et al., 2007; Coops, 2002; Mougin et al., 1999). Operating using visible/infrared wavelengths, optical remote sensing has not been considered very useful for modeling forest structure in general. Except for certain local studies, reported relationships between optical imagery and tree height were typically weak (Donoghue & Watt, 2006; Franklin et al., 2003), although better relationships were also reported when multi-angle measurements were used (Chopping et al., 2008).

The spatial and temporal coverage of available optical imagery is substantially better than that of available lidar and radar data. Images acquired by optical instruments are typically spatially contiguous. Some optical remote sensing systems, often with a series of similar instruments, have been in operation for several decades. In particular, a series of six Landsat instruments have established an imagery record dating back to 1972, which can be used to evaluate forest disturbance over the last several decades (e.g. Cohen et al., 1998; Healey et al., 2005; Kennedy et al., 2007). Recently, a vegetation change tracker (VCT) algorithm was developed for reconstructing forest disturbance history and age structure using Landsat time series stacks (LTSS) (Huang et al., 2009a,b, 2010). Similar algorithms for mapping forest change using time series Landsat observations have also been developed in other studies (Kennedy et al., 2007, 2010). Because age is often used as a good predictor of forest growth and yield (Pretzsch, 2001; von Gadow & Hui, 1999), the age since disturbance calculated using the LTSS–VCT approach may be used to improve forest height modeling, especially for young forests reestablished following recent disturbances. In addition, the LTSS provides a spectral record of the growing history of those young forests. Such a record may reflect the collective impact of vegetation species composition, site conditions and local climate on forest growth, and therefore may provide additional improvements to the height–age relationships.

The main purpose of this study is to develop an approach for modeling the height of "young" forests by integrating historical Landsat observations with GLAS lidar data, and to assess its performance over the entire state of Mississippi. Here, "young" forest is not necessarily defined following an ecologically-based age grouping method. It refers to forests reestablished following stand clearing disturbance events that were mapped using the LTSS–VCT approach. As will be discussed later, the earliest Landsat images used in this study were acquired in 1984, and the modeling target year was 2004. Therefore, the "young" forests in the context of this study had age since disturbance values of 20 years or less. With an average disturbance rate of about 2% per year (Li et al., 2009b), such "young" forests accounted for about 40% of all forests in Mississippi. The modeling approach developed here consists of three major steps: 1) development of forest disturbance products using the LTSS–VCT approach, 2) derivation of forest height using ICESat GLAS data, and 3) forest height modeling and prediction. The last two steps will be the focus of the methodology description in this paper, because details on the first step have been provided in previous publications (Huang et al., 2009a, 2010). Only a brief overview of the LTSS–VCT method and the disturbance products derived using this method over Mississippi is provided for completeness.

2. Data and methods

2.1. Study area

Mississippi was selected as the study area because spatially contiguous disturbance products had already been developed through previous efforts (Huang et al., 2009a, 2010; Li et al., 2009b). It is located in the deep south of the United States (Fig. 1), extending from 88.12°W to 91.68°W and 30.22°N to 35°N, with a land area of 125,443 km^2. It has a humid subtropical climate with long summers and short, mild winters. Composed mostly of low hills, the majority of the state belongs to the East Gulf Coast Plain ecosystem. The northwest of the state is made up of a section of the Mississippi Alluvial Plain, which is narrow in the south and widens north of Vicksburg.

Approximately 65% of Mississippi is covered by forests (Morgan et al., 2007). Major forest types include pine forest, hardwood forest and oak–pine forest. Major disturbances to the forest here include flooding, hurricane damage, wild land fire, timber harvest and reforestation. Lumber is a prevalent industry in Mississippi. Forest change occurs frequently due to forest harvesting, and rapid regrowth or reforestation is a common practice to ensure a sustainable supply of sawtimber and pulpwood for producing forest-dependent products. Because of their rapid juvenile growth and early commercial maturity, loblolly and slash pines are the major species planted by large land owners (Pinder & Rea, 1999). Such pine plantations are typically harvested with rotation cycles of 20 to 30 years (Walker, 1994). To avoid labor costs incurred by multiple visits, clearcut is often the preferred harvesting method (Faulkner et al., 1993). Before a clearcut harvest, however, a forest stand may be thinned multiple times to boost the yield of merchantable volume.

2.2. Approach overview

As discussed earlier, our approach for modeling forest height using Landsat time series observations and GLAS lidar data consists of three major steps. In the first step, forest disturbances that occurred during the observing period of each LTSS are mapped using the LTSS–VCT method. For the detected disturbances, age since disturbance and several cumulative indices designed to track the spectral progression of forest growth are calculated. In the second step, available GLAS laser shots are filtered to identify samples suitable for reliable forest height retrieval, and forest height values are derived for those samples using the GLAS waveform data. Finally, the selected GLAS samples are used to train and evaluate forest height models. The best model is then used to make predictions for all "young" forests across Mississippi. A flowchart of these steps is provided in Fig. 2.

2.3. LTSS–VCT disturbance products

2.3.1. Product description

Mississippi is covered by 11 Landsat path/row tiles as defined using the World Reference System (WRS) (Fig. 1). For each path/row

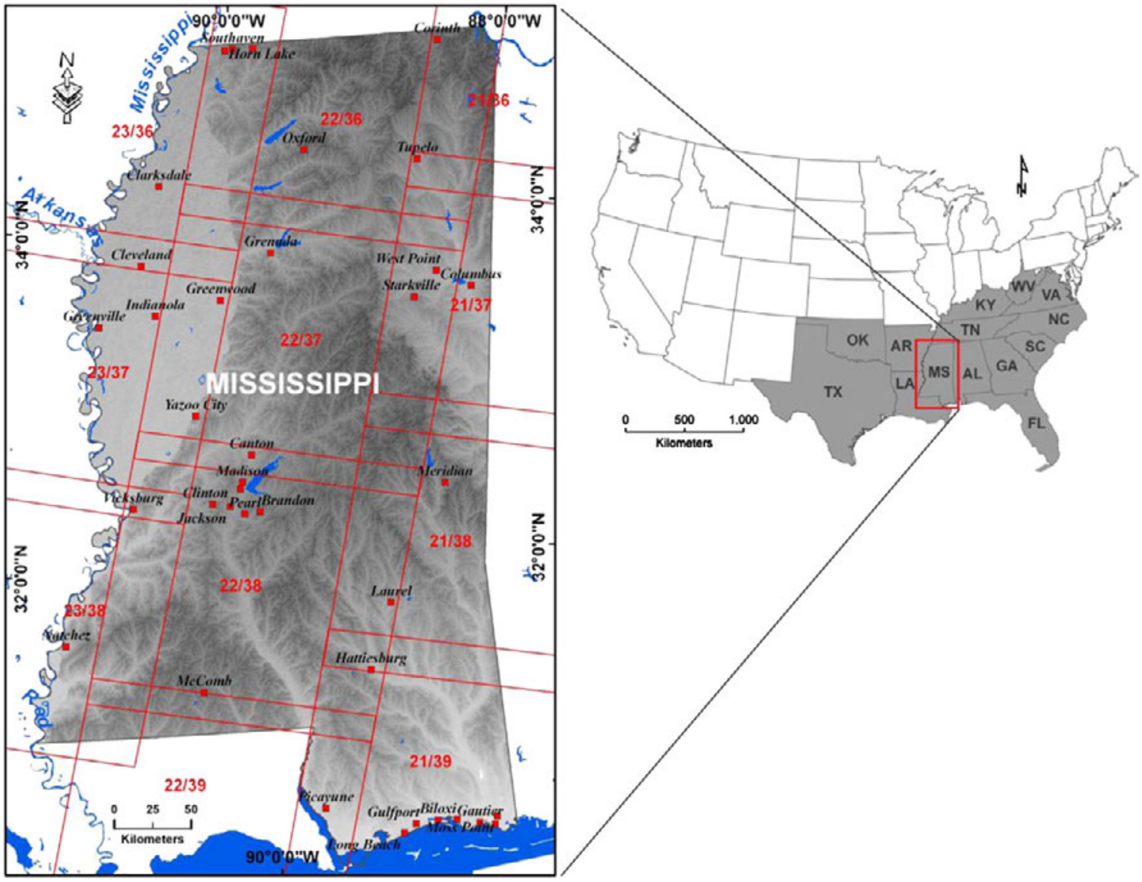

Fig. 1. Location of the study area, Mississippi (left), in the United States (right). The study area (left) is shown with topography as the background (elevation range: 0–323 m). The quadrangles and the numbers inside each of them (e.g. 22/38) show the boundary and the path/row numbers of the WRS tiles needed to cover the study area.

tile, we assembled a Landsat time series stack (LTSS) consisting of approximately one Landsat TM or ETM+ image every two years from 1984 to around 2006 (Table 1). In this study, the TM and ETM+ images were used interchangeably, because they had near identical spectral and spatial characteristics. The selected images were acquired during the leaf-on growing season and had minimal or no cloud contamination. A total of 155 images were used to develop the LTSS for the 11 path/row tiles needed to cover Mississippi (Table 1). These images were corrected to achieve subpixel geolocation accuracy and high levels of radiometric consistency. A detailed description of the procedures for developing LTSS has been provided by Huang et al. (2009a).

Each LTSS was analyzed using a vegetation change tracker (VCT) algorithm to produce disturbance products. VCT is a highly automated forest change mapping algorithm designed for analyzing the LTSS. It consists of two major steps: individual image analysis and time series

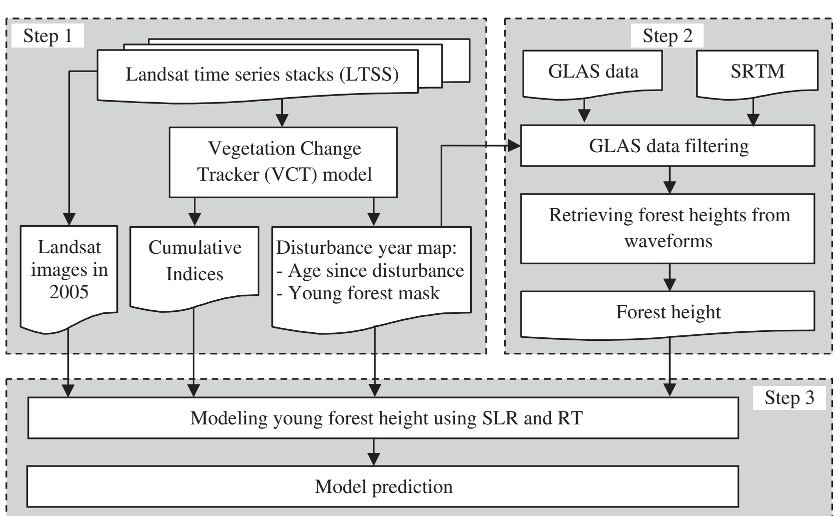

Fig. 2. A flowchart of the overall approach for modeling young forest height using Landsat time series observations and ICESat GLAS data.

