James A. Moore Keynote Address

Biometrics in the era of remotely sensed forest inventory: Opportunity and Risk

INTERMOUNTAIN FORESTRY COOPERATIVE
Tuesday, March 23, 2021

John Paul McTague
Southern Cross Biometrics

Advertisement in the SAF *E-Forester*

Did the keynote speaker misplace the correct presentation?

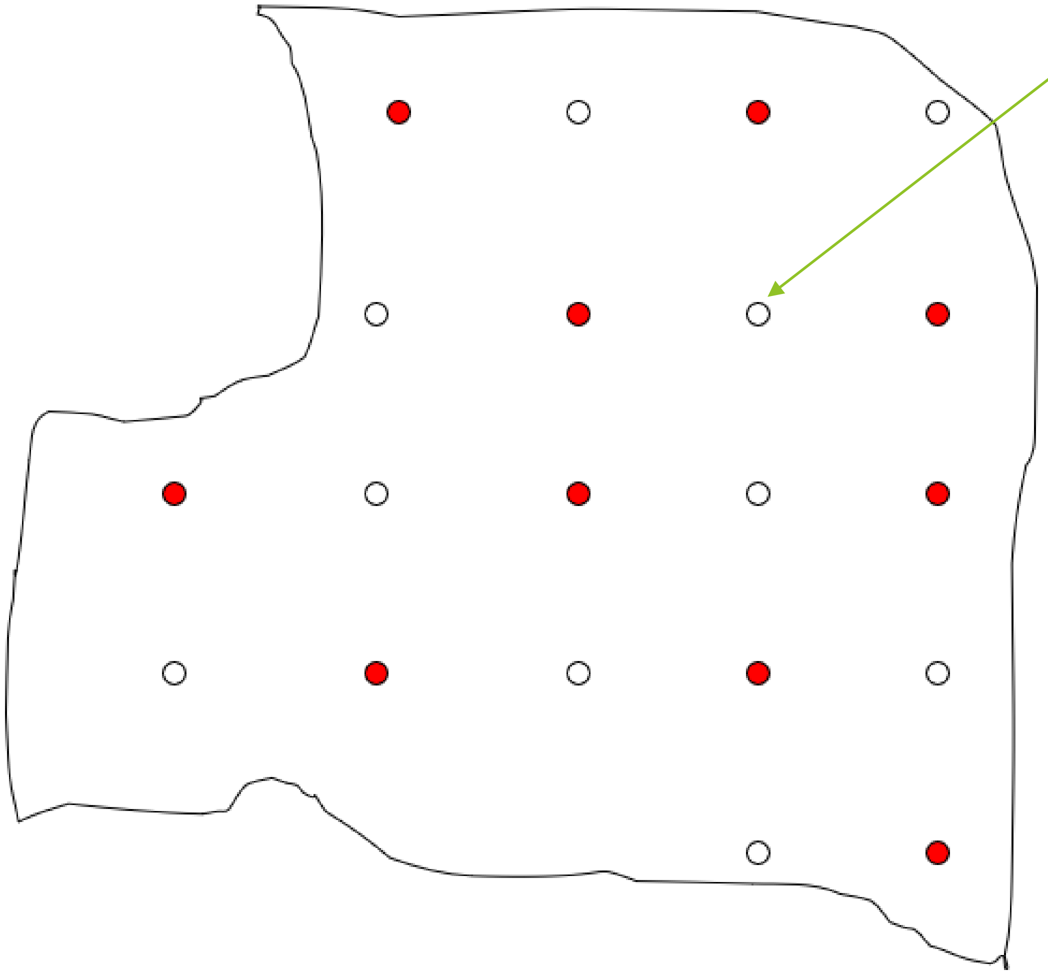
or

Is he manifesting a 'senior' moment?

Emphasize the concept of 'auxiliary' information. Prof. Oderwald is proposing to add more auxiliary information.

Volume point (**measure** dbhs and heights of "in" trees)

○ Basal area point (sweep 360° and **count** "in" trees)



Oderwald wants to augment the number of these plots

Large sample (n_1) = number of **volume** points + number of **count** points

small sample (n_2) = number of **volume** or **measure** points

Double sampling

with regression

small sample slope of V-B

Large sample basal area

$$\bar{y}_{Rd} = \bar{y}_2 + b\left(\bar{x}_1 - \bar{x}_2\right)$$

small sample volume

small sample basal area

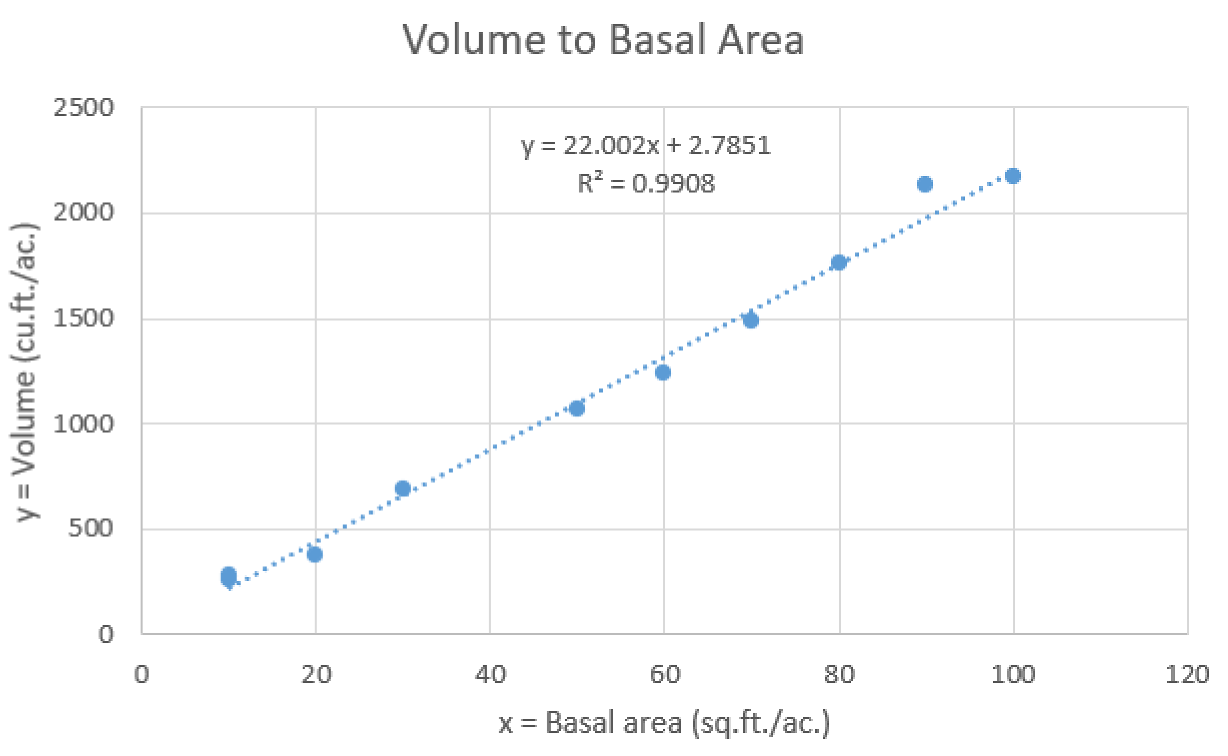
x = auxiliary information (variable)

