Producing a canopy height map over a large region using heterogeneous lidar datasets

Ranjith Gopalakrishnan, Dr. Valerie Thomas , Dr. John Coulston and Dr. Randolph Wynne

VirginiaTech

Motivation: Why care about a tree height map?

• Estimation of biomass, for carbon accounting, fuel load (eg: annual federal spending on wildfires increased from $1.4 billion (1990s) to $3.5 billion (2000s), adjusted for inflation.
•Currently, there is no good strategy to get forest parameter estimates over large areas using a combination of airborne LIDAR and (ground based) national forest inventory data. This work is a beginning in that direction.

Previous work

1. Over a large area: There has been attempts to estimate tree heights over large areas using space-borne LIDAR sensors. The RMSE involved in these efforts varied from 5.0 to 12.5 meters (Lefsky et al 2005, Simard et al 2011).
2. For smaller areas: Most attempts are over similar ecozones, and use the same airborne sensor. Relative RMSE of ~10% have been reported (Yu et al 2011).

Methods

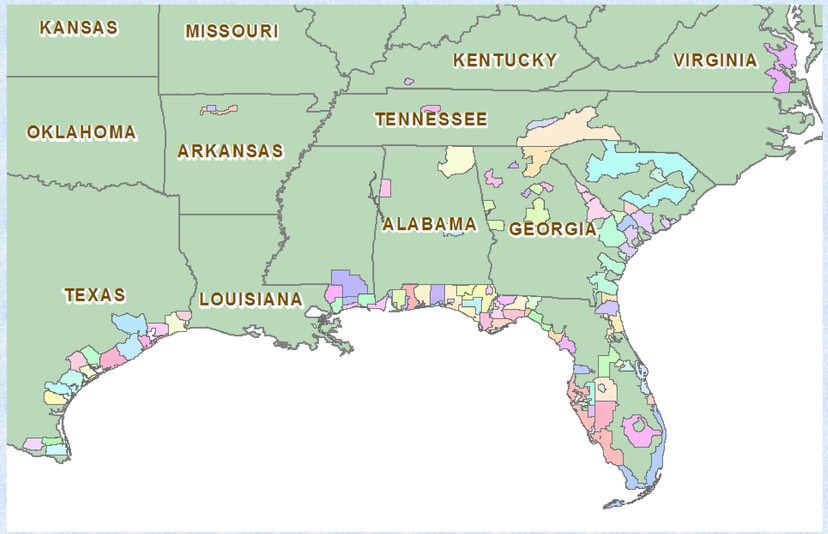


Fig 1: LIDAR coverage, different projects are shown as different colors.

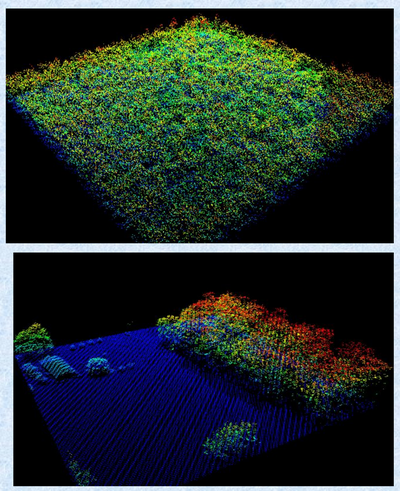


Fig 2: Homogeneous (top) and non-homogeneous lidar plots

- First, we acquired LIDAR data from USGS, NOAA and NRCS (AL) for a large area (fig 1), with LIDAR acquisition dates ranging from 2005 to 2011. Over 90 separate LIDAR projects are involved.
- Then, we intersected this data with FIA plot location data. This gave us 2084 plots (120 x 120 m) where we had both LIDAR and FIA field data (fig 2).
- Then, simple linear regression models between the LIDAR parameters and FIA field measurements of tree heights, were made. The factors affecting the residual of the fit were then examined (next section).

Results and conclusions

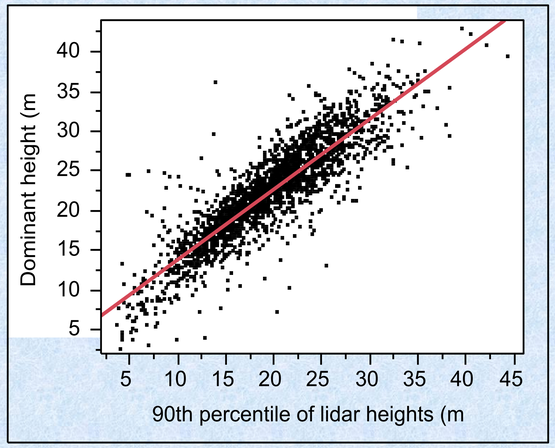


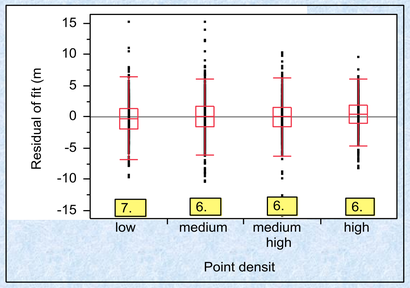
Fig 3: Simple linear regression fit of a lidar-based height metric, and the FIA-measured height, for n = 2084 plots. R_{adj}^2 = 0.78, RMSE = 2.94 meters.

PD more than…	Homogeneity more than… 2.0 (least homog.)	2.50	3.33	5.00	6.67	10.0 (most homog.)
0.20	2.94	2.84	2.63	2.56	2.16	**1.46**
1.00	2.87	2.75	2.56	2.36	2.11	1.47
2.00	2.81	2.66	2.41	2.21	2.21	NA
4.00	2.56	2.52	2.44	2.06	NA	NA
6.00	2.93	2.99	2.86	2.23	NA	NA

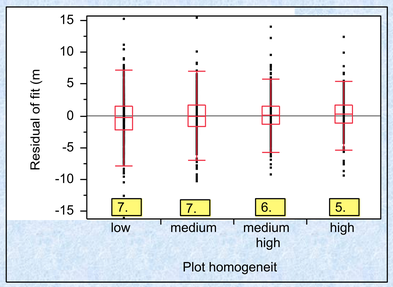
Table 1: The effect on RMSE (in meters) of thresholding homogeneity and PD (point density).

Scatter: factors investigated

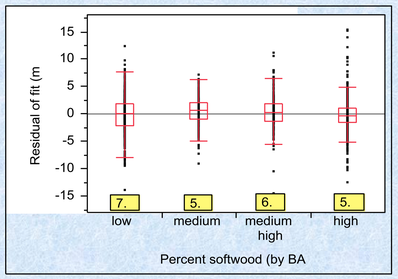
- Point density (num. lidar pulses/m^2)
- Plot homogeneity (surrogate for co-registration errors). Quantified by 1/CV (coefficient of variation of lidar pulse heights)
- % trees that are softwoods in the FIA plots (vs hardwoods)



IQR (interquartile range) = 1.0



IQR = 1.9



IQR = 2.5

Fig 4: The effect of various factors, explored by boxplots. The values at the bottom (yellow box) are the quantile ranges, between the 10th and the 90th quantiles.

Conclusion

- The RMSEs of our effort (~ 2.94 m) are lesser than that of other similar efforts, for large areas .
- Homogeneity of the plot (surrogate for co-registration accuracies), and species grouping (hardwoods versus softwoods) were found to be important factors for accuracy. Point density was relatively less important.
- We selected a set of plots informed by our findings (homogeneity >= 10.0). These “better quality” plots had a significantly improved RMSE of 1.46 meters.