Table 1
Acquisition dates (yyyy/mm/dd) of Landsat images used in this study.

	WRS path/row										
	21/36	21/37	21/38	21/39	22/36	22/37	22/38	22/39	23/36	23/37	23/38
1	1984/06/18	1984/09/06	1984/09/06	1984/09/06	1984/10/31	1984/10/31	1984/09/13	1984/09/13	1984/07/18	1984/07/18	1986/08/25
2	1986/09/28	1987/06/27	1986/06/24	1986/06/24	1986/07/17	1986/10/21	1986/03/27	1985/08/31	1986/07/24	1986/07/24	1987/07/27
3	1988/06/13	1988/06/13	1987/06/27	1987/06/27	1988/09/08	1988/07/31	1987/08/21	1987/10/08	1988/05/26	1988/07/29	1989/10/20
4	1990/06/19	1990/09/07	1989/10/22	1989/10/22	1990/07/28	1991/07/31	1990/06/10	1989/05/06	1990/08/20	1990/08/04	1991/08/23
5	1991/09/26	1991/09/26	1990/06/19	1991/09/26	1991/07/31	1993/09/22	1991/07/13	1991/07/31	1991/08/23	1992/05/05	1992/05/05
6	1993/10/01	1993/10/01	1991/09/26	1993/10/01	1993/06/18	1995/08/27	1993/06/02	1993/06/02	1993/05/08	1993/09/29	1993/09/29
7	1995/10/07	1995/06/17	1993/10/01	1995/10/07	1995/08/27	1997/10/03	1995/05/23	1995/08/27	1995/07/17	1995/08/02	1995/08/02
8	1997/08/25	1997/08/25	1995/10/07	1997/08/25	1997/10/03	1999/08/22	1997/10/03	1996/04/07	1997/07/06	1997/08/23	1997/09/08
9	2000/08/17	1999/09/16	1997/08/25	1999/08/15	1999/08/06	2000/07/15	1999/09/23	1998/09/04	1999/08/29	1999/10/24	1999/10/24
10	2001/10/15	2001/09/29	1999/08/15	2001/10/15	2000/08/16	2001/05/15	2000/07/15	2000/07/15	2001/10/29	2001/10/21	2001/05/14
11	2002/08/07	2003/06/23	2000/07/08	2002/10/18	2001/04/29	2003/05/29	2002/10/17	2001/11/07	2003/06/21	2002/06/18	2002/08/21
12	2004/05/08	2005/10/18	2001/10/15	2004/10/15	2002/07/05	2004/09/20	2004/09/20	2002/09/15	2005/05/25	2004/09/27	2004/09/27
13	2005/05/27	2006/06/15	2002/08/07	2005/05/11	2004/08/19	2005/09/07	2006/09/26	2004/09/20	2007/08/03		2006/05/12
14	2006/06/15	2007/05/17	2003/10/29	2006/08/22	2006/08/22	2006/05/21		2005/09/07			
15			2004/10/15					2006/06/06			
16			2005/05/27								
17			2006/07/17								
18			2007/05/17								

analysis. During the first step, each image is analyzed separately to identify some forest samples, which are used to calculate an integrated forest z-score (IFZ) index as a measure of forest likelihood (Huang et al., 2008). During this step, water, cloud and shadow are also masked, and several other spectral indices are calculated. Once this step is completed for all images of a LTSS, forest changes are detected and tracked through a time series analysis step. Detailed descriptions of the VCT algorithm have been provided in previous publications (Huang et al., 2009b, 2010).

The VCT algorithm produces a suite of products, including disturbance maps and measures for characterizing the detected disturbances and for tracking post-disturbance processes (Huang et al., 2009b, 2010). A disturbance map indicates where and when disturbances occurred (Fig. 3). For each detected disturbance, a

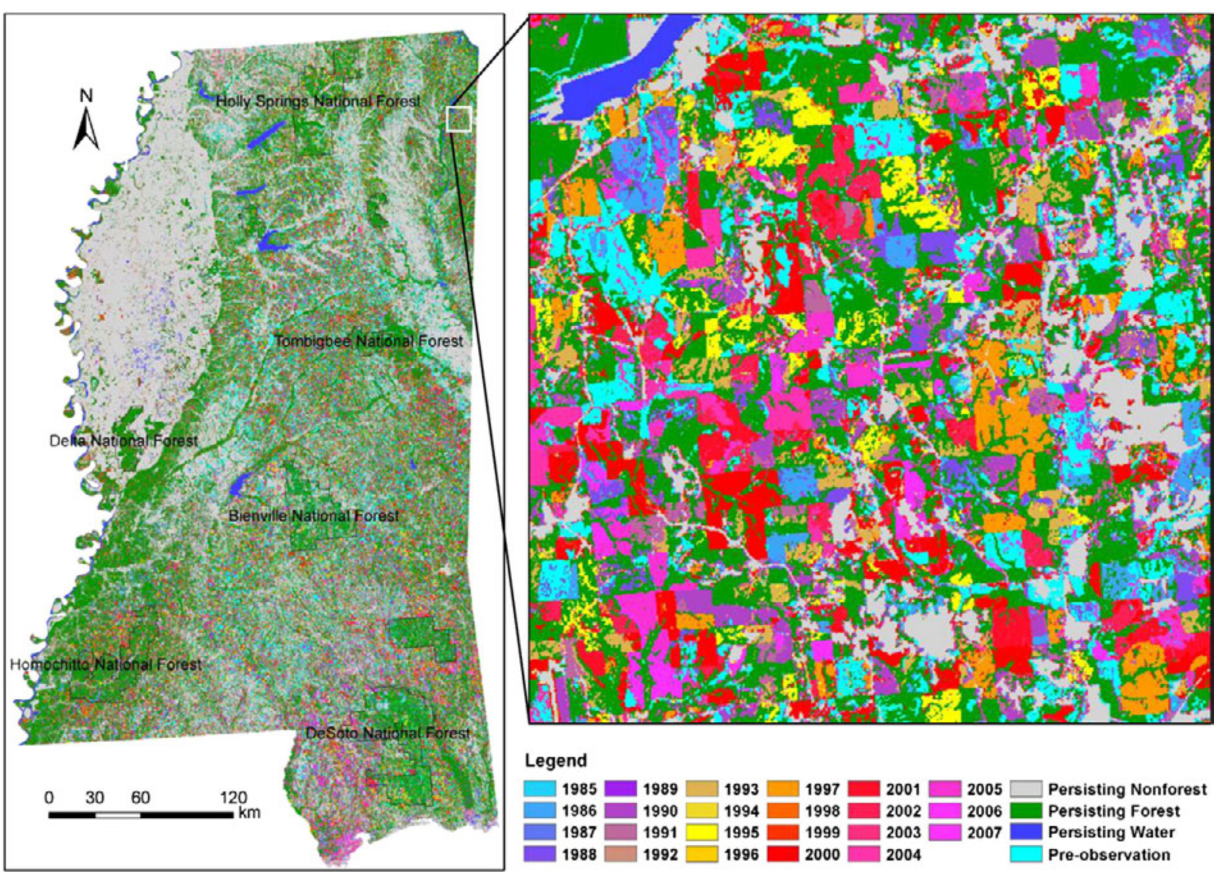

Fig. 3. Overview (left) and full resolution (right) disturbance year map of Mississippi developed using the LTSS–VCT approach. The full resolution map on the right side covered a ground area of 14.5 km by 15 km. In the legend each year number indicates a disturbance year. The "Pre-observation" category refers to disturbances that occurred during or before the first observation of the LTSS and regrowth was observed in later years.

disturbance year value, i.e., the year when that disturbance occurred, is recorded. Because most disturbances in Mississippi were timber harvests that were often followed by regeneration of young forests, the disturbed pixels mapped by the VCT were used to identify "young" forest pixels whose height will be modeled in this study. For each "young" forest pixel, an age since disturbance (AGE_SD) is calculated as the difference between model target year and the disturbance year. For forest stands regenerated immediately following stand clearing disturbances, this measure should be close to their actual age. Pixels that were not mapped as disturbed were labeled with one of the following three classes: *persisting water*, *persisting non-forest*, or *persisting forest*. Here we used "persisting" to indicate that a pixel had the same land cover type throughout the entire observing period of the LTSS.

For each detected disturbance, VCT uses three indices to track the post-disturbance recovery process, including the IFZ, the normalized difference vegetation index (NDVI), and a normalized burn ratio index (NBRI). IFZ is an inverse measure of the likelihood of a pixel being a forest pixel, and is calculated using the mean ($\bar{b}_i$) and stand deviation (SD_i) of forest sample as follows:

$$IFZ = \sqrt{\frac{1}{NB}\sum_{i=1}^{NB}\left(\frac{b_i - \bar{b}_i}{SD_i}\right)^2} \tag{1}$$

where NB is the number of spectral bands and b_i the spectral value of a pixel in band i. VCT uses Landsat bands 3, 4, and 7 to calculate the IFZ (Huang et al., 2010). NDVI and NBRI are spectral indices calculated using TM bands 3 (R_3), 4 (R_4), and 7 (R_7) as follows:

$$NDVI = \frac{R_4 - R_3}{R_4 + R_3} \tag{2}$$

$$NBRI = \frac{R_4 - R_7}{R_4 + R_7}. \tag{3}$$

NDVI is an indicator of vegetation greenness. While NBRI was designed primarily for measuring burn severity (Chen et al., 2008; Escuin et al., 2008), its correlations with forest structure variables were found significant in some studies (e.g. Pascual et al., 2010). Therefore, it may be useful for tracking forest growth.