y = principal or target information (variable)

Auxiliary information should be highly correlated to y while easy and inexpensive to acquire. In our example, complete enumeration of x_i is still too expensive, hence $\bar{X}$ is not known. The variance is expressed as:

$$v(\bar{y}_{rd}) = \frac{MSE}{n_2} + \frac{b^2 S_x^2}{n_1}$$

b is the slope coefficient or 22.002 in this example



Goal: Select a good auxiliary variable, which leads to a high R^2 and low MSE

If $\bar{X}$ could be acquired, say from remote sensing techniques, then ratio or regression sampling should be used. In other words, '**Wall-to-Wall**' auxiliary information.

Variance now becomes: $v(\bar{y}_r) = \frac{MSE}{n_2}$

Is remote sensing mensuration any different?

Terminology is different

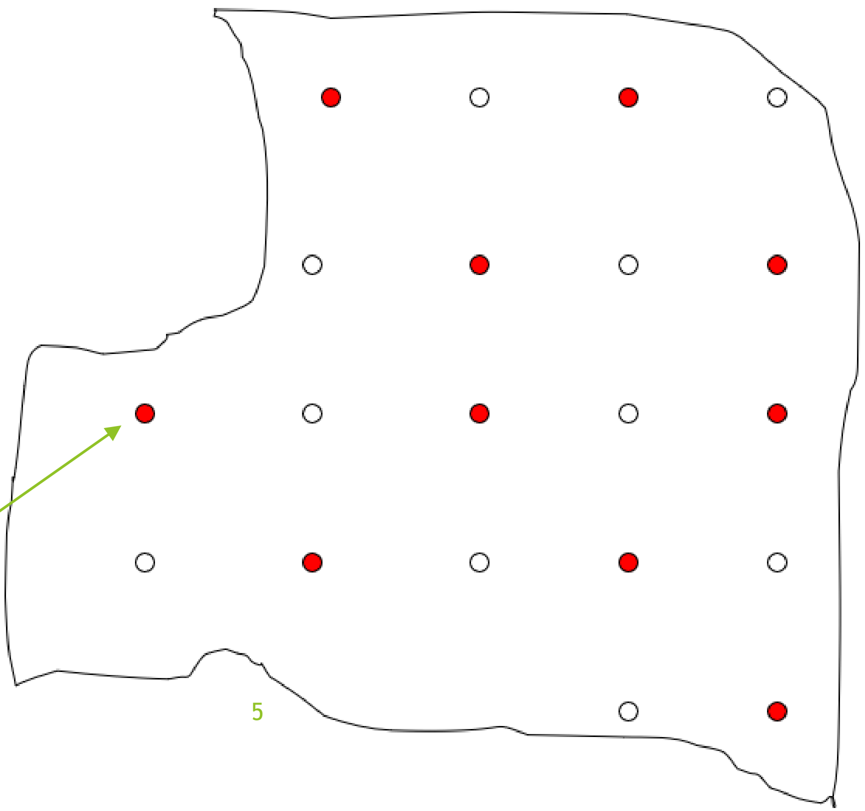
Design-based inference through model-assisted estimation

Models can be used to improve estimators under the design-based framework. An important category of such estimators are known as model-assisted estimators (Särndal et al. 1992). The general form of such estimators, for estimating a population total, is

$$\hat{\tau}_{ma} = \sum_{i\in U}\hat{y}_i + \sum_{i\in s}\frac{(y_i-\hat{y}_i)}{\pi_i}$$

The 'wall-to-wall' remote sensing plots constitute auxiliary information. (we obviously meet Oderwald's plea for more auxiliary plots with wall-to-wall coverage)

We still need a sample (probability-based) design to measure the 2nd-phase volume plots of the target population. Yep, that's right …. Boots on the ground! We also obtain the diameter distribution and species mix.



