# EXTRAPOLATION OF LIDAR FOR FOREST STRUCTURE ESTIMATION USING SAR, IFSAR, AND OPTICAL DATA

*Leland Pierce, Michael Benson, Kamal Sarabandi*

The University of Michigan,
EECS Dept, Radiation Lab,
1301 Beal Ave, Ann Arbor, MI
lep@umich.edu

## 1. INTRODUCTION

One of the most fundamental new technical challenges of a DESDynI spaceborne mission is the fusion of the several sensor modalities - LiDAR, SAR, InSAR, and Optical - in order to accurately estimate desired Vegetation 3D and biomass parameters at their point of intersection and to extrapolate them over continuous areas.

The objective of this paper is to use both our simulation models and measured dataset to develop and validate fusion and extrapolation methods while simulating DESDynI-type missions.

We use existing datasets to develop and validate our fusion and extrapolation approach, which involves using our four sensor simulators, including our fractal-based tree geometry generator, in tandem with our in-house simulated-annealing-based parameter estimation software which performs fusion and retrieval functions. We use existing field and radar-lidar-VNIR data for the north and south Boreas sites, as well as simulated data.

## 2. TECHNICAL APPROACH AND METHODOLOGY

### 2.1. Vegetation 3D Simulation Models

Over the past years we have developed several forward models for remote sensing applications [Ulaby, et al., 1990; Sarabandi and Lin, 2000; Lin and Sarabandi, 1999b]. Recently we have focused on enhancing our capabilities for description of image scenes (the input data) through development of our fractal tree software, and incorporation of these into the forward models themselves. All of our forward models use as their input a description of a particular forested area from the fractal tree software, including every branch, needle, and leaf, its position, orientation, length, diameter, thickness, moisture, and bulk density. Each of the simulators has been developed using the same 3D forest structure input data and validated against measured data [Sarabandi and Lin, 2000; Lin and Sarabandi, 1999b].

### 2.2. Data Fusion

Several recent studies have focused on assessing accuracies of forest structure estimates using two or more sensors including LIDAR, radar, and Visible-Infrared (VIR) [Hyde, et al., 2007; Walker, et al., 2007; Hyde, et al., 2006]. These studies, based primarily on empirical analysis, and have produced rms error assessments and interpretations of sensor utilities, and have also suggested development of physical-based models as a needed advancement. One of the first attempts at fusion of four modalities (VIR, SAR/INSAR, and LIDAR) [Pierce, et al., 2002] used only several features: height from LIDAR, vegetation community from VIR, INSAR heights, and SAR powers. This study used the empirical Bayesian approach and demonstrated that using this approach, multiple features quickly resulted in poorly-estimated multi-dimensional density functions. In a step towards physical or model-based methods, Moghaddam et al. [2002] used AirSAR and Landsat TM to estimate foliage mass. Their method is a combination of Bayesian and model-based, with training areas used to develop forward models instead of probability density functions, yielding an improvement in the error over VIR-alone from about 30% to 15%. Recently Kimes et al. [2006] studied fusion of lidar with multi-angle and used an optical model.

Extrapolation of LIDAR heights using VIR has been used in Hudak [2002], where an empirical relationship between the two was used with kriging and cokriging. The study concluded that the spacing of the LIDAR data needed to be 250 meters or less for an accurate extrapolation. SRTM (INSAR), Landsat (tasseled-cap), and a canopy density layer were used to extrapolate LIDAR heights [Kellndorfer, et al., 2006], using non-spatial regressions, resulting in rms errors of 3 meters. Another study showed that radar and multispectral data could be used to extend LIDAR samples [Hyde, et al., 2006].

### 2.3. Boreas Test Site

We are using the Boreas test site in Canada because it has a wealth of ground truth data, including the necessary forest cover and structure measurements needed for this study, as well as a large set of SAR, IFSAR, and large-footprint LiDAR data that is separted by less than one year in time. Given the slow growth at this site, this small separation in time is not expected to impact the research effort. The site encompasses different forest types, including various types of Pine, mixed Conifer, and Aspen. There are also a variety of stand ages, and hence biomasses and heights for each of these forest types. The completeness and near-simultaneity of this dataset makes it ideal for this study.

## 3. RESULTS

### 3.1. Simulation-based Fusion Study

To study the capabilities of data fusion for determining forest structure, we devised a simulation study using uniform plantations of sugar pine trees. These were generated using the fractal tree simulator and made into realistic stands which were then fed into each of the four simulators.

The generated dataset consists of sugar pine stands with biomasses ranging from near zero to 2000 tons/ha, and heights ranging from 5 meters to 25 meters. The remaining forest parameters were either kept constant or varied in a mechanistic way based on the height, using allometric equations.

