ESTIMATION OF FOREST STRUCTURAL PARAMETERS FROM LIDAR AND SAR DATA

Z. Zhang [a, b, *], W. Ni [b], A. Fu [b], Z. Guo [b], Guoqing Sun [c], and D. Wang [b]

[a] Remote Sensing and GIS Research Center, Beijing Normal University, Beijing, china - lwzzzy@gmail.com
[b] State Key Laboratory of Remote Sensing, Institute of Remote Sensing Applications, Chinese Academy of Sciences, P. O. Box 9718, Beijing, 100101, China - zhifeng_guo@hotmail.com
[c] Dept. of Geography and ESSIC, University of Maryland, College Park, MD USA - guoqing.sun@gmail.com

Commission VIII, WG VIII/11

KEY WORDS: Forest structure, Lidar, Radar, GLAS, PALSAR

ABSTRACT

Vegetation spatial structure including plant height, biomass, vertical and horizontal heterogeneity, is an important factor influencing the exchanges of matter and energy between landscape and atmosphere, and the biodiversity of ecosystems. Regional and global forest biomass and forest structure estimation is essential for understanding and monitoring ecosystem responses to human activities and climate change. Lidars with capabilities of recording the time-varying return signals provide vegetation height, ground surface height, and the vertical distribution of vegetated surfaces intercepted by laser pulses, while radar responds to the amount of water in a forest canopy, as well as its spatial structure. Large footprint lidar has been shown to be an effective technique for measuring forest canopy height, and biomass from space. Data from these sensors contain information relevant to different aspects of the biophysical properties of the vegetation canopy. Synergy of lidar's vertical profile sampling and radar's mapping capabilities provides new opportunity for regional and global mapping of forest structural parameters.Field sampling data, lidar data, PALSAR and InSAR data were obtained and processed in our study area in Changbai Mountain area, northeastern China. The capabilities for forest height and biomass estimation through combined use of lidar and radar data with lidar footprints scale were investigated. The results show that combined lidar and radar data improved the prediction accuracy by individual sensors, which was validated using field measurements and forest inventory data.

1. INTRODUCTION

The terrestrial ecosystem is currently undergoing dramatic changes coinciding with global warming in recent decades. Researches have shown that 3D vegetation structure exerts a strong influence on radiation transmission, which changes radiation and energy balance on the land surface. Vegetation spatial structure including plant height, biomass, vertical and horizontal heterogeneity, is an important factor influencing the exchanges of matter and energy between the landscape and atmosphere, and the biodiversity of ecosystems. Estimates of regional and global forest biomass and forest structure are essential for understanding and monitoring ecosystem responses to human activities and climate change. Lidars with capabilities of recording the time-varying return signals provide the vegetation height, ground surface height, and the vertical distribution of vegetated surfaces intercepted by laser pulses. Large footprint lidar has been shown to be an effective technique for measuring forest canopy height, and biomass from space. Essentially, radar responds to the amount of water in a forest canopy, as well as its spatial structure. Data from these sensors contain information relevant to different aspects of the biophysical properties of the vegetation canopy.

Researchers have started using GLAS data for forest studies (Harding & Carabajal, 2005; Ranson et al., 2004a, b). Tree heights and forest biomass estimation using GLAS data has been tried in several studies (Lefsky, 2005; Sun et al., 2008a,b). In this study, the synergy of GLAS and PALSAR data for forest parameters estimation were studied. The combination of lidar (GLAS) and radar (PALSAR) data in forest structural parameter estimation was investigated over the test site in Changbai Mountain, Northern China.

2. DATA AND TEST SITE

The Geoscience Laser Altimeter System (GLAS) instrument aboard the Ice, Cloud, and land Elevation (ICESat) satellite, launched on 12 January 2003, and it is the first lidar instrument designed for continuous global observation of the Earth (Zwally et al., 2002). GLAS data acquired in fall 2003 (L2A) were used in this study. GLA01 L-1 waveform data and GLA14 L-2 land surface products were used in this study. From GLAS GLA14 data, the record index, the serial number of the shot within the index, acquisition time, latitude, longitude, elevation, range offsets of signal beginning, signal ending, waveform centroid, and fitted Gaussian peaks of a GLAS footprint were retrieved. The original waveform and some associated parameters were extracted from GLA01 record and were processed to generate various waveform indices.