While the growth rate of a forest stand is controlled by many factors, including species composition, local environmental conditions, and microclimate (Amaro et al., 2003), we hypothesized that the aggregate impact of these factors on forest growth could be tracked using the temporal profiles of these indices. Other spectral indices, if found more effective in tracking forest growth, could be used in the place of these indices in the modeling approach developed here. For each index, a cumulative value calculated using the following equation is used to characterize the temporal profile (Fig. 4):

$$y = \sum_{i=m}^{n} x_i \tag{4}$$

where y is the cumulative value for one of the three indices, m the disturbance year as detected by the VCT, and n the ending year for calculating the cumulative index. For the GLAS samples used as training data for model development, the ending year was set to match the acquisition year of each GLAS sample. For statewide prediction, the ending year was set to 2004 for most LTSS. This target year for model prediction was chosen because one of the LTSS ended in 2004 (Table 1). No images acquired after this year would be available for calculating the cumulative indices for that LTSS if a later year were chosen as the target year for model prediction. For LTSS that did not have a 2004 image, the target year for model prediction was set to 2003.

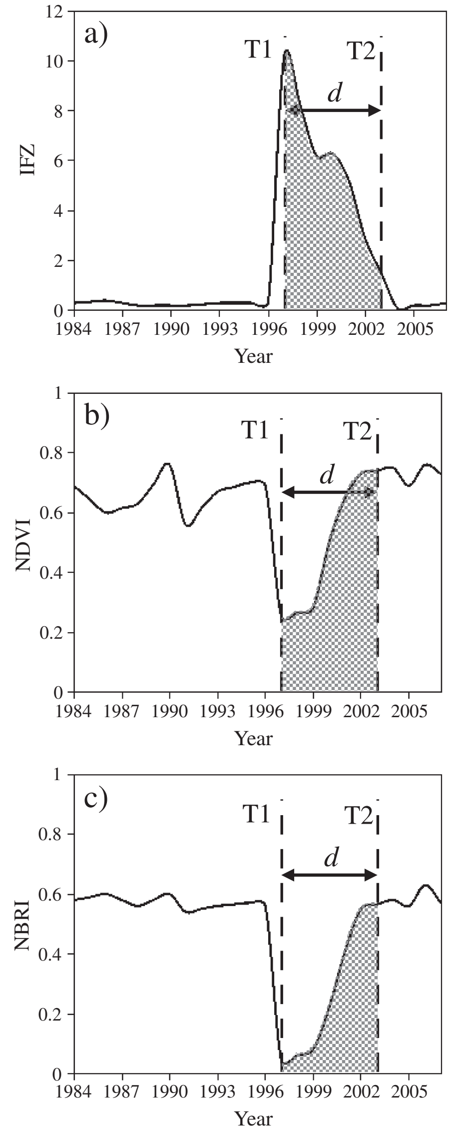

Fig. 4. Calculation of cumulative indices using *IFI* (a), *NDVI* (b), and *NBRI* (c). *T1* and *T2* are the same as the *m* and *n* in Eq. (4). Each cumulative index is represented by the shaded area in the corresponding figure, and age since disturbance is represented by *d*.

Because we did not have an image for every year in the LTSS, for each of the three indices (x) its value in a year (i) during which no image was acquired was calculated through linear interpolation using the immediately previous (p) and next (n) LTSS acquisitions as follows:

$$x_i = x_p + (i - p) \times \frac{x_n - x_p}{n - p}. \tag{5}$$

2.3.2. Assessment of the disturbance products

VCT has been tested in many places of the U.S., including Mississippi (Li et al., 2009b), Alabama (Li et al., 2009a), and the locations where LTSS have been assembled through the North American Forest Dynamics (NAFD) project (Goward et al., 2008; Huang et al., 2009a). Efforts to assess the VCT disturbance products have so far focused on the disturbance year products as shown in Fig. 3, because reference data sets needed to calibrate or validate other VCT products are extremely difficult to find. Through the NAFD project, the VCT disturbance year maps were evaluated using a design-based accuracy assessment method over 6 sites selected across the U.S., including one site in Mississippi (WRS path 21/row 37). On average, the disturbance year products had overall accuracies of about 80%, with the disturbance classes having user's accuracies around 80% and producer's accuracies around 60% (Huang et al., 2009b; Thomas et al., 2010). Furthermore, a comparison with field data collected through the USDA Forest Inventory and Analysis (FIA) program revealed that over 75% of the age variance of young forests in the Mississippi site could be explained by the AGE_SD calculated using the VCT disturbance year map (Thomas et al., 2010).

2.4. Deriving forest height from GLAS data

2.4.1. GLAS data

The GLAS sensor on board NASA's ICESat satellite was designed to collect high precision altimetry measurements using lidar. The lidar operates using the 1064-nm and 532-nm wavelengths. The laser pulses are sent at a frequency of 40 Hz, and the returned energy is recorded in waveform (Fig. 5). The lidar footprint has a diameter of about 65 m, but its size and ellipticity have varied significantly through the course of the mission as a function of the power output from the laser (Schutz et al., 2005). The GLAS samples have an along-track interval of 170 m. The cross-track interval varies as a function of latitude (Abshire et al., 2005), and is about 25 km in Mississippi. The GLAS mission operates with a 91-day repeat orbit and a 33 day sub-cycle. GLAS carries three laser altimeters named as L1, L2, and L3. L1 was turned off shortly after the spring campaign in 2003. L2 and L3 operated from September 24, 2003 to June 21, 2004 and October 3, 2004 to October 19, 2008, respectively. The waveform data recorded by the GLAS has a vertical resolution of 15 cm, resulting in a height range of 81.6 m for L1 and L2A, and 150 m for L2B, L2C and L3 (Harding & Carabajal, 2005). The GLAS laser altimeter has a range precision of 3 cm and a pointing determination accuracy of better than 2 arc-second (Sirota et al., 2005).

The GLAS had a total of 18 operational periods during its 5-year mission. The data acquired during the L2A and L3C periods were used in this study. The L2A and L3C data sets were acquired using the L2 laser from September 25 to November 19, 2003 and the L3C laser from May 20 to June 23 in 2005, respectively. In Mississippi, most forests should be in the leaf-on growing season during the dates when the two data sets were acquired, although some deciduous trees in the northern part of the state may lose some leaves by mid-November. Therefore, most GLAS samples in these two data sets were deemed suitable for retrieving forest height. The National Snow and Ice Data Center (NSIDC)[1] distributes 15 GLAS data products. Two of them were used in this study, including the waveform data (GLA01) and the global land surface altimetry data (GLA14). The latter provided various parameters including surface elevation, laser range offsets for signal beginning and end, acquisition time, location, amplitude, waveform centroid, and width of the fitted Gaussian peaks. The waveform produced by each laser shot was extracted from GLA01 data, while the exact location of that shot was determined by combining the two data sets (Fig. 5).

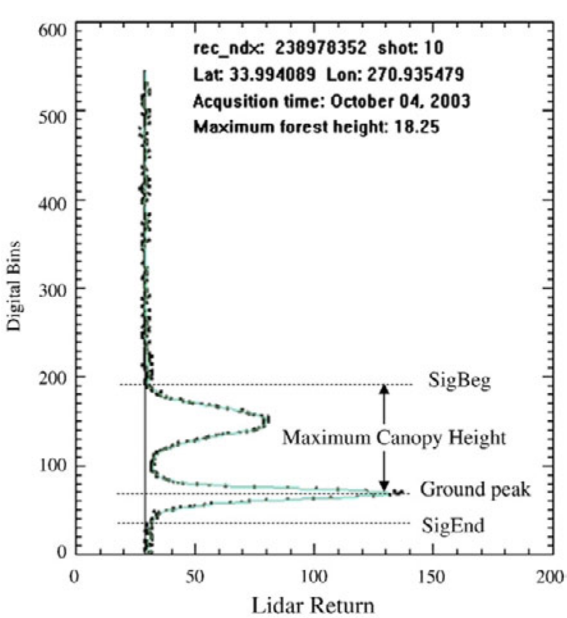

Fig. 5. A typical GLAS waveform over forest. The ground peak is the first significant peak found from the signal end (SigEnd) upwards. The Maximum Canopy Height was defined as the distance between the signal beginning (SigBeg) and the ground peak, both of which were determined by visually inspecting the waveform in this study.

2.4.2. Filtering of the GLAS data

Lidar technology provides a relatively straightforward way for measuring tree height. A lidar instrument determines its distance to a target according to the difference between the time a laser pulse is sent out and the time a return signal is received. For forest land tree height can be calculated as the difference between the distance from the instrument to the ground and to tree canopy (Fig. 5). However, terrain relief within the footprint of a laser shot can introduce substantial uncertainties in the height value derived this way. With each GLAS laser shot having a ground footprint of 65 m, even a moderate slope can produce a surface relief comparable with or more than the height of the overlying vegetation (Sun et al., 2008). While methods for decoupling the effect of terrain relief from tree canopy in waveform lidar data are being investigated (Lefsky et al., 2007), in this study we chose to avoid this problem by using laser shots over flat areas only.

To identify the GLAS laser shots over flat areas, we used the digital elevation model (DEM) data set produced through the Shuttle Radar Topography Mission (SRTM) to calculate terrain relief (Rabus et al., 2003). This data set is available at the 30 m spatial resolution for the U.S. Since the 65 m footprint of each GLAS laser shot most likely intersects with 9 SRTM pixels within a 3×3 window, for each GLAS laser shot the terrain relief within its footprint was calculated as the difference between the maximum and minimum values of the SRTM pixels within the 3×3 window centered at the centroid of that laser shot. This screening method is more rigorous in identifying flat areas than the use of slope as suggested in other studies (e.g. Nelson et al., 2009), because an area with a slope value of 0° calculated using a window of 3 by 3 pixels or larger can still have terrain relief. In this study, only laser shots with a calculated terrain relief of less than 1 m were kept for further analysis. Furthermore, the Landsat images may have residual geolocation errors of up to 30 m (Huang et al., 2009a), and the GLAS waveform data can have geolocation errors of up to 10 m (Brenner et al., 2003). To reduce the impact of potential uncertainties that may arise from residual geolocation errors in both the Landsat and the GLAS waveform data, we also excluded GLAS laser shots that were located within 2 pixels from the edge of disturbance patches.