Good examples to emulate

Remote Sensing of Environment 117 (2012) 348–357

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

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

ELSEVIER

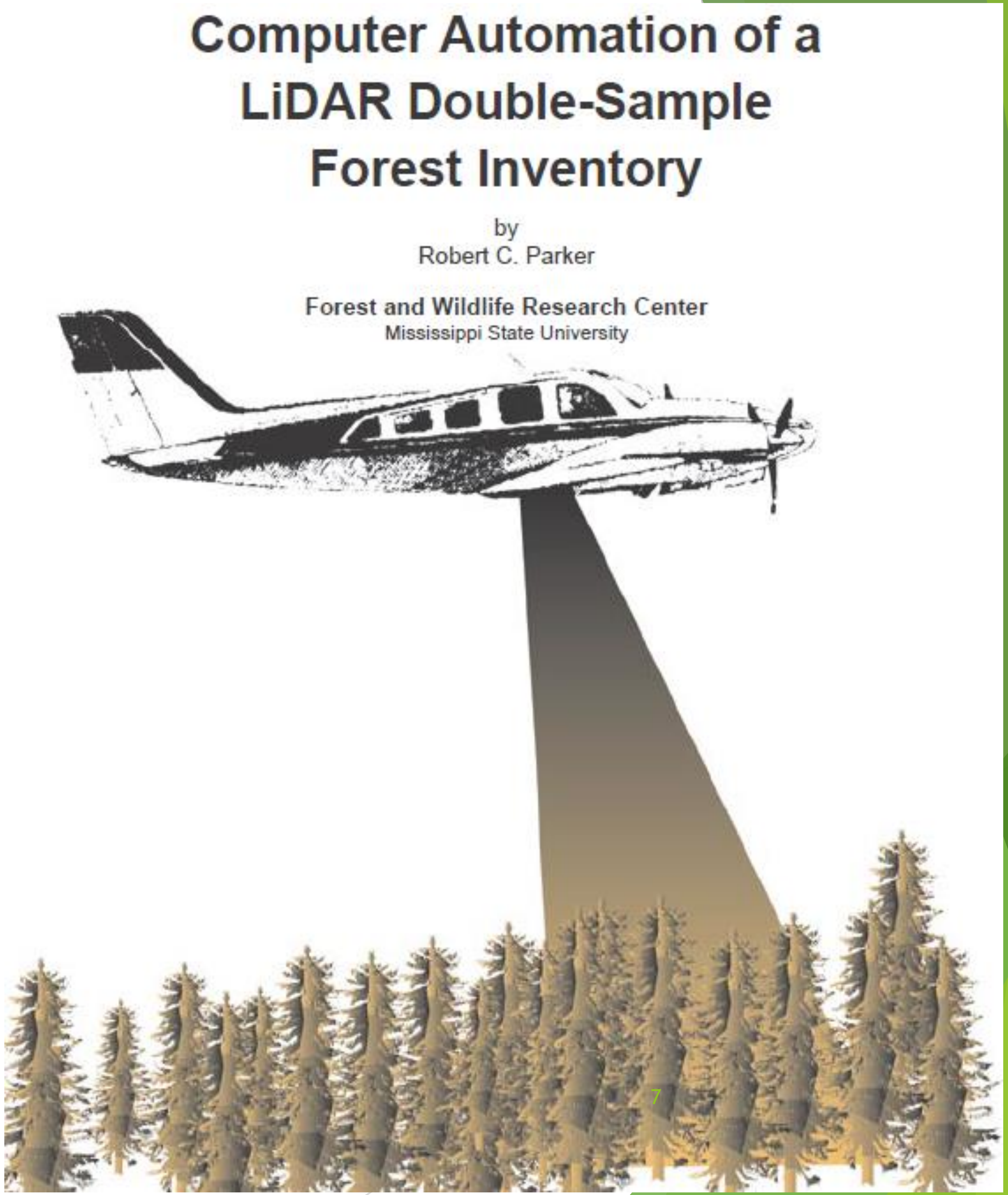
Airborne scanning LiDAR in a double sampling forest carbon inventory

Peter R. Stephens [a,1], Mark O. Kimberley [b], Peter N. Beets [b,*], Thomas S.H. Paul [b], Nigel Searles [a], Alan Bell [c], Cris Brack [d], James Broadley [d]

Model-Assisted Forest Yield Estimation with Light Detection and Ranging

Jacob L. Strunk, Stephen E. Reutebuch, Hans-Erik Andersen, Peter J. Gould, and Robert J. McGaughey

Early work by Prof. Parker. The computer automation contains a component on sampling computations



One step forward, one step back

- Remote sensing 'interpretation' is glamorous. It is a dynamic field, with aircraft, satellites, drones, sensors, filters, algorithms, and energetic scientists. Moving from passive imagery to active imagery (LiDAR) only heightens the mystique.

- Survey sampling is boring. Forest biometricians have a mastery of sampling designs and techniques, but no soul ever got promoted in the private sector by exhibiting their skills in sampling. Biometricians frequently get involved in the last step or damage control of a project.

- Problem: Glamorous interpretation and information derived from remotely sensed images is still auxiliary information, in essence, no different than 'basal area plots'.

- In the rush to display fast turnaround and inexpensive results, the sampling component of remote sensing mensuration is often cast aside, with almost always, poor or inconclusive results.

Opposing views about the role of auxiliary information: Unmanned aerial systems with laser scanning sensors

Remote Sens. **2015**, *7*, 9632-9654; doi:10.3390/rs70809632

OPEN ACCESS

remote sensing

ISSN 2072-4292

www.mdpi.com/journal/remotesensing

Article

Inventory of Small Forest Areas Using an Unmanned Aerial System

Stefano Puliti *, Hans Ole Ørka, Terje Gobakken and Erik Næsset

Most of the published studies have focused on using UAS-LS to map forest biophysical properties from flying above the tree canopy. <u>The main objectives of the studies have been to determine the feasibility of sampling forests remotely without the need of field measurements</u>, ….

Opposing view from a major international forest research organization

Problem Statement

- Use LiDAR derived metrics in a multiple regression modeling approach to determine the relationship with stand parameters such as DBH, Basal Area, etc.
- This approach has been widely used in forestry applications.
- The weakness of this approach is that it relies on the correlations between the LiDAR metrics and the stand parameters. These correlations change in different stand conditions and in different years. The LiDAR metrics correlated at one time in an area are likely different at another time in the same area and in different areas. So the correlations need to be redone for each locate and time.

Opposing view:

The proposed solution

- Use traditional biometrics approach. Measure height and diameter in a series of plots and determine the relationship between Training and dbh-ht eq.
- Use LiDAR to determine height of each tree and then predict DBH based on the calculated relationship between height and DBH. Every tree can be identified
- The calculate Basal Area or Volume of each tree using standard equations.
- Sum individual trees to give stand level data. The inventory is then completed

Opposing view about role of auxiliary information … continued

- In the previous slide, the series of plots used to construct the dbh-ht relationship originated from research trials in the Northern and Southern hemisphere
- The equation form is simple: dbh = $a_0 + a_1$ht
- Stand density, age, and location were not used as predictor variables
- To my knowledge, this proposed solution has not been implemented.