PALSAR data, in the form of dual-pol InSAR pair, were acquired in June and July, 2007. SRTM data, which is also used in this study, was the first spaceborne fixed-baseline InSAR and operated two InSAR instruments including the United States C-band (5.6cm, 5.3GHz) sensor and the German X-band (3.1cm, 9.6GHz) sensor (USGS, 2003; Kellndorfer et al., 2004). Here we used the C-band data to get the scattering phase center height of forest canopy associate with DEM.

[*] Corresponding author. E-mail address: lwzzzy@gmail.com (Z. Zhang)

Changbai mountain Natural Preserve (42.8° N, 128.5° E) and nearby Luoshuihe Bureau of Forestry (42.5° N, 127.8° E) in Northern China provide a good test site for forest structure studies. The broadleaf-korean pine (Pinus koraiensis) mixed forest is the most diversified forest in species and ecosystem, and is the most productive one in various resources in northern China. In this test site, tree dbh, height in 30 GLAS shots were obtained through field measurement in June, 2007 (Figure 1). The geographic location of each GLAS footprint was located by GPS. Four sampling plots with a radius of 7.5m within a GLAS footprint were sampled. One plot is at the center, and the others were 22.5m from center to north, southeast and southwest directions, respectively. Diameter at breast height (DBH), heights to the top and base of crown for all trees with DBH > 3cm were measured.
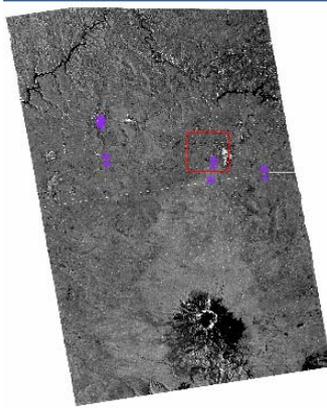

Figure1. PALSAR L-HH backscattering coefficients image of Changbai Mountain area overlaid by GLAS footprints, where the trees were measured in field in 2007.

3. DATA PROCESSING

3.1 InSAR Data Processing

Some researches have been conducted for estimating forest height from InSAR data. Kellndorfer et al.(2004) combined SRTM and NED（National Elevation Datasets） data to estimate vegetation height, and received good results. In this paper, we reported a similar method to estimate forest parameter. Using a pair of PALSAR data of the same scene but at a slightly different radar location, the interferegram as well as the coherence of the scenes can be generated. The surface elevation model (DEM) then can be generated from the unwrapped interferogram. If the bare surface DEM is available, the difference between InSAR DEM and bare surface DEM will be the height of the scattering phase center. In this study, SRTM DEM was used. We knew that C-band SRTM DEM data represents the scattering phase center height of forest area at C-band. Because L-band SAR has more penetration into forests than C-band, the DEM from PALSAR data represents the lower elevation than SRTM DEM. The 3D structure of forest stand will affect the scattering center at both wavelengths. So the difference of SRTM and PALSAR DEM would include some information on the stand structure:

$$H_{C-L} = H_{SRTM} - H_{InDEM} \qquad (1)$$

Where H_{C-L} is scattering phase center height differences between C-band and L-band; H_{SRTM} is SRTM DEM；and H_{InDEM} is PALSAR DEM.

This elevation differences H_{C-L} are indicators of the canopy structure (Sarabandi & Lin, 2000). We noticed that the SRTM data and PALSAR DEM use different coordinates system. The SRTM data are referenced to EGM96 Geoid, but PALSAR DEM is referenced to WGS84 ellipsoid. So, we chose some bare points in these two DEM images to calculate the difference between SRTM and PALSAR DEM, and developed a correction grids, which makes the elevation difference at bare surface of these DEMs the same and is used to correct the PALSAR DEM. Through this processing, part of vertical error on PALSAR DEM could be removed, and a differential DEM scene was generated. Test site of Luoshuihe Bureau of Forestry is covered by two PALSAR images shown in Figure 1. Both of them were corrected using the differences between the bare points.

The backscattering coefficients of HH and HV polarizations were also extracted from PALSAR data. To calculate the backscattering coefficients of PALSAR data, geometric correction was an important step. The pixel size of SRTM DEM data used for correction is 90 m. In the co-registration of SRTM and PALSAR DEM, both were re-sampled to 28.5m as the ETM+ images. ETM+ images were used to help to select the control points. From previous studies, there is a correlation between radar backscattering coefficient and forest biomass, these data will be used in statistical analysis.