2.4.3. Forest height retrieval

Given a GLAS lidar waveform as shown in Fig. 5, many height measures of the forest canopy can be derived (Brenner et al., 2003; Harding & Carabajal, 2005; Lefsky et al., 2005, 2007). In our study we used a maximum canopy height (MCH) to measure the height of dominant trees. It was calculated as the difference between the signal beginning (SigBeg) and the ground peak (Fig. 5). The signal beginning was identified by searching downwards from the top of the waveform until the point where the return signal was larger than 3 standard deviation of an estimated noise level, which was provided in the

[1] The GLAS data sets used in this study were downloaded from http://nsidc.org/data/icesat/order.html. More information on the GLAS mission and operational periods is also available at this website.

GLA01 product. Searching backwards from the signal ending, the first significant peak was the ground peak. To reduce noises in the original waveform signal, a Gaussian filter with a width similar to that of the transmitted laser pulse was applied. To ensure that the MCH was derived reliably, we visually inspected the waveform of each laser shot that passed the filtering described in Section 2.4.2. The following rules were followed in deriving MCH from the waveform data:

- For the waveforms with easily identifiable first and last peaks, the MCH was calculated as shown in Fig. 5;
- Waveforms that did not have obvious first and last peaks were discarded from further analysis, because they likely were contaminated by atmosphere effects or background noise;
- If a waveform had a single peak, the Landsat image acquired immediately before or in the same year as the GLAS data acquisition was visually inspected to determine whether the GLAS sample location had no forest cover or the forest at that location was so dense that no ground return signal was received. In the former case MCH was set to 0 m. In the latter case the sample was discarded because there was no way to derive MCH from the waveform;
- For each GLAS sample where a MCH value was retrieved, a height growth rate was calculated as the ratio between MCH and AGE_SD. Assuming a maximum growth rate of 1.5 m per year for trees in the south (Carmean et al., 1989), a calculated growth rate exceeding this maximum rate was considered illogical. Such illogical growth rates were likely the result of thinning and other non-stand clearing disturbance events, because for such events the AGE_SD values were often lower than actual forest age. GLAS samples having an illogical growth rate were excluded from the height modeling work described in Section 2.5.

After applying these filtering and screening rules, 373 and 197 GLAS samples were selected from the 24400 L2A samples acquired in 2003 and 11188 L3C samples acquired in 2005, respectively. These samples were scattered across the entire study area (Fig. 6), and had an AGE_SD distribution similar to that of the entire study area (Fig. 7).

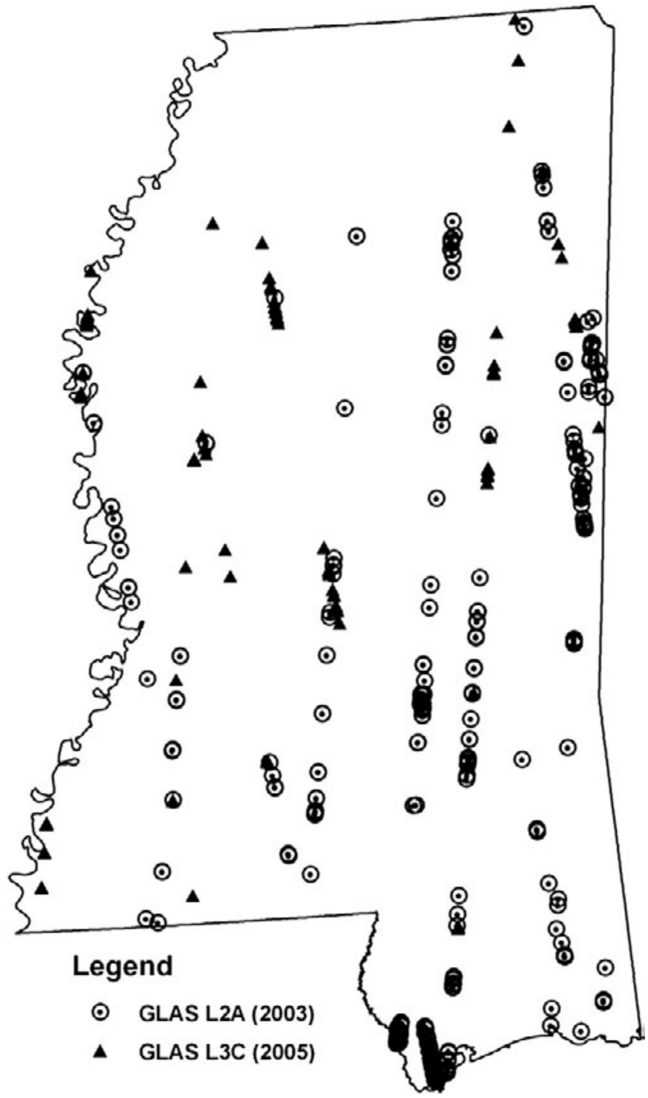

Fig. 6. Spatial distribution of the GLAS samples used in this study.

2.5. Height modeling

The GLAS samples selected using the filtering procedures described in Section 2.4 were used to develop models for predicting forest height. To determine which models provided the best predictive power, we evaluated eight models developed using two data mining techniques and different combinations of three groups of predictor variables.

2.5.1. Predictor variables

The following three groups of variables were considered for modeling forest height (Table 2):

- Landsat images that had acquisition years closest to those of the GLAS data (G1): All 6 spectral bands as well as the IFZ, NDVI, and NBRI were used as predictor variables. Because use of Landsat images and reference height measurements acquired at roughly the same time to model forest height was a common practice in previous studies (e.g. Donoghue et al., 2004; Franklin et al., 2003; Freitas et al., 2005; Jakubauskas & Price, 1997), the results derived using this group of variables were considered as a baseline for evaluating the more complex models developed in this study.
- Age_SD calculated the VCT (G2): As discussed earlier, for a forest stand established immediately following a previous disturbance, its age should be close to the AGE_SD measure (Thomas et al., 2010). In forestry, age together with site conditions has long been used to derive models for predicting forest growth and yield (Pretzsch, 2001; von Gadow & Hui, 1999).
- Cumulative indices calculated using the IFZ, NDVI, and NBRI according to Eq. (4) and Fig. 4 (G3): In addition to age, forest growth is also affected by species composition and many site conditions, including soil moisture, nutrient, and micro-climate variables. We hypothesized that since these cumulative indices track the spectral progression of forest growth, they likely reflect some of the collective effects of these conditions, and therefore may provide additional predictive power.

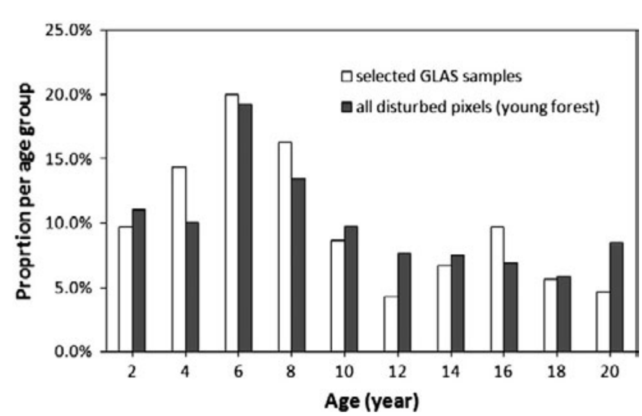

Fig. 7. Distribution of age since disturbance (AGE_SD) of young forests as represented by the selected GLAS samples and for the entire study area.

Table 2
Predictor variables used to model forest height.

Variables	Group	Definition
b1,b2,b3,b4,b5,b7 IFI,NDVI,NBRI	G1	The Top-Of-Atmosphere (TOA) reflectance of Landsat images (bands 1–5 and 7) acquired in the year closest to the acquisition of the GLAS data, and three indices calculated using Eqs. (1)–(3).
Age	G2	The age since disturbance, derived from VCT-LTSS disturbance year map.
Cum_IFI, Cum_NDVI, Cum_NBRI	G3	Cumulative indices calculated using Eqs. (1)–(5) (see Section 2.3.1 for details).

2.5.2. Modeling approaches

Two groups of data mining techniques were used to explore the relationships between forest height and the predictor variables, including stepwise linear regression (SLR) and regression tree (RT). The SLR assumes a linear relationship. Given the many available predictor variables as described above, SLR can be used to identify the ones that provide the most predictive power and remove those that do not add much to the overall predictive performance (Effroymson, 1960). For data sets characterized by nonlinear relationships, RT has a theoretical advantage over SLR, because RT can approximate non-linear relationships using a set of linear models (Huang & Townshend, 2003). This technique does not require prior knowledge on the mathematical form of a nonlinear relationship (Breiman et al., 1984), and produces rules in a decision tree format, which are easier to understand than those produced using neural network or other nonlinear modeling methods (De'ath & Fabricius, 2000).

RT has been implemented in many computer software packages. The Cubist software was used in this study.[2] It has been used to model land cover and biophysical variables using remote sensing data in many studies (e.g. Blackard et al., 2008; Walker et al., 2007). In addition to the basic RT algorithm concept, Cubist provides two options for improving model performance, including *instance* and *committee*. A model of N committee consists of N sets of RT models generated using an ensemble approach, while the *composite* option allows predictions be made using both regression tree models and a nearest neighbor approach called *instance* (Quinlan, 1996). Based on repeated trials, we chose the *composite* option and 5-*committee* to model forest height in this study.

To evaluate the predictive power of the predictor variables described in Section 2.5.1, we first applied the two data mining techniques to each of the three groups of predictor variables separately. We then pooled the three groups of variables together and used them as inputs to the two data mining techniques. The combinations of the two data mining techniques and the four groups of predictor variables used in model development are listed in Table 3.