The process of estimating the biomass and height uses the simulator to produce measured data using a guess that consists of the biomass and height alone. This guess is used to generate a forest stand, which is used by the simulators to produce the measured data (SAR, LIDAR, etc.). Various measures of the data are used together to determine an error measure as compared with the dataset under study. When this measure is close enough, the biomass/height used for that guess is the estimate as determined for that datapoint by the process. This is diagrammed in Fig. **??**. It is very important in this process that we are only trying to estimate 2 unknowns, rather than the many that would be needed if the geometry were represented using standard statistical descriptions of forests.

After applying this algorithm for hundreds of simulated forests, the resulting RMS error in estimation of height was 1.5 meters, while for biomass the RMS error was 226 tons/ha. In both cases the error is approximately 10% of full-scale, which is quite good. The results are plotted in Fig. **??**. We expect that in areas where the range of heights and biomasses are known to within some smaller range of values this algorithm can provide high accuracy estimates of both biomass and canopy height.

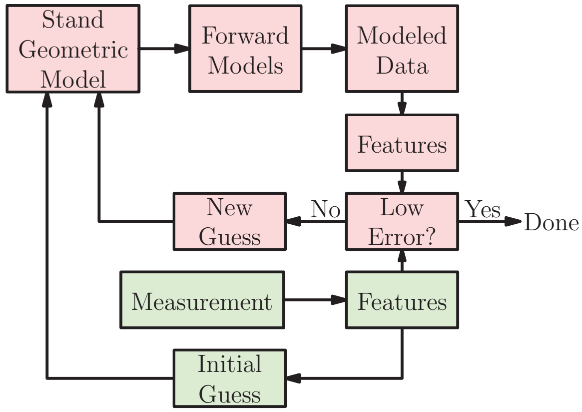
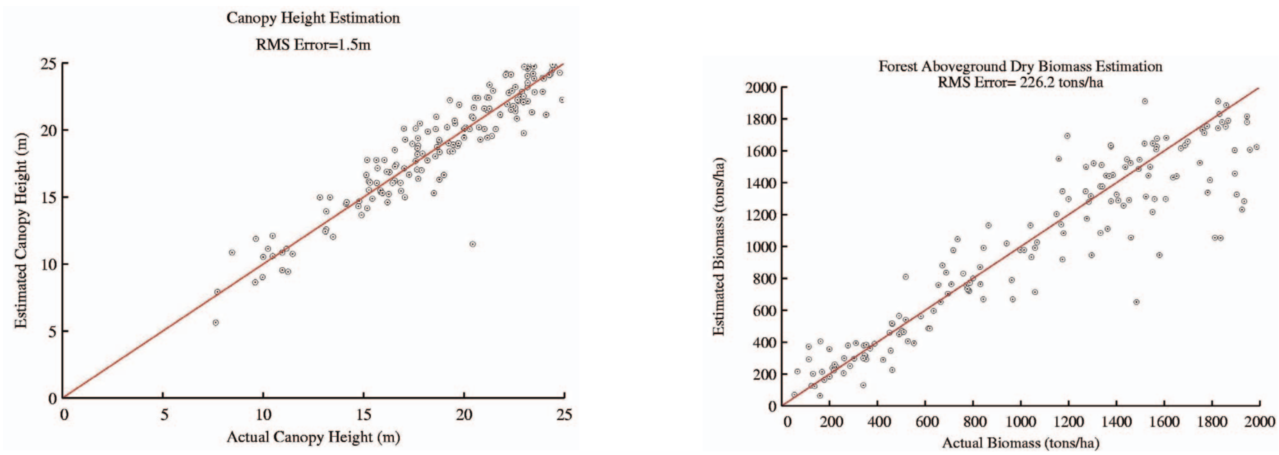



**Fig. 1**. Block diagram of data fusion process for estimation of forest biomass and height.




(a) Height Estimation Results

(b) Biomass Estimation Results

**Fig. 2**. Estimation of forest height and biomass. Results for the entire range of biomasses and heights show a 10% RMS error for each.

### 3.2. Simulation-based Extrapolation

Using the same simulated dataset as previously, but limiting the biomass to 800 tons/ha, we repeated the previous analysis to set a baseline performance measure when using all available data. This resulted in RMS errors of 1.7m for height and 184 tons/ha for biomass. The two questions we then wanted to answer were: (1) How did these errors change when LIDAR was not available? (2) Could we use a subset of the LIDAR data to devise a regression equation to estimate the LIDAR data from the rest, and use that to estimate height and biomass more accurately?