3.2 GLAS Data Processing

From GLAS GLA14 data, all necessary information were retrieved from the record index, and the serial number of the shot within the index, including acquisition time, latitude, longitude, elevation, range offsets of signal beginning, signal ending, waveform centroid, and fitted Gaussian peaks. Since the elevation in GLA14 is the height of the waveform centroid, it has to be converted into ground surface height (the last Gaussian peak, i.e. ground peak). The offset difference between signal beginning and ending is defined as waveform length, and the distance between the signal beginning and the last Gaussian peak is defined as the top tree height (Sun et al., 2008).

The waveform extracted from GLA01 data was first filtered by Gaussian filter with a width similar to the transmitted laser pulse. The GLA01 product gives the estimated noise level, i.e. the mean and standard deviation of background noise values in the waveform. In many cases, the noise level before the signal beginning was lower than the noise after the signal ending. Consequently, we estimated the noise levels before the signal beginning and after the signal ending from the original waveform separately using a method based on the histogram. Using three standard deviations as a threshold above the noise level, the signal beginning and ending were located. The total waveform energy was calculated by summing all the return energy from signal beginning to ending. Starting from the signal ending, the position of the 25%, 50%, and 75% of energy were located by comparing the accumulated energy with total energy. Since the heights of these quartiles refer to the ground surface, not the signal ending, the ground peak in the waveform needs to be located. Searching backward from the signal ending, the peaks can be found by comparing a bin's value with those of the two neighboring bins. If the first peak is too close to the signal ending, i.e. the distance from signal ending to the peak is less than the half width of the transmitted laser pulse, this peak was

discarded. The first significant peak found is the ground peak. With terrain slope and surface roughness increasing, the ground peak of the waveform becomes wider and the signal beginning moves upwards in a proportional manner. The distance between the signal ending and the assumed signal ending was used as an adjustment to the signal beginning.

4. DATA ANALYSIS

The diameter of GLAS footprint is about 70 m, and the spatial resolution of PALSAR data is about 20-30m. The resolution of SRTM DEM is about 90m with 3 arc second. PALSAR and SRTM data were co-registered and re-sampled into 28.5m pixel size. Therefore the radar data (H_{C-L}, LHH and LHV backscattering coefficients) were extracted using 3 by 3 window. Four values were calculated: value of the center pixel, maximum, minimum and mean values of the 3 by 3 window.

Variables	Description
ht	Calculated top canopy height from GLA01 data (slope corrected)
h_{25}, h_{50}, h_{75}, h_{100}	Height of waveform energy quartiles calculated from GLA01 waveform data
wflen, h_{14}	Total length of waveform and top canopy height from GLA14 data
H_{C-L}	Phase center height between L and C-band
σhh_0, σhh_{stdev} σhv_0, σhv_{stdev}	Mean and standard deviation of radar backscattering coefficient for 90 m pixel
$\sigma hh_{0-28.5}$, $\sigma hh_{stdev-28.5}$ $\sigma hv_{0-28.5}$, $\sigma hv_{stdev-28.5}$	Mean and standard deviation of radar backscattering coefficient for 28.5 m pixel

Table.1. GLAS waveform indices and radar signature used for forest parameters retrieval

Table 1 lists the variables derived from GLAS and PALSAR SAR data. Multi-variable correlation analyses were performed using GLAS data or SAR data only, and combined GLAS and SAR data respectively. The step-wise regression is performed using S-plus software.

Through data processing, we constructed the regression relationships or prediction models for estimation of maximum tree height and biomass at GLAS footprints from GLAS and SAR data. The models and related prediction accuracy were discussed below.

4.1 Tree Height Data

When using GLAS data only, the step-wise regression picked five variables (wflen, ht, h_{25}, h_{75}, and h_{100}) from seven GLAS indices list in Table 1. Figure 2 shows the comparison of GLAS predicted maximum tree heights with field measurements. The R^2 is 0.715. The tree height from GLAS waveform is the distance from the signal beginning to the ground peak. Because of the certain width of the transmitted lidar signal, only trees above certain height can be estimated. This limited the range of the tree heights, and reduced the correlation.

If only SAR data were used, two variables (H_{C-L} and σhh_0) were selected, and the R^2 of the regression is only 0.314 and the RSE

(residual standard error) is 4.82m. There may be several reasons for this phenomenon. One is the mis-registration because these points were on the edge of the PALSAR scene. Secondly, we only have 30 GLAS shots and need more for a better correlation.