2.5.3. Model assessment

The models developed in Section 2.5.2 were evaluated using cross-validation, a technique for deriving relatively independent accuracy estimates when only limited reference samples are available for model development (Quinlan, 1993). For an N-fold cross-validation, the reference samples are divided into N equal-sized subsets. By holding out each subset at a time for testing a model developed using the remaining (N-1) subsets, N models are developed and tested. The overall results derived from the N tests are used to represent the performance of the model developed using all reference samples.

[2] More details on Cubist can be found at http://www.rulequest.com/cubist-info.html.

Table 3
Combinations of predictor variables and data mining algorithms in modeling forest height.

SLR	RT
G1	G1
G2	G2
G3	G3
G1 + G2 + G3	G1 + G2 + G3

In this study, we used 10-fold cross validation to assess each of the models developed in Section 2.5.2. Model performance was measured by the agreement between model predictions and actual values derived using GLAS data. Specifically, for each model, we calculated an R^2 value following standard statistical textbooks and a root mean square of the difference (RMSD) using the following equation:

$$\mathrm{RMSD}(h_1, h_2) = \sqrt{\sum_{i=1}^{n}\left(h_{1,i}-h_{2,i}\right)^2 / n} \tag{6}$$

where h_1 refers to the model-predicted height values, h_2 the values derived from GLAS data, and n the number of samples.

2.5.4. Model prediction and validation

The model that gave the best cross validation results was used to make predictions for all "young" forest pixels as determined by the VCT for the entire study area. The predicted values were validated using a field data set collected through the US Forest Service Forest Inventory and Analysis (FIA) program (Smith, 2002) over Mississippi. This data set was collected between 2006 and 2007 over a systematically arranged network of plots distributed across the state with a density of roughly one plot for every 2428 ha of land base (Oswalt & Oswalt, 2008). Each plot consisted of four 7.3-m radius subplots, with one at the center of a triangle formed by the other three subplots (USDA Forest Service, 2005). The distance between the center subplot and each of the three surrounding subplots was 36.5-m. Within each subplot, all trees having a diameter at breast height (DBH) value of 12.5 cm or more were measured, and stand height was calculated as the average height of dominant and codominant trees weighted by basal area. For plots where no tree had a DBH >12.5 cm, stand height was the average height of sapling-sized trees (i.e., trees with DBH between 2.5 cm and 12.5 cm) which were measured in one 2.1-m radius microplot within each subplot. Heights of trees with DBH <2.5 cm were not measured.

For each plot, a stand age was also estimated by the field crew as the average age of live trees in the stand that are not overtopped (USDA Forest Service, 2005). Estimates of stand age were based on the time of tree establishment, not age at the point of diameter measurement. For planted stands, stand age was based on the year the stand was planted (i.e., the age of the planting stock was not added). An examination of the height–age relationship revealed that some plots had illogical combinations of height–age values, i.e., the growth rate calculated using stand height and age exceeded the maximum possible forest growth rate over Mississippi. Such illogical combinations may indicate errors in either the age or height measurements, or both. Stand age may be difficult to measure under certain circumstances, and hence can have large measurement errors (Pollard et al., 2006). For plots consisting of mostly small, young trees and a few big trees that were left for seeding purposes, the stand height may not reflect the height of the young trees, but that of the seeding trees. To minimize the impact of such measurement errors, plots with illogical height–age combinations were not used in this study. Following the same logic used to filter the GLAS samples (see Section 2.4.3), FIA plots with illogical height–age combinations were identified using a maximum possible growth rate of 1.5 m/year (Carmean et al., 1989). While some intensively managed pine forests may have a growth rate of >1.5 m/year (Siry, 2004), use of a different maximum possible growth rate value to identify illogical height–age combinations likely will not have much impact on the conclusions derived through this study, although it may result in a slightly different number of FIA plots that would be considered suitable for validating model predictions.

3. Results and analysis

3.1. Height measurements from the two GLAS data sets

The MCH values derived using the L2A and L3C data sets had similar relationships with age since disturbance (AGE_SD) when AGE_SD was 12 years or less (Fig. 8). For forests with AGE_SD greater than 12 years, the MCH values derived using the L3C data set appeared to be slightly lower than those derived using the L2A data set. A similar phenomenon was reported by Sun et al. (2008) when they compared the height values of some old forest stands derived using L2A and L3C GLAS data. Further studies are needed to investigate as to why the L2A GLAS data gave slightly higher height values than the L3C data for old forests. To obtain a larger sample size and better spatial distribution of the selected samples (Fig. 6), however, the samples from the two GLAS data sets were pooled together in developing the height models. The validity of our conclusions on the usefulness of the developed modeling approach likely will not be affected by the slightly different MCH–AGE_SD relationships given by the two GLAS data sets, although the performance of the developed models may be improved slightly if the height values derived using the two data sets were more consistent for all AGE_SD groups.

3.2. Performance of the height models

A total of 8 models were developed using the two data mining techniques and four groups of input variables (Table 3). Accuracies of these models as evaluated through cross validation are summarized in Table 4. It shows that the RT and SLR gave similar trends regarding the relative predictive power of the individual groups of predictor variables. The AGE_SD variable (G2) had the best relationships with MCH, explaining nearly 90% of its total variance. The cumulative indices (G3) calculated using Eq. (4) had the second best relationships, explaining nearly 70% of the total variance of the GLAS based forest height measurements. The Landsat images with acquisition year closest to the GLAS measurement year (G1) explained less than 20% of the total variance of MCH, which was expected, because in general spectral data is not very sensitive to height structure. Similarly poor relationships between forest height and Landsat images were

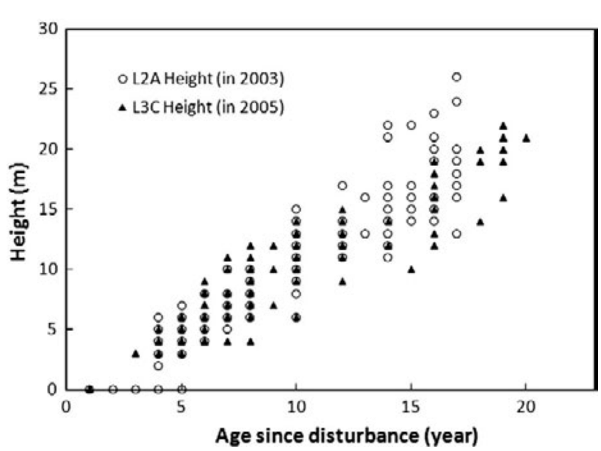

Fig. 8. Relationships between forest height values derived using the L2A and L3C data sets and age since disturbance.

Table 4
Cross-validation assessment of the forest height models developed in this study.

Variable groups	SLR models		RT models	
	R^2	RMSD (m)	R^2	RMSD (m)
G1	0.15	10.32	0.19	6.07
G2	0.88	2.42	0.89	2.13
G3	0.69	3.85	0.68	3.73
G1, G2, G3	0.91	2.08	0.91	1.97

reported in many previous studies (De La Cueva, 2008; Franklin et al., 2003; Hyyppa et al., 2000; Lefsky et al., 2001), although better correlations between modeled and actual height have also been reported in studies conducted within very small areas (Hall et al., 2006; Hudak et al., 2002; Jakubauskas & Price, 1997; Puhr & Donoghue, 2000). When pooled together, the three groups of variables resulted in models that performed better than the models developed using each individual group of variables.

The RT models yielded substantially lower RMSD than the SLR models when the G1 variables were used as predictor variables. However, the RMSD differences between the two groups of models were small when the AGE_SD (G2) and cumulative indices (G3) were included as predictor variables, suggesting that the relationships between MCH of and those predictor variables were linear or close to linear (e.g., see Fig. 8). As a result, although RT can use multiple linear equations to approximate nonlinear relationships, many RT models developed in this study had only two linear regression equations (Table 5). The best RT model was developed when all three variable groups were used together. When evaluated using cross validation, this model had an RMSD of 1.97 m and R^2 of 0.91 (Fig. 9), and most of the predicted values were within ± 5 m of GLAS based height measurements. A detailed examination of that model revealed that none of the G1 variables were used by the model. Therefore, we regenerated the RT model using the G2 and G3 variables. This model was used as the final model to make predictions for all "young" forest pixels identified by the VCT.

3.3. Comparison with FIA field measurements

A total of 586 FIA plots were located within the "young" forest areas as determined by the VCT. After excluding the plots that had illogical combinations of height and age values (see Section 2.5.4), 504 plots were left for validating the predictions of the final RT model. While the model predictions had good agreements with the GLAS-derived MCH values (Fig. 9), they had much larger disagreements with the FIA height measurements (Fig. 10(a)). In general, the FIA height measurements were higher than the model predictions, with the differences being more than 10 m for many plots. A detailed examination of the FIA data reveals that many of them (186 in total) had stand age values of 20 years or more, suggesting that the disturbances mapped by the VCT over those plot locations were not stand clearing disturbances and therefore did not reset the age of

Table 5
Example decision rules in forest height models produced by the regression tree algorithm.