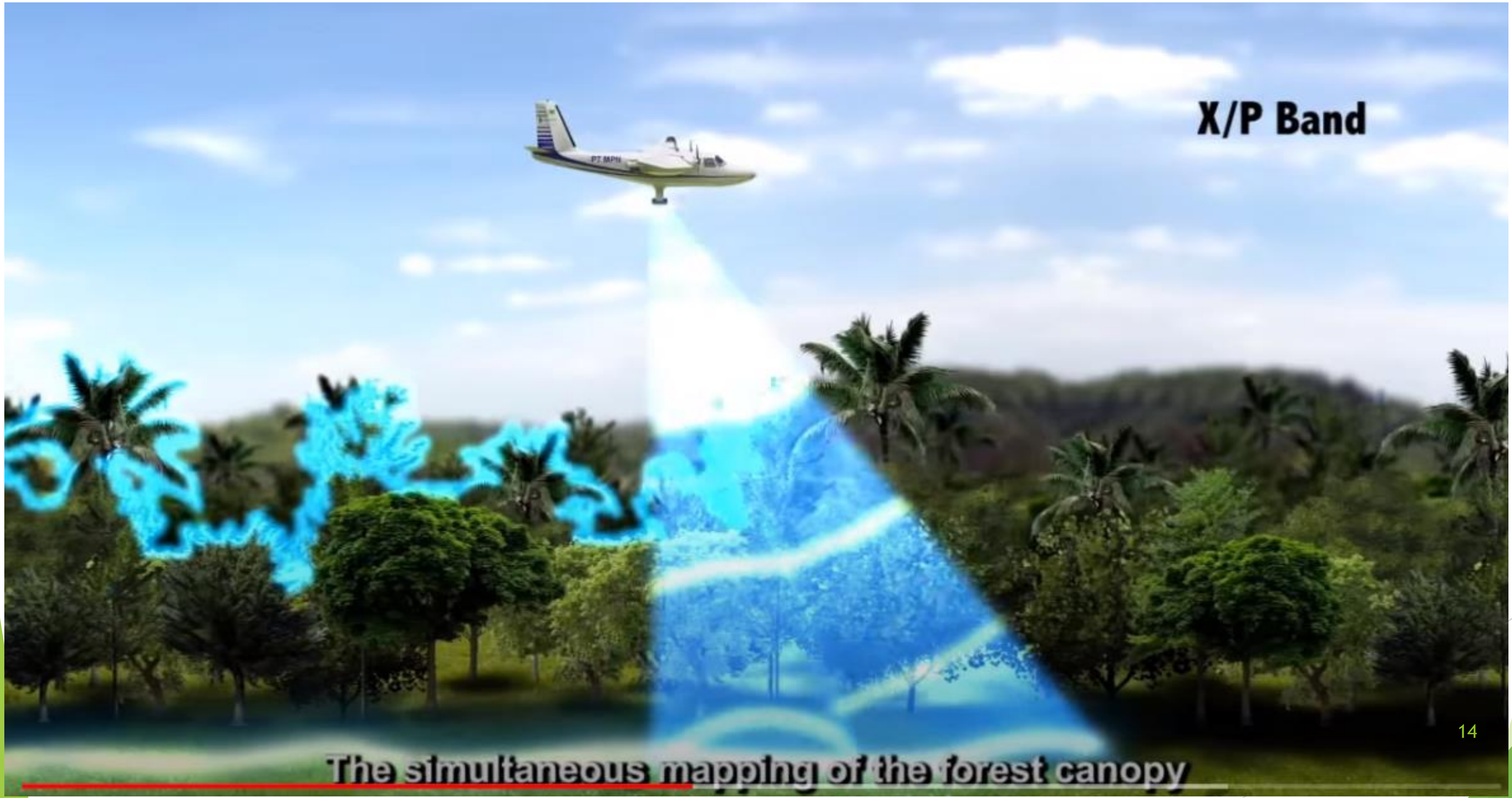
> The weakness of this approach is that it relies on the correlations between the LiDAR metrics and the stand parameters. These correlations change in different stand conditions and in different years.

- In my mind, a stand-specific inventory is not a weakness.