When GLAS and SAR data combined, the regression model picked four variables (wflen, h_{25}, h_{75}, and H_{C-L}) and the prediction model is as follows:

$$maxHt = 29.93 - 0.46 \, wflen - 1.09 \, h_{25} + 0.90 \, h_{75} + 0.13 \, H_{C-L} \quad (2)$$

and $R^2=0.767$, RSE=2.415m.

Figure 3 shows the predicted maximum tree height from combined GLAS and SRA data. The combined GLAS and SAR data improved the prediction accuracy.

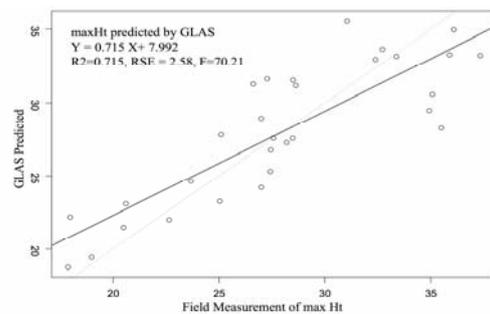Fig. 2.Predicted maximum tree heights vs. field measurements using GLAS waveform indices only. Dash line if 1:1 line.
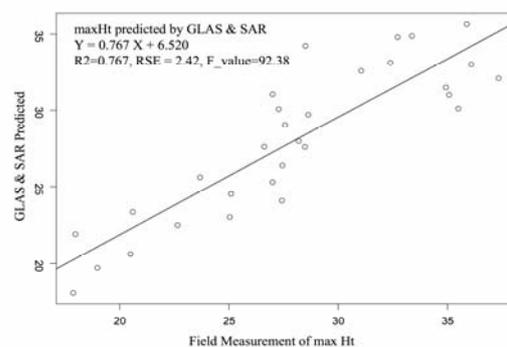

Figure.3. Predicted biomass using GLAS and SAR data together vs. field measurements

4.2 Biomass Data

When using GLAS data only, the step-wise regression picked four variables (h_{25}, h_{50} h_{100} and h_{t4}) from GLAS indices. Figure 4 shows the comparison of GLAS predicted biomass with field measurements. The R^2 is 0.568 (Fig. 4). We found that the two

points with high biomass level made the relation worse. There are problems in sampling or calculating biomass for these two points. The values are too high for the forests in this region, so these two high biomass points were eliminated. Then a

step-wise regression was used over again, three variables (wflen, h_{25} and h_{75}) were selected. Figure 5 shows the comparison of picked GLAS data predicted biomass and field measurements and $R^2=0.642$, and RSE=6.373 Kg/m^2.
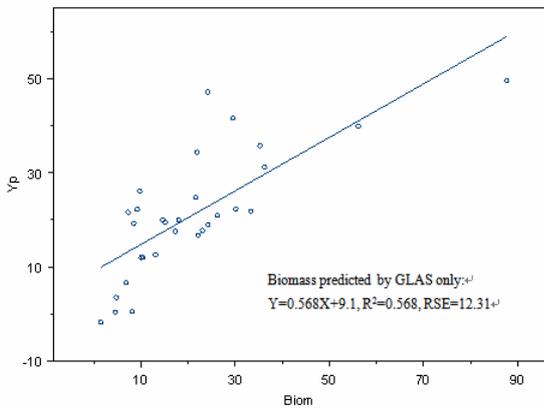

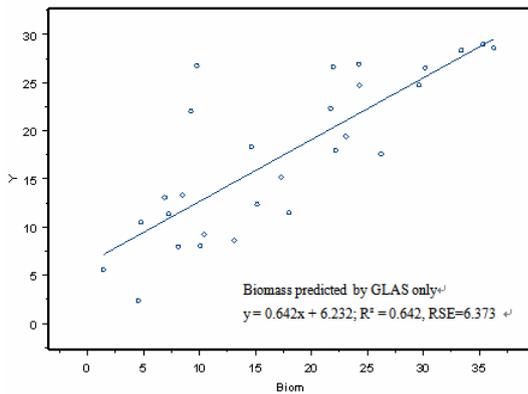Figure.4. Predicted biomass using GLAS only vs. field measurements.

Figure.5. Predicted biomass using GLAS data vs. field measurements after two high points were eliminated.

If only SAR data were used, three variables (H_{C-L}, σhh_0 and $\sigma hh_{0-28.5}$) were selected, and the R^2 is only 0.39. Just like the tree heights prediction using SAR data, too few GLAS shots got was the reason for this. After eliminating the high biomass points, the prediction using SAR data were still not well with a $R^2=0.43$.