Variable groups	Predictors	RT models
G2	Age	Rule 1: if age <=3, then height = −0.5 + 0.23 × age
		Rule 2: if age >3, then height = −1 + 1.14 × age
G2, G3		Rule 1: if age <=3, then height = −0.4 + 0.34 × age
	Age	−0.0012 × Cum_IFI − 0.0044 × Cum_NDVI
	Cum_IFI	+ 0.0045 × Cum_NBRI
	Cum_NDVI	Rule 2: if age >3, then height = −1.2 + 1.04 × age
	Cum_NBRI	−0.0035 × Cum_IFI − 0.013 × Cum_NDVI + 0.0133 × Cum_NBRI

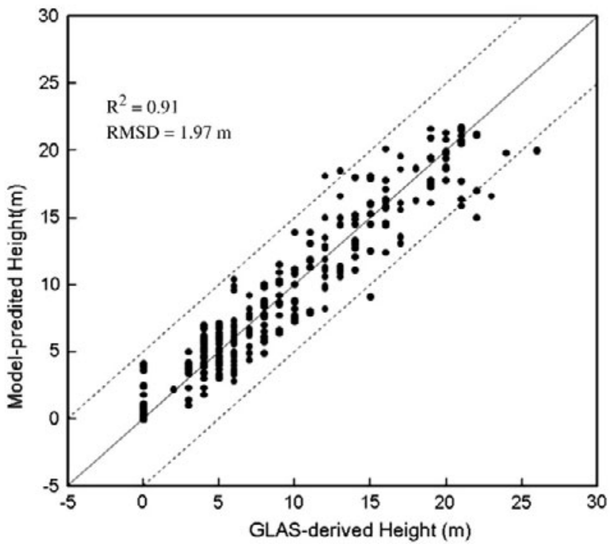

Fig. 9. Comparison of forest height values predicted by the best RT model developed in this study and those derived using GLAS data. Each point is a hold-out point used in the 10-fold cross validation (see Section 2.5.3 for details). The diagonal is the 1:1 line, and the dotted lines show the [−5, 5] range.

those forests. Previous studies revealed that VCT was capable of detecting most stand clearing disturbances and some partial disturbance events (Huang et al., 2010; Thomas et al., 2010), but future research is needed to achieve good separation between stand clearing events and partial disturbance events in the VCT products. For this study, we excluded the FIA plots that had stand age values of 20 years or more from further analysis. The remaining plots gave a much tighter relationship between modeled height values and the FIA measurements (Fig. 10(b)).

Fig. 10(b) reveals a systematic underestimation of about 5 m in forest height by the developed model. Considering the fact that forests in this area typically grow at about 1 m or more per year (Borders & Bailey, 2001; Carmean et al., 1989; Dolan et al., 2009), nearly half of the underestimation is likely the result of forest growth over 2–4 years, which are the differences between FIA field measurement year (2006–2007) and the model target year (2003–2004, see Section 2.3.1). The rest of the difference between FIA height measurements and the modeled MCH values may imply a general trend of underestimation of forest height when the MCH value is derived using the method shown in Fig. 5. While good agreement between lidar based forest height estimates and ground measurements has been reported in many studies (Carabajal & Harding, 2005; Dolan et al., 2010; Duncanson et al., 2010; Lefsky et al., 2001, 2005), underestimation by the method shown in Fig. 5 was reported or implied in several other studies (e.g. Hyde et al., 2005; Ni-Meister et al., 2010; Rosette et al., 2008). Such underestimation can be substantial for the "young" forests considered in this study, because these forests were composed mostly of pine trees that had conical shapes. The topmost portion of such trees may be too small to yield a detectable signal in the GLAS waveform data and therefore may not be measurable using GLAS data.

While cross validation reveals that about 90% of the variance of the GLAS based MCH value was explained by model predictions (Fig. 9), only 45% of the variance of the FIA height measurements was explained by model prediction (Fig. 10(b)). Such a difference between cross validation accuracy estimate and that derived using independent reference data is not likely the result of spatial automation (Friedl et al., 2002), because the selected GLAS samples did not appear to be spatially autorelated. Rather, it is likely due to the aggressive filtering of the GLAS samples designed to minimize the impact of terrain relief on the retrieval of forest height. Because only GLAS samples over flat areas were used in model development, the performance of the developed model was not controlled in areas with significant terrain relief. The substantially lower R^2 value derived using the FIA plot data may suggest that the developed model did not work well in areas with significant terrain relief.

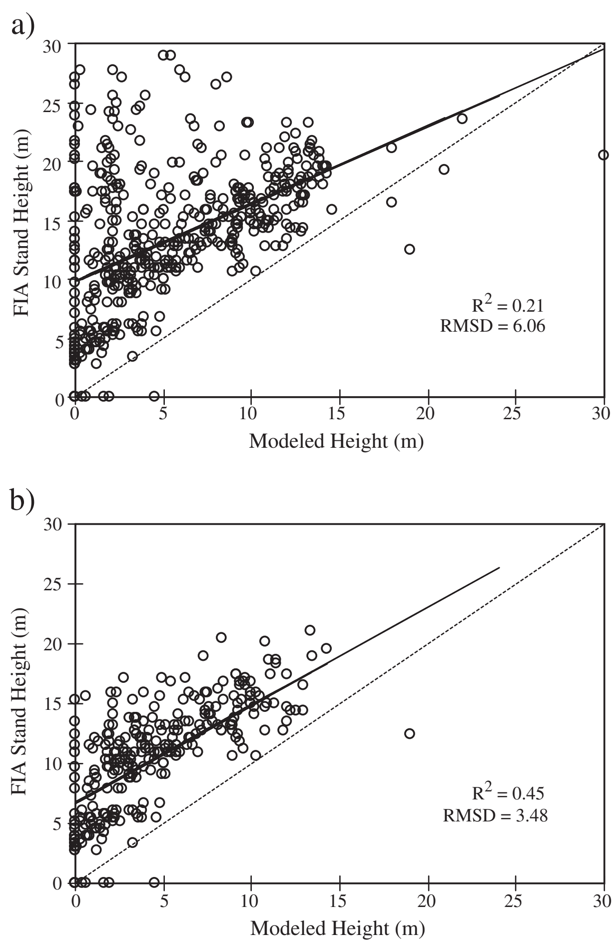

Fig. 10. Comparison of modeled forest height values with FIA stand height measurements for FIA plots with AGE_SD <20 years (a) and plots with both AGE_SD <20 years and FIA stand age <20 years.

4. Discussions and conclusions

Young forests are typically characterized by strong photosynthesis that exceeds respiration, providing a major carbon sink mechanism in the forest segment. Quantifying the height and growth rate is therefore critical for improved understanding of carbon pools and fluxes associated with these forests. This paper describes a new approach for modeling the height of "young" forest regenerated from previous disturbances by integrating time series Landsat observations with ICESat/GLAS lidar data. While there have been many studies exploring the use of Landsat images for forest height modeling (e.g. Franklin et al., 2003; Freitas et al., 2005; Hall et al., 2006; Jakubauskas & Price, 1997), the approach developed in this study is unique in its explicit use of age since disturbance (AGE_SD) and a spectral record of forest growth history represented by cumulative indices derived using the LTSS–VCT method. In Mississippi, about 90% and 70% of the variance of the maximum canopy height (MCH) values derived based on GLAS data were explained by age since disturbance (AGE_SD) and the cumulative indices, respectively. Less than 20% of that variance was explained by the Landsat images acquired during or near the acquisition years of the GLAS data. The best cross validation results were achieved when the AGE_SD and the cumulative indices were used together in an RT model. The Landsat images acquired during or near the acquisition years of the GLAS data did not provide additional explanatory power to this model.

When evaluated using field data collected through the FIA program, the predictions of the best model developed in this study were substantially less accurate than indicated by cross validation accuracy estimates, suggesting that major improvements to the developed modeling approach are needed in future studies. One of the error sources was the inability of the VCT to provide relatively reliable information on forest age when the mapped disturbances were non-stand clearing events. The actual age of a forest stand over a disturbed area mapped by VCT is often much older than that indicated by AGE_SD if a mapped disturbance was a non-stand clearing disturbance. This type of error can be reduced by separating stand clearing disturbances from and non-stand clearing events. The developed model should be applied only to the pixels mapped as stand clearing disturbances.

In this study, GLAS lidar data were used to derive the training samples needed for model development. While the GLAS has accumulated over 1.5 billion laser shots[3] distributed across the globe since its launch in 2003, only those having minimum terrain relief within their footprints allow relatively reliable retrieval of forest height using the method illustrated in Fig. 5. A model developed using training samples selected from flat areas may not work well for pixels having significant terrain relief within their footprints. Furthermore, in regions consisting mostly of rugged terrain, there may not be enough GLAS samples that have a flat ground footprint and therefore can be used for training data development using the method illustrated in Fig. 5. One way to mitigate this problem is to use lidar data with substantially smaller footprint sizes than the GLAS data to derive the required training data. Given the same slope, terrain relief within an area generally becomes smaller as the area size decreases. Depending on flight altitude, lidar data acquired by the airborne Laser Vegetation Imaging Sensor (LVIS) typically had ground footprint sizes of 10 m to 20 m (Blair et al., 1999). The lidar instrument onboard the planned Deformation, Ecosystem Structure, and Dynamics of Ice (DESDynI) will have a spatial resolution of 25 m (Freeman et al., 2009). The performance of the developed modeling approach likely will improve when such lidar data sets are available for training data development.

Where available, ground-based forest height measurements can also be used to improve the performance of the developed modeling approach. In particular, field measurements of forest height and other variables have been collected through many national and regional forest inventory programs. To achieve their inventory goals, such programs were typically designed to sample the landscape regardless of terrain relief (e.g. Smith, 2002). Therefore, the inventory data collected through such programs can represent a much wider range of terrain conditions than the training samples derived using GLAS data. In addition, the potential underestimation of forest height by the GLAS-based MCH values (see Section 3.3) will be removed or greatly reduced when field measurements are used as training data in the developed modeling approach. Finally, algorithms have been developed to improve forest height retrieval using GLAS waveform data by decoupling the signal from terrain and tree canopy (Lefsky et al., 2007). When such algorithms become mature enough to allow reliable retrieval of forest height over areas with different levels of terrain relief, available GLAS data will allow training data development over a much wider range of terrain conditions than allowed by using the method shown in Fig. 5.

Use of the LTSS–VCT method to calculate age since disturbance and the cumulative indices requires temporally dense, cloud free or near cloud free Landsat acquisition over a decade or longer. The availability of such Landsat acquisitions, however, varies from place to place. Based on knowledge gained through the North American Forest Dynamics (NAFD) project (Goward et al., 2008; Huang et al., 2009a) and the LANDFIRE project (Rollins, 2009), and an in-depth analysis of the USGS Landsat archive (Goward et al., 2006), the Landsat images needed for assembling annual or biennial LTSS exist for most areas in the U.S. Such data may also exist in many areas outside the U.S. However, less than half of all available Landsat images are in the USGS archive. The rest are in International Cooperator (IC) archives (Goward et al., 2006). An inventory of the Landsat images held in the IC archives is needed to determine which areas have the images required for assembling LTSS. In areas where available Landsat acquisitions are not adequate for assembling LTSS, images acquired by other instruments that have spatial and spectral characteristics similar to those of the Landsat may be used in the place of Landsat data, provided that the VCT can be adapted for handling a mixture of images from different instruments.