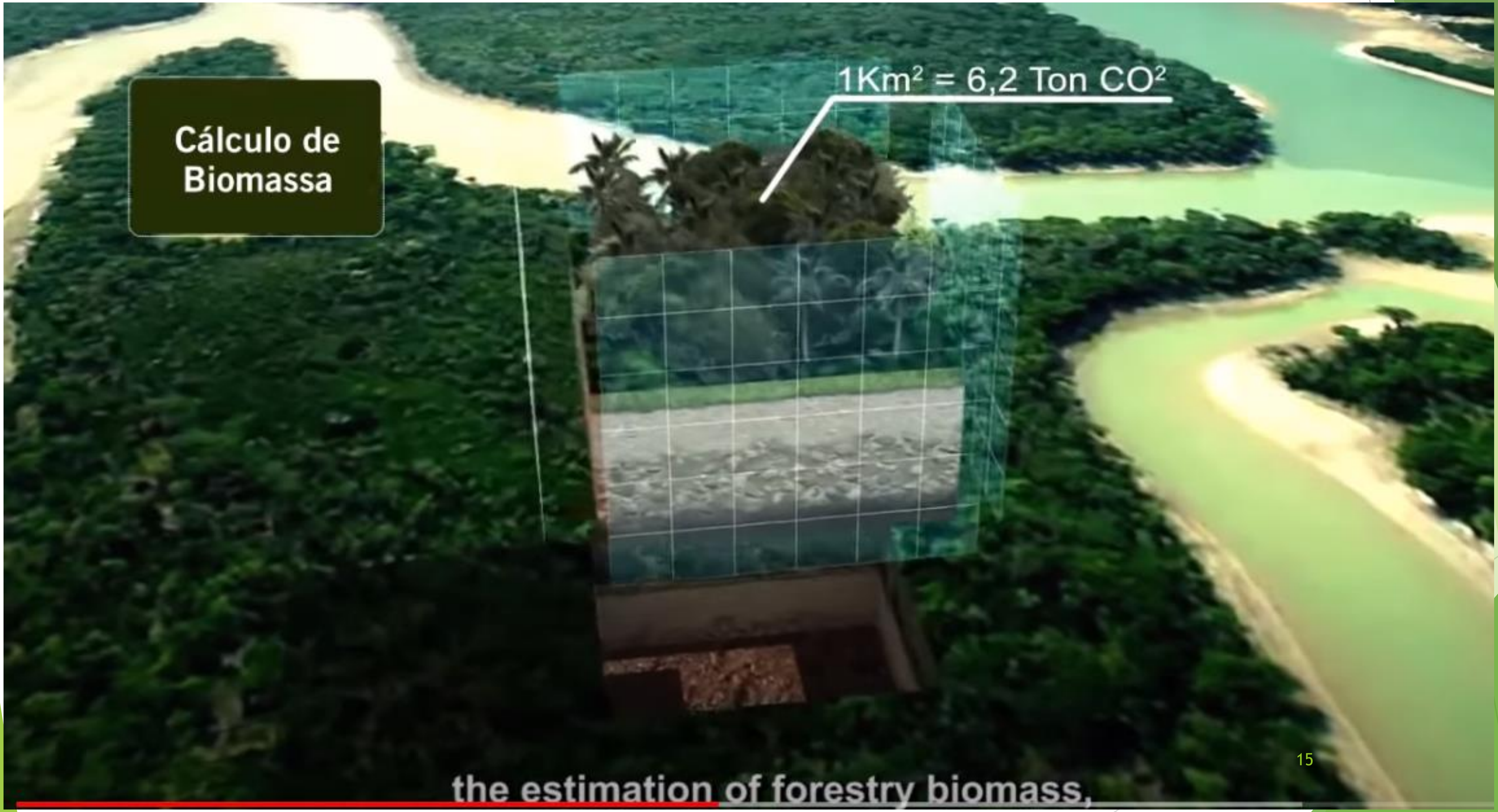
The perils of creative marketing

- Marketing blurbs touted advanced forest management practices with the use of remote sensing mensuration
- In addition, savings were claimed by removing the need of expensive field measurements
- The website claims of the elimination of ‘boots on the ground’ were removed in a matter of months. I assume this occurred after a land sale or appraisal.

Another solution for eliminating difficult field measurements in tropical forests
(in this case with radar)



Bradar no longer has an active website



Speaking of carbon

- What are the dangers of using only auxiliary information from remote sensing sources to estimate carbon sequestration or growth of above ground biomass (AGB)?

- Most likely, this is not a problem.

- Dr. Paul Van Deusen (1997) demonstrated that additive biases at time 1 and time 2 cancel out when estimating change. This of course assumes that no substantive changes were made to the sensors from time 1 to time 2.

- Carbon stock estimates at time 1 and carbon stock at time 2 will, most likely, continue to possess a bias

Goodness of fit for the prediction models

Mean top height

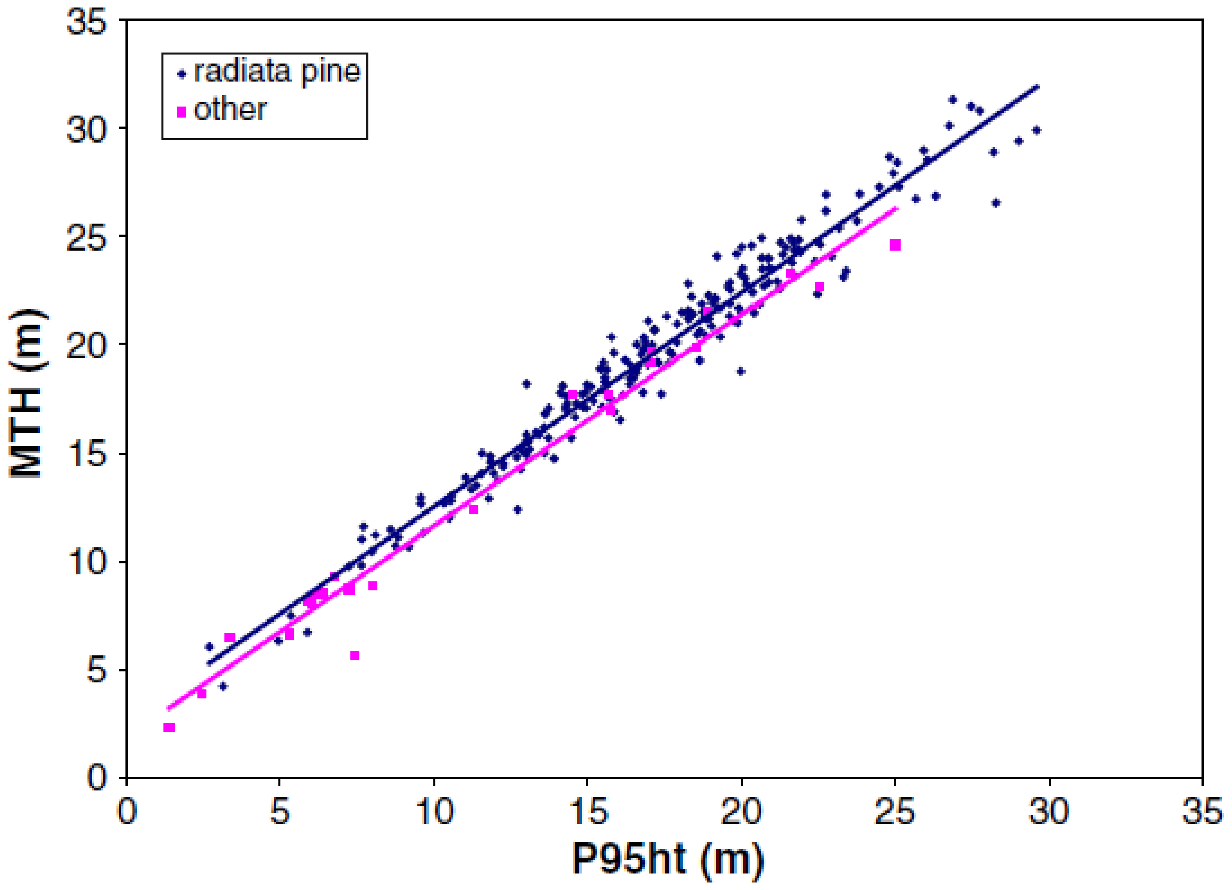


Fig. 2. MTH versus LiDAR P95ht showing sites planted in radiata pine, and sites planted in other species. The overall regression model is y = 2.31 + 1.004 × P95ht; $R^2 = 0.96$; root mean square error, RMSE = 1.09 m.

Plethora of candidate predictor variables tested and used

Metric	Description
MEAN	Mean height (m)
STD_DEV	Standard Deviation
ABS_DEV	Absolute Standard Deviation
SKEW	Skewness
KURTOSIS	Kurtosis
MIN	Minimum height (m)
P01	First percentile LiDAR Height (m)
P05	Fifth percentile LiDAR Height (m)
P25	Twenty-fifth percentile LiDAR Height (m)
P75	Seventy-fifth percentile LiDAR Height (m)
P95	Ninety-fifth percentile LiDAR Height (m)
P99	Ninety-ninth percentile LiDAR Height (m)
P10	First Decile LiDAR Height (m)
P20	Second Decile LiDAR Height (m)
	↓
P80	Eighth Decile LiDAR Height (m)
P90	Ninth Decile LiDAR Height (m)
MAX	Maximum height (m)
D1	Cumulative percentage of the number of returns found in Bin 1 of 10
D2	Cumulative percentage of the number of returns found in Bin 2 of 10
	↓
D8	Cumulative percentage of the number of returns found in Bin 8 of 10
D9	Cumulative percentage of the number of returns found in Bin 9 of 10
DA	First returns/ All Returns
DV	First Vegetation Returns/All Returns
DB	First and only return / All Returns
VDR	Vertical Distribution Ratio = [Max–Median]/Max
Canopy_relief_ratio	(Mean - Min) / (Max – Min) – Using All Returns
Covar	Std Dev (all Returns)/ mean ((All Returns)
CanCovar	Std Dev (First Returns Only)/ mean ((First Returns Only) [represents the canopy layer]
FIRST	Number of First Returns
ALLRETURNS	Number of all Returns
FIRSTVEG	Number of First Vegetation Returns only
ALLGROUND	Number of Ground Returns
cc2	Crown closure ≥ 2 m
Cc4	Crown closure ≥ 4 m
	↓
cc24	Crown closure ≥ 24 m
cc26	Crown closure ≥ 26 m
STRAT1	Percentage of Total Return Count Below 0m
STRAT2	Percentage of Total Return Count Between 0m and 2m
	↓
STRAT14	Percentage of Total Return Count Between 24m and 26m
STRAT15	Percentage of Total Return Count Above 26m

New Nouveau Brunswick

Height

Run 1	Run 2	Run 3
P95	P95	cc12
cc12	cc12	P95
STRAT7	P80	cc18
P80	STRAT7	P99
P99	P99	STRAT7
STRAT8	cc18	STRAT4
STRAT4	STRAT4	STRAT6
DA	STRAT8	STDDEV
P90	STDDEV	P90
cc16	DA	STRAT5
STDDEV	P90	STRAT8
cc18	STRAT5	P80
cc14	AAD	cc16
STRAT5	STRAT6	cc14
DB	DB	DA
STRAT6	cc16	AAD
cc4	cc14	P75
AAD	P75	DB
P75	cc6	cc6
P70	cc10	cc10

Results in the tropics (Acre, Brazil)

Lidar-based above ground biomass, volume, and basal area regression models developed using the FEA ground plots ($n = 50$). Square root transformation of response variables and back transformed equation with bias correction factors (CF) are included for biomass and volume regressions.

Forest structure variable	Regression model	CF	Multiple R^2	Root-mean-square error
Sqrt(Biomass) (Mg ha^{-1})	$\sqrt{AGB} = 3.119^{*} + 0.564^{***} \times P25 + 0.062^{***} \times VAR$	-	0.70	1.28
Biomass (Mg ha^{-1})	$AGB = (3.119 + 0.564 \times P25 + 0.062 \times VAR)^2 + CF$	1.74	0.72	40.20
Sqrt(Biomass) (Mg ha^{-1})	$\sqrt{AGB} = -1.583 + 0.796^{***} \times QElevMean$	-	0.71	1.37
Biomass (Mg ha^{-1})	$AGB = (-1.583 + 0.796 \times QElevMean)^2 + CF$	1.95	0.67	43.25
Sqrt(Biomass) (Mg ha^{-1})	$\sqrt{AGB} = -0.834 + 0.837^{***} \times ElevMean$	-	0.63	1.43
Biomass (Mg ha^{-1})	$AGB = (-0.834 + 0.837 \times ElevMean)^2 + CF$	2.12	0.63	46.12

Low pulse density (1.1 pulses/m^2) - Maine

Attributes	Key variables (mean square error)	R^2 (Adj R^2)	MB (SD)	RMSE
Maximum tree height (m)	Maximum height	0.869 (0.867)	1.89 (± 2.06)	2.80
Stem density (trees ha^{-1})	Fifth percentile height (3.302), Height kurtosis (5.982), Height L-skewness (6.198)	0.287 (0.280)	9 (± 5013)	4993
QMD (cm)	Percent of the first return above mean (6.5909) Percent of the first return above 1 m (7.8544) Twenty-fifth percentile height (8.3618)	0.489 (0.434)	−0.05 (± 3.69)	3.68
Basal area (m^2 ha^{-1})	Percent all returns above 1 m (7.262) Height L-kurtosis (7.564) Ninety-ninth percentile height (7.614)	0.344 (0.339)	0.03 (± 13.07)	13.01
Stem volume (m^3 ha^{-1})	Ninetieth percentile height (7.795) Twentieth percentile height (8.724) Seventy-fifth percentile height (9.757)	0.721 (0.719)	1.81 (± 66.96)	66.70

The quest for parsimony

LiDAR-supported estimation of change in forest biomass with time-invariant regression models

S. Magnussen, E. Næsset, and T. Gobakken

With more than 20 years of experience with LiDAR-aided AB estimation of key forest inventory attributes, there is sufficient knowledge to guide an analyst towards a reasonable a priori choice of model. A priori chosen models are, by necessity, parsimonious and rarely present an optimal fit to the observed data. In other words, the analyst deliberately chooses to forgo a minimization of estimates of uncertainty for the sake of achieving nearly unbiased estimates of uncertainty. In this study, we chose a regression model with two explanatory variables: the mean canopy height raised to a power of 1.5 (*mz*), and the standard deviation of canopy heights (*sz*).

Segmentation - individual tree detection (ITD)

individual tree information is important for merchandizing multiple products and valuation

Individual Tree Delineation & Attribution

Calculate and catalog valuable forest metrics such as tree height, canopy cover, stem density, and crown area for individual trees in both forest and urban settings. Quantum Spatial has developed cutting-edge methodology that generates forest metrics directly from LiDAR points, which produces an accurate, detailed tree databases for entire study areas.

Dr. Jim Flewelling has conducted pioneering work in this area of tree segmentation.

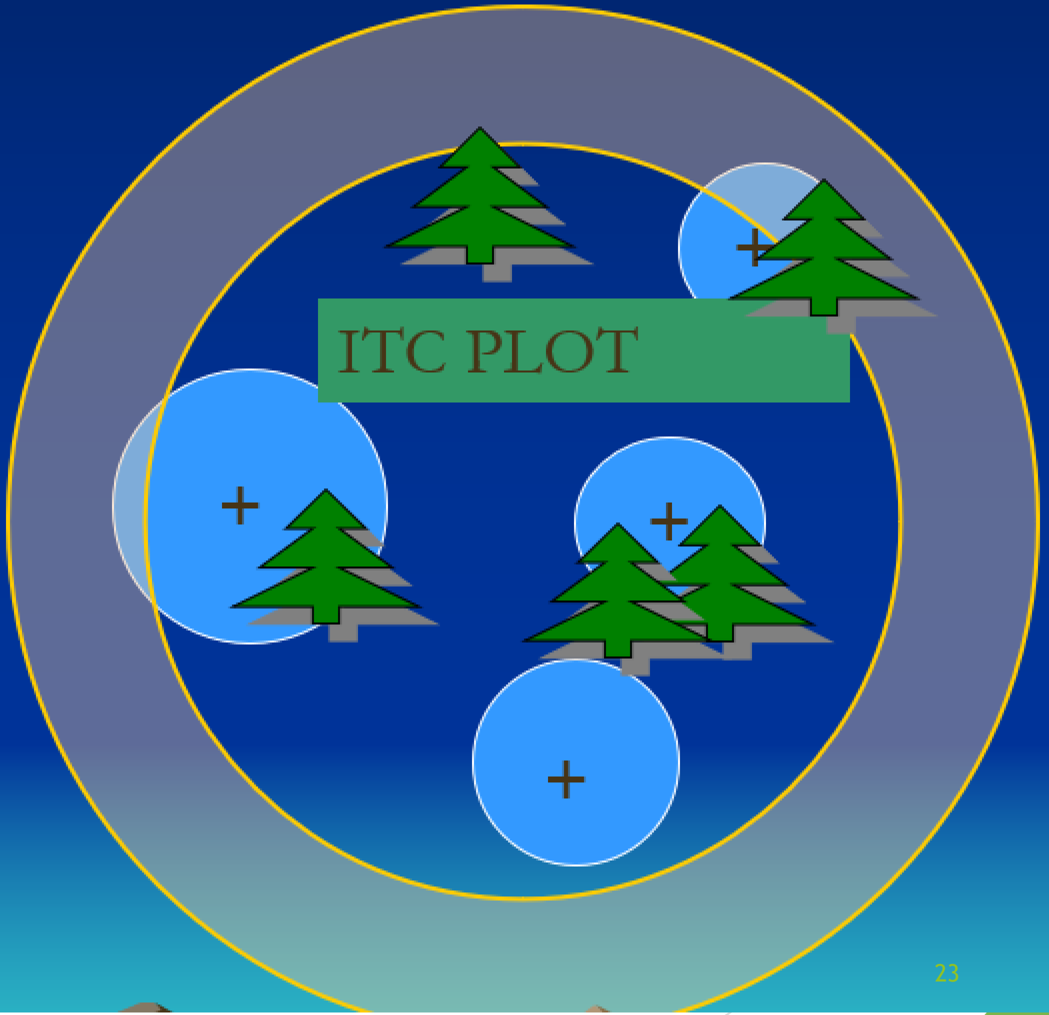
ITC and Tree Matching

Rule-based matching system.

Accept errors.

ITC may have 0, 1, 2 … trees assigned.

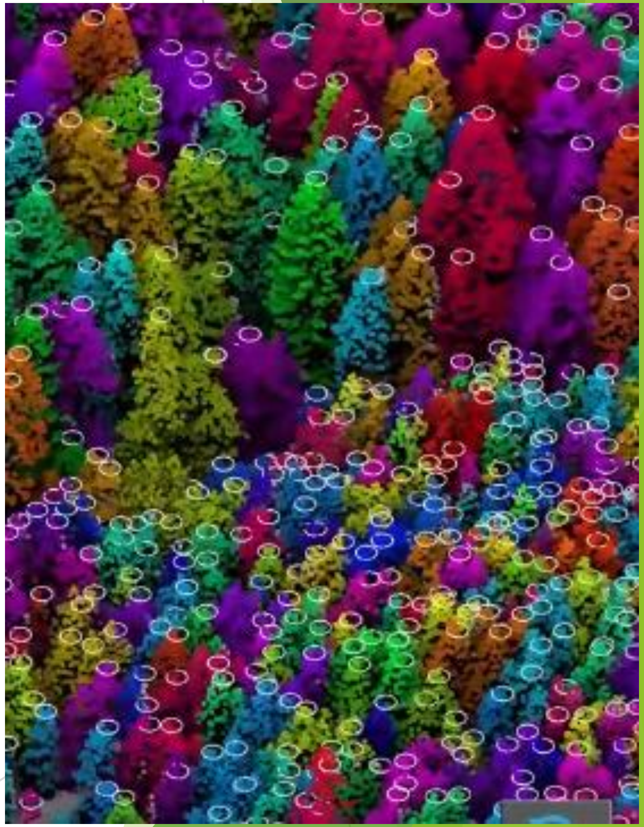
Unseen trees tabulated separately.



Jim Flewelling’s assignment of probabilities of 0,1, or 2 trees on ground for the image circle, circumscribed on the tree top

Tree count –example eqn.