When GLAS and SAR data combined, a step-wise regression model picked six variables (H_{C-L}, σhv_0, $\sigma hv_{0-28.5}$, h_{25}, h_{75} and h_{100}) and the prediction model is:

Biomass=25.0862+ 0.2460 H_{C-L} - 3.1876 σhv_0 +3.5214 $\sigma hv_{0-28.5}$+1.7030 h_{75}-1.9080h_{25} -0.8232wflen   (3) and $R^2=0.8012$, RSE=5.081.

Figure 6 shows the predicted biomass vs. field measurements using the combined model, and from the figure, we found the model had much high prediction accuracy than using one sensor only.
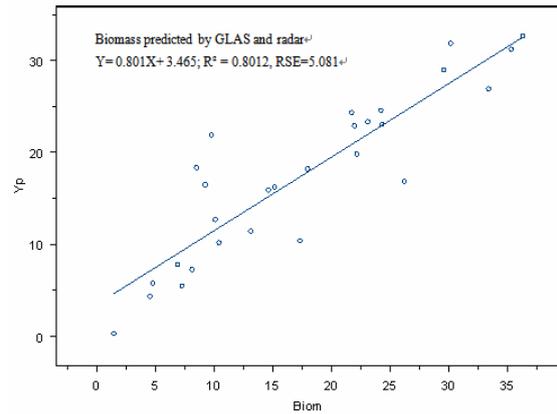

Figure.7. Predicted biomass using GLAS and SAR data vs. field measurements after two high points were eliminated.

.

5. CONCLUSION

This paper presented preliminary results of estimation of forest canopy height and biomss using GLAS and radar data. The results show that GLAS data has better prediction capability than PALSAR data, and the combined use of lidar and radar data provides the best prediction results. There still are some issues for data processing and regression. Mis-registration may be a primary problem for poor predicting results using PALSAR data while surface slope poses a problem for both SAR and lidar data. What is the best spatial scale is also a question. Further work will be done for answering these questions.

REFERENCES

Harding, D. J., Carabajal, C. C, 2005. ICESat waveform measurements of within footprint topographic relief and vegetation vertical structure. *Geophysical Research Letters*, 32. L21S10.doi: 10/1029/2005GL023471

Lefsky, M. A., Harding, D. J., Keller, M., Cogen, W. B., Carabajal, C. C, Espirito-Santo, F. D., et al. 2005. Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters*. 32(22).

Josef Kellndorfer, Wayne Walker, Leland Pierce, Craig Dobson, Jo Ann Fites, Carolyn Hunsaker, John Vona, Michael Clutter, et al. 2004. Vegetation height estimation from Shuttle Radar Topography Mission and National Elevation Datasets. *Remote Sensing of Environment*, 93(3), pp.338-358.

Ranson, K. J., Sun, G., Kovacs, K., Kharuk, V. I. 2004a. Landcover attributes from ICESat GLAs data in central Siberia. I*GRASS 2004 Proceedings.* 20–24September 2004, Anchorage, Alaska, USA.

Ranson, K. J., Sun, G., Kovacs, K., & Kharuk, V. I. 2004b. Use of ICESat GLAS data for forest disturbance studies in central Siberia. *IGARSS 2004 Proceedings*, 20–24September 2004, Anchorage, Alaska, USA.

Sarabandi, K., Lin Y. C. 2000. Simulation of interferometric SAR response for characterizing the scattering phase center statistics of forest canopies. *IEEE Transactions on Geoscience and Remote Sensing*, 38(1), pp.115-125.

Sun, G., Ranson, K.J., Masek, J., Z. Guo, Y. Pang, A. Fu, D. Wang et al. 2008a. Estimation of tree height and forest biomass from GLAS data. *Journal of Forest Planning*, 13, pp.157-164.

Sun, G., Ranson, K. J., Kimes, D. S., Blair, J. B., Kovacs, K. 2008b. Forest vertical structure from GLAS: An evaluation using LVIS and SRTM data. *Remote Sensing of Environment*, 112(15), pp.107-117

United States Geological Survey. 2003. Shuttle Radar Topography Mission documentation: SRTM Topo. http://decftp.cr.usgs.gov/pub/data/srtm/Doc umentation/SRTM_Topo.tex

Zwally, H. J., Schutz, B., Abdalati, W., Abshire, J., Bentley, C., Brenner, A., et al. 2002. ICESat's laser measurements of polar ice, atmosphere, ocean, and land. *Journal of Geodynamics,* 34(3-4), pp.405-445

ACKNOWLEDGEMENTS