[3] Based on information at http://nsidc.org/data/icesat/visge/. Last visited on November 11, 2010.

Acknowledgments

Funding support for this study was provided by the U.S. Geological Survey, and by NASA's Terrestrial Ecology, Carbon Cycle Science, and Applied Sciences Programs. Partial support was also provided by the Knowledge Innovation Program of the Chinese Academy of Sciences (grant no. KZCX2-YW-QN313). It contributes to the North American Carbon Program, and the joint USDA-DOI LANDFIRE project sponsored by the intergovernmental Wildland Fire Leadership Council of the United States. Portions of this work were performed in accordance with a memorandum of understanding between the Forest Inventory and Analysis program of USDA Forest Service and the interagency LANDFIRE program, in support of vegetation structure mapping in LANDFIRE.

References

Abshire, J. B., Sun, X. L., Riris, H., Sirota, J. M., McGarry, J. F., Palm, S., et al. (2005). Geoscience Laser Altimeter System (GLAS) on the ICESat mission: On-orbit measurement performance. *Geophysical Research Letters*, *32*.

Amaro, A., Reed, D., & Soares, P. (Eds.). (2003). *Modelling forest ecosystems* (pp. 401). Cambridge, MA: CABI Publishing.

Bachman, C. G. (1979). *Laser radar systems and techniques.* Norwood, MA: Artech House.

Balzter, H. (2001). Forest mapping and monitoring with interferometric synthetic aperture radar (InSAR). *Progress in Physical Geography*, *25*, 159—177.

Balzter, H., Rowland, C. S., & Saich, P. (2007). Forest canopy height and carbon estimation at Monks Wood National Nature Reserve, UK, using dual-wavelength SAR interferometry. *Remote Sensing of Environment*, *108*, 224—239.

Blackard, J. A., Finco, M. V., Helmer, E. H., Holden, G. R., Hoppus, M. L., Jacobs, D. M., et al. (2008). Mapping US forest biomass using nationwide forest inventory data and moderate resolution information. *Remote Sensing of Environment*, *112*, 1658—1677.

Blair, J. B., Rabine, D. L., & Hofton, M. A. (1999). The Laser Vegetation Imaging Sensor (LVIS): A medium-altitude, digitization-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*, 115—122.

Borders, B. E., & Bailey, R. L. (2001). Loblolly Pine — Pushing the limits of growth. *Southern Journal of Applied Forestry*, *25*, 69—74.

Breiman, L., Friedman, J. H., Olshend, R. A., & Stone, C. J. (1984). *Classification and regression trees* (pp. 358). Belmont, California: Wadsworth International Group.

Brenner, A. C., Zwally, H. J., Bentley, C. R., Csathó, B. M., Harding, D. J., Hofton, M. A., et al. (2003). *Derivation of range and range distributions from laser pulse waveform analysis for surface elevations, roughness, slop, and vegetation heights: Geosciences Laser Altimeter System (GLAS) Algorithm Theoretical Basis Document (Version 4.1).* Geosciences Laser Altimeter System (GLAS): Algorithm Theoretical Basis Document (Version 4.1): National Aeronautics and Space Administration.

Brown, S. L., & Schroeder, P. E. (1999). Spatial patterns of aboveground production and mortality of woody biomass for eastern U.S. forest. *Ecological Applications*, *9*, 968—980.

Carabajal, C. C., & Harding, D. J. (2005). ICESat validation of SRTM C-band digital elevation models. *Geophysical Research Letters*, *33*, L22S01.

Carmean, W. H., Hahn, J. T., & Jacobs, R. D. (1989). *Site index curves for forest tree species in the Eastern United States.* St. Paul, Minnesota: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station.

Chen, X., Zhu, Z., Ohlen, D., Huang, C., & Shi, H. (2008, November 18–20). Use of multiple spectral indices to estimate burn severity in the Black Hills of South Dakota. *Pecora 17 — The future of land imaging … going operational.* ASPRS (American Society for Photogrammetry and Remote Sensing), Denver, Colorado.

Chopping, M., Moisen, G. G., Su, L. H., Laliberte, A., Rango, A., Martonchik, J. V., et al. (2008). Large area mapping of southwestern forest crown cover, canopy height, and biomass using the NASA Multiangle Imaging Spectro-Radiometer. *Remote Sensing of Environment*, *112*, 2051−2063.

Cohen, W. B., Fiorella, M., Bray, J., Helmer, E., & Anderson, K. (1998). An efficient and accurate method for mapping forest clearcuts in the Pacific Northwest using Landsat imagery. *Photogrammetric Engineering and Remote Sensing*, *64*, 293−300.

Coops, N. C. (2002). Eucalypt forest structure and synthetic aperture radar backscatter: A theoretical analysis. *Trees*, *16*, 28−46.

De La Cueva, A. V. (2008). Structural attributes of three forest types in central Spain and Landsat ETM plus information evaluated with redundancy analysis. *International Journal of Remote Sensing*, *29*, 5657−5676.

De'ath, G., & Fabricius, K. E. (2000). Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology*, *81*, 3178−3192.

Dolan, K. A., Hurtt, G. C., Chambers, J. Q., Dubayah, R. O., Frolking, S., & Masek, J. G. (2010). Using ICESat's Geoscience Laser Altimeter System (GLAS) to assess large-scale forest disturbance caused by hurricane Katrina. *Remote Sensing of Environment*, *115*, 86−96.

Dolan, K., Masek, J. G., Huang, C., & Sun, G. (2009). Regional forest growth rates measured by combining ICESAT GLAS and Landsat data. *Journal of Geophysical Research, Biogeosciences*, *114*, G00E05, doi:10.1029/2008JG000893.

Donoghue, D. N. M., & Watt, P. J. (2006). Using LiDAR to compare forest height estimates from IKONOS and Landsat ETM+ data in Sitka spruce plantation forests. *International Journal of Remote Sensing*, *27*, 2161−2175.

Donoghue, D. N. M., Watt, P. J., Cox, N. J., Dunford, R. W., Wilson, J., Stables, S., et al. (2004). An evaluation of the use of satellite data for monitoring early development of young Sitka spruce plantation forest growth. *Forestry*, *77*, 383−396.

Dubayah, R. O., & Drake, J. B. (2000). Lidar remote sensing for forestry. *Journal of Forestry*, *98*, 44−46.

Duncanson, L. I., Niemann, K. O., & Wulder, M. A. (2010). Estimating forest canopy height and terrain relief from GLAS waveform metrics. *Remote Sensing of Environment*, *114*, 138−154.

Effroymson, M. A. (1960). Multiple regression analysis. In A. Ralston, & H. S. Wilf (Eds.), *Mathematical methods for digital computers* (pp. 191−203). New York: Wiley.

Escuin, S., Navarro, R., & Fernandez, P. (2008). Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images. *International Journal of Remote Sensing*, *29*, 1053−1073.

Faulkner, J. L., Miller, P. E., Hartsell, A. J., & London, J. D. (1993). *Forest statistics for south Mississippi counties, 1994. forest service resource bulletin.* Technical Report PB-94-142379/XAB; FSRB-SO–185. New Orleans, LA: USDA Forest Service Southern Forest Experiment Station.

Franklin, S. E., Hall, R. J., Smith, L., & Gerylo, G. R. (2003). Discrimination of conifer height, age and crown closure classes using Landsat-5 TM imagery in the Canadian Northwest Territories. *International Journal of Remote Sensing*, *24*, 1823−1834.

Freeman, A., Rosen, P., Jordan, R., Johnson, W. T. K., Hensley, S., Sweetser, T., et al. (2009, January 26–30). DESDYNI — A NASA mission for ecosystems, solid earth, and cryosphere science. *Proc. of '4th int. workshop on science and applications of SAR polarimetry and polarimetric interferometry — PolInSAR 2009'* Frascati, Italy.

Freitas, S. R., Mello, M. C. S., & Cruz, C. B. M. (2005). Relationships between forest structure and vegetation indices in Atlantic Rainforest. *Forest Ecology and Management*, *218*, 353−362.

Friedl, M. A., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., et al. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, *83*, 287−302.

Garestier, F., & Dubois-Fernandez, P. C. (2008). Forest height inversion using high-resolution P-band Pol-InSAR data. *IEEE Transactions on Geoscience and Remote Sensing*, *46*, 3544−3559.

Goward, S., Irons, J., Franks, S., Arvidson, T., Williams, D., & Faundeen, J. (2006). Historical record of landsat global coverage: Mission operations, NSLRSDA, and international cooperator stations. *Photogrammetric Engineering and Remote Sensing*, *72*, 1155−1169.

Goward, S. N., Masek, J. G., Cohen, W., Moisen, G., Collatz, G. J., Healey, S., et al. (2008). Forest disturbance and North American carbon flux. *EOS. Transactions of the American Geophysical Union*, *89*, 105−106.

Hall, R. J., Skakun, R. S., Arsenault, E. J., & Case, B. S. (2006). Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. *Forest Ecology and Management*, *225*, 378−390.

Harding, D. J., & Carabajal, C. C. (2005). ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure. *Geophysical Research Letters*, *32*, 1−4.

Harding, D. J., Lefsky, M. A., Parker, G. G., & Blair, J. B. (2001). Lidar altimeter measurements of canopy height profiles: Methods and validation for closed canopy, broadleaf forests. *Remote Sensing of Environment*, *76*, 283−297.

Healey, S. P., Cohen, W. B., Zhiqiang, Y., & Krankina, O. N. (2005). Comparison of tasseled cap-based Landsat data structures for use in forest disturbance detection. *Remote Sensing of Environment*, *97*, 301.

Hinsley, S. A., Hill, R. A., Gaveau, D. L. A., & Bellamy, P. E. (2002). Quantifying woodland structure and habitat quality for birds using airborne laser scanning. *Functional Ecology*, *16*, 851−857.