The estimates were done using all the data except the LIDAR data and the resulting RMS errors were: 3.3m for height and 380 tons/ha for biomass. This is a significant degredation from the previous result, as expected.

Next, we used a subset of the LIDAR data points, to simulate that we had only a few areas with LIDAR data that overlapped SAR data, as expected for DesDynI. The regression equations used the IFSAR scattering phase center heights (Lhv-Chh), NDVI, and SAR power. With these regression equations the estimates resulted in RMS errors of: 2.7m for height, and 268 tons/ha for biomass. Both of these show 20 to 30% improvement over the no-LIDAR case, with the resulting height error close to 10% full-scale, and the biomass near 30% full-scale, only 7% worse than when using all the data available.

### 3.3. Application to Boreas dataset

The next step is to apply this estimation process to a real dataset, in this case data from the Boreas study. We are in the midst of processing the data in order to obtain the data in a usable form for the estimation procedure as described earlier. This involves extracting the ground truth for each forest stand, and the corresponding SAR and LIDAR measurements. Once tabulated, this data will be used to determine the usefulness of the previously-described techniques on actual measurements.

## 4. REFERENCES


[1] Hudak, Andrew T., Lefsky, Michael A., Cohen,Warren B., Berterretche, Mercedes, "Integration of lidar and Landsat ETM+ data for estimating and mapping forest canopy height," *Remote Sensing of Environment,* Vol. 82, pp. 397–416, 2002.

[2] Hyde, P., Dubayah, R., Walker, W., Blair, J.B., Hofton, M., Hunsaker, C., "Mapping forest structure for wildlife habitat analysis using multi-sensor (LIDAR, SAR/INSAR, ETM+, Quickbird) synergy," *Remote Sensing of Environment,* Vol. 102, pp. 63–73, 2006.

[3] Hyde, P., R. Nelson, D. Kimes, E. Levine, "Exploring LiDAR-RaDAR synergy – Predicting aboveground biomass in a southwestern ponderosa pine forest using LiDAR, SAR, and InSAR," *Rem. Sensing of Environment,* pp. 28–38, Vol. 106, 2007.

[4] Kellndorfer, J. M., Walker, W., Bishop, J., Lapoint, E., Hoppus, M., Westfall, J., "InSAR/Lidar/Optical Data Fusion for Vegetation Height and Biomass Estimation in Support of the North American Carbon Program," *American Geophysical Union,* Fall Meeting 2006, abstract #B41A-0164, Dec. 2006.

[5] Kimes, D.S., K.J. Ranson, G. Sun and J.B. Blair, "Predicting lidar measured forest vertical structure from multi-angle spectral data," *Remote Sensing of Environment,* Vol. 100, No. 4, 503-511, 2006.

[6] Lin, Y.C., and K. Sarabandi, "A Monte Carlo Coherent Scattering Model for Forest Canopies Using Fractal-Generated Trees," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 37, no. 1, pp. 440–451, January 1999a.

[7] Lin, Y.C., and K. Sarabandi, "Retrieval of Forest Parameters Using a Fractal-Based Coherent Scattering Model and a Genetic Algorithm," *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 37, No. 3, pp. 1415–1424, May 1999b.

[8] Moghaddam, M., J.L. Dungan, S. Acker, "Forest variable estimation from fusion of SAR and multispectral Optical data," *IEEE Trans. Geosci. Rem. Sensing,* pp. 2176–2187. Vol. 40, No. 10, Oct. 2002.

[9] Pierce, L.E., W.S. Walker, M.C. Dobson, C.T. Hunsaker, J. Fites-Kaufman, R. Dubayah, "Fusion of optical and SAR data for forestry applications in the Sierra Nevada of California," IGARSS'02, pp. 1771–1773, Vol. 3, 2002.

[10] Sarabandi, K., and Y.C. Lin, "Simulation of Interferometric SAR Response for Characterizing the Scattering Phase Center Statistics of Forest Canopies," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 38, no. 1, pp. 115–125, January 2000.

[11] Ulaby, F.T, Sarabandi, K., McDonald, K., Whitt, M., Dobson, M.C., "Michigan microwave canopy scattering model," *Intl. J. Remote Sensing,* Vol. 11, No. 7, pp. 1223–1253, July 1990.

[12] Walker, W.S., J.M. Kellndorfer, E. LaPoint, M. Hoppus, J. Westfall, "An empirical InSAR-optical fusion approach to mapping vegetation canopy height," *Rem. Sensing of Environment,* pp. 482–499, Vol. 109, 2007.