- ITC whose preliminary species is Pine, area is A, and height is H_{ITC}
- $Pr\{Count \geq 1\} = logit^{-1}(L) = exp(L)/[\,1 + exp(L)]$
- Where $L = -2.43 + 0.455 \times A + 0.1667 \times H_{ITC} - 0.01553 \times (A) \times (H_{ITC})$.



Experience with the forest products industry using remote sensing mensuration

- New, energetic firms with innovative scientists, have proprietary techniques for remote sensing interpretation. They align themselves with venture capitalists and marketing managers.

- None of the vendor partners have a firm grasp of forestry. The remote sensing specialists are initially dismissive about any need for expertise in statistics or sampling.

- Capital is limited, so when the remote sensing firm approaches the forest products management, there exists a strong desire by the vendor to go ‘operational’ quickly and ramp-up company-wide.

- Foresters are conservative by nature. They demand to inspect the remote sensing mensuration results with a pilot study.

Optimism vs caution

- Executives of the forest products companies are optimistic. Remote sensing is glamourous. Promises of lower costs and higher precision are alluring.
- Operational foresters are alarmed by some of the costs. Timber inventory or cruising is typically conducted at specific age intervals, oftentimes prior to an intermediate or final harvest. The remote sensing protocol calls for flying large contiguous tracts with images for all timber types and age classes.

Finding common ground

- The forest products company biometrician is delighted by the opportunity to get more auxiliary information for an inventory. He/she has inspected NAIP imagery and knows of many 'holes' or understocked patches in a timber tract that ground plots never seem to capture.

- Great expectations build

- The naïve biometrician might inadvertently disclose over an evening beverage that the precision of the inventory is driven by the R^2 between the auxiliary and principal variable. As long as no major management activity has occurred, the image could be several years old and still very valuable.

Finding the common ~~ground~~ foe

- Trouble brews. The vendor makes it clear that other than training plots, they want no part of a probability-based inventory of 2nd-phase ground plots. The vendor is offended that the company does not want new imagery every year for every project (image acquisition is part of their mark-up and business model)

- The biometrician finds some R&D funding to pay the vendor for the pilot study, but must rely on company Operations to pay for installing 'validation' plots. The validation plots do not meet the criteria of 1st or 2nd phase plots and have GPS coordinates with large Hdop settings (wide tolerance).

- More trouble: Operations is miffed by the unbudgeted cost of the 'validation' inventory. Pressure builds on the biometrician to achieve desired results.

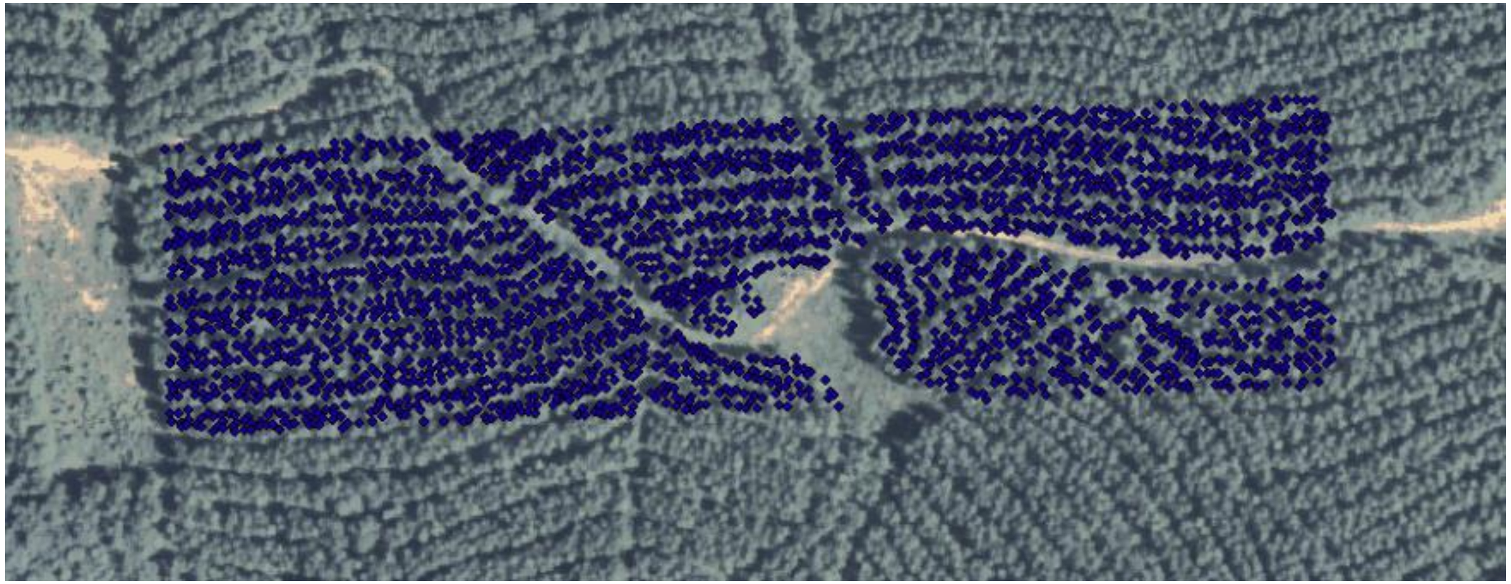
Results are disappointing

- Yep, no doubt. The estimate from remote sensing inventory is different than the estimate derived from the validation plots.
- The validation plots are nothing more than a simple random sample (SRS) for the stand of interest.
- Since there is no real design-based, model-assisted sample, the biometrician can't compare the statistical efficiency between SRS and remote sensing mensuration.
- All parties are displeased. The biometrician gives serious thought to a career change. Others scramble to find other potential uses for LiDAR such as measuring leaf area index (LAI).

Sage remarks by the late Bob Curtis seem appropriate here. Curtis, R.O. 1982. Some notes on validation of stand simulator models. West. Mens. Portland, OR.

- Curtis is emphasizing the difficulty of validation with temporary plots. His comparison is between simulated yields and observed volumes. Let’s replace the words ‘simulation model’ with remote sensing (RS) inventory.
- Allied to this is the suggestion, …, that one should run out and establish some plots for the specific purpose of validating an existing ~~simulation model~~ RS inventory. This sounds plausible, but again