Houghton, R. A. (2005). Aboveground forest biomass and the global carbon balance. *Global Change Biology*, *11*, 945−958.

Huang, C., Goward, S. N., Masek, J. G., Gao, F., Vermote, E. F., Thomas, N., et al. (2009a). Development of time series stacks of Landsat images for reconstructing forest disturbance history. *International Journal of Digital Earth*, *2*, 195−218.

Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., & Vogelmann, J. E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*, *114*, 183−198.

Huang, C., Goward, S. N., Schleeweis, K., Thomas, N., Masek, J. G., & Zhu, Z. (2009b). Dynamics of national forests assessed using the Landsat record: case studies in eastern U.S. *Remote Sensing of Environment*, *113*, 1430−1442.

Huang, C., Song, K., Kim, S., Townshend, J. R. G., Davis, P., Masek, J., et al. (2008). Use of a dark object concept and support vector machines to automate forest cover change analysis. *Remote Sensing of Environment*, *112*, 970−985.

Huang, C., & Townshend, J. R. G. (2003). A stepwise regression tree for nonlinear approximation: Applications to estimating subpixel land cover. *International Journal of Remote Sensing*, *24*, 75−90.

Hudak, A. T., Lefsky, M. A., Cohen, W. B., & Berterretche, M. (2002). Integration of LIDAR and Landsat ETM+ data for estimating and mapping forest canopy height. *Remote Sensing of Environment*, *82*, 397−416.

Hyde, P., Dubayah, R., Peterson, B., Blair, J. B., Hofton, M., Hunsaker, C., et al. (2005). Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. *Remote Sensing of Environment*, *96*, 427−437.

Hyyppa, J., Hyyppa, H., Inkinen, M., Engdahl, M., Linko, S., & Zhu, Y. H. (2000). Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management*, *128*, 109−120.

Jakubauskas, M. E., & Price, K. P. (1997). Empirical relationships between structural and spectral factors of Yellowstone lodgepole pine forests. *Photogrammetric Engineering and Remote Sensing*, *63*, 1375−1381.

Kennedy, R. E., Cohen, W. B., & Schroeder, T. A. (2007). Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment*, *110*, 370−386.

Kennedy, R. E., Yang, Z. G., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms. *Remote Sensing of Environment*, *114*, 2897−2910.

Lefsky, M. A., Carabajal, C. C., Del Bom Espirito-Santo, F., Hunter, M. O., de Oliveira, R., Jr., Harding, D. J., et al. (2005). Estimates of forest canopy height and aboveground biomass using ICESat.*Geophysical Research Letters*, *32*, 1−4 Article Number L22S02.

Lefsky, M. A., Cohen, W. B., Parker, G. G., & Harding, D. J. (2002). Lidar remote sensing for ecosystem studies. *Bioscience*, *52*, 19−30.

Lefsky, M. A., Cohen, W. B., & Spies, T. A. (2001). An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Canadian Journal of Forest Research*, *31*, 78−87.

Lefsky, M. A., Keller, M., Pang, Y., Camargo, P. B. d., & Hunter, M. O. (2007). Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms. *Journal of Applied Remote Sensing*, *1*, 013537.

Li, M., Huang, C., Zhu, Z., Shi, H., Lu, H., & Peng, S. (2009a). Assessing rates of forest change and fragmentation in Alabama, USA, using the vegetation change tracker model. *Forest Ecology and Management*, *257*, 1480−1488.

Li, M., Huang, C., Zhu, Z., Shi, H., Lu, H., & Peng, S. (2009b). Use of remote sensing coupled with a vegetation change tracker model to assess rates of forest change and fragmentation in Mississippi, USA. *International Journal of Remote Sensing*, *30*, 6559−6574.

Morgan, C., DeMatteis, J., & Barber, E. (2007). *Mississippi's forest legacy program-assessment of need 2007–2012*. Jackson, Mississippi: Mississippi Forestry Commission.

Mougin, E., Proisy, C., Marty, G., Fromard, F., Puig, H., Betoulle, J. L., et al. (1999). Multifrequency and multipolarization radar backscattering from mangrove forests. *IEEE Transactions on Geoscience and Remote Sensing*, *37*, 94−102.

Næsset, E., & Økland, T. (2002). Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment*, *79*, 105−115.

Nelson, R., Boudreau, J., Gregoire, T. G., Margolis, H., Naesset, E., Gobakken, T., et al. (2009). Estimating Quebec provincial forest resources using ICESat/GLAS. *Canadian Journal of Forest Research-Revue Canadienne de Recherche Forestiere*, *39*, 862−881.

Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment*, *56*, 1−7.

Ni-Meister, W., Lee, S. Y., Strahler, A. H., Woodcock, C. E., Schaaf, C., Yao, T. A., et al. (2010). Assessing general relationships between aboveground biomass and vegetation structure parameters for improved carbon estimate from lidar remote sensing. *Journal of Geophysical Research, Biogeosciences*, *115*, 12.

NRC (2007). *Earth science and applications from space: National imperatives for the next decade and beyond.* Washington D.C.: The National Academies Press.

Oswalt, S. N., & Oswalt, C. M. (2008). Relationships between common forest metrics and realized impacts of Hurricane Katrina on forest resources in Mississippi. *Forest Ecology and Management*, *255*, 1692−1700.

Pascual, C., Garcia-Abril, A., Cohen, W. B., & Martin-Fernandez, S. (2010). Relationship between LiDAR-derived forest canopy height and Landsat images. *International Journal of Remote Sensing*, *31*, 1261−1280.

Pinder, J., & Rea, T. (1999). Deforestation, reforestation and forest fragmentation on the upper coastal plain of South Carolina and Georgia. *American Midland Naturalist*, *142*, 213−228.

Pollard, J. E., Westfall, J. A., Patterson, P. L., Gartner, D. L., Hansen, M., & Kuegler, O. (2006). *Forest Inventory and Analysis national data quality assessment report for 2000 to 2003.* Gen. Tech. Rep. RMRS-GTR-181. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Pretzsch, H. (2001). Models for pure and mixed forests. In J. Evans (Ed.), *The forest handbook*, *Vol. 1*. (pp. 210−228)London: Blackwell Science Ltd.

Puhr, C. B., & Donoghue, D. N. M. (2000). Remote sensing of upland conifer plantations using Landsat TM data: A case study from Galloway, south-west Scotland. *International Journal of Remote Sensing*, *21*, 633−646.

Quinlan, J. R. (1993, June 27–29). Combining instance based and model-based learning. *The 10th international machine learning conference* (pp. 236−243). Amherst, MA, USA: University of Massachusetts.

Quinlan, J. R. (1996, August 4–8). Bagging, boosting, and C4.5. *The thirteenth national conference on artificial intelligence and eighth innovative applications of artificial intelligence conference* Portland, Oregon.

Rabus, B., Eineder, M., Roth, A., & Bamler, R. (2003). The shuttle radar topography mission — A new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing*, *47*, 241−262.

Rollins, M. G. (2009). LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment. *International Journal of Wildland Fire*, *18*, 235−249.

Rosette, J. A. B., North, P. R. J., & Suarez, J. C. (2008). Vegetation height estimates for a mixed temperate forest using satellite laser altimetry. *International Journal of Remote Sensing*, *29*, 1475−1493.

Schutz, B. E., Zwally, H. J., Shuman, C. A., Hancock, D., & DiMarzio, J. P. (2005). Overview of the ICESat mission. *Geophysical Research Letters*, *32*.

Scott, J. H., & Burgan, R. E. (2005). *Standard fire behavior fuel models: a comprehensive set for use with Rothermel's surface fire spread model.* Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Sirota, J. M., Bae, S., Millar, P., Mostofi, D., Webb, C., Schutz, B., et al. (2005). The transmitter pointing determination in the Geoscience Laser Altimeter System. *Geophysical Research Letters*, *32*.

Siry, J. P. (2004). Southern forest resource conditions and management practices from 1950–2000: Benefits of research. In H. M. Rauscher, & K. Johnsen (Eds.), *Southern forest science: past, present, and future* (pp. 23−31). Asheville, NC: Department of Agriculture, Forest Service, Southern Research Station.

Smith, W. B. (2002). Forest inventory and analysis: a national inventory and monitoring program. *Environmental Pollution*, *116*, S233−S242.

Sun, G., Ranson, K. J., Khairuk, V. I., & Kovacs, K. (2003). Validation of surface height from shuttle radar topography mission using shuttle laser altimeter. *Remote Sensing of Environment*, *88*, 401−411.

Sun, G., Ranson, K. J., Kimes, D. S., Blair, J. B., & Kovacs, K. (2008). Forest vertical structure from GLAS: An evaluation using LVIS and SRTM data. *Remote Sensing of Environment*, *112*, 107−117.

Thomas, N., Huang, C., Goward, S. N., Powell, S., Rishmawi, K., Schleeweis, K., et al. (2010). Validation of North American forest disturbance dynamics derived from Landsat time series stacks. *Remote Sensing of Environment*, *115*, 19−32.

USDA Forest Service (2005). Forest inventory and analysis national core field guide, version 3.0. available at. http://fia.fs.fed.us/library/field-guides-methods-proc/docs/2006/core_ver_3-0_10_2005.pdf Date accessed: January 25, 2011.

von Gadow, K., & Hui, G. (1999). *Modelling forest development. Forestry Science*, *57*. (pp. 213) Dordrecht: Kluwer Academic Publishers.

Walker, L. C. (1994). The southern pine region. In J. W. Barrett (Ed.), *Regional silviculture of the United States* (pp. 271−333). New York: Wiley.

Walker, W. S., Kellndorfer, J. M., LaPoint, E., Hoppus, M., & Westfall, J. (2007). An empirical InSAR-optical fusion approach to mapping vegetation canopy height. *Remote Sensing of Environment*, *109*, 482−499.

Zarnetske, P. L., Edwards, T. C., Jr., & Moisen, G. G. (2007). Habitat classification modeling with incomplete data: Pushing the habitat envelope. *Ecological Applications*, *17*, 1714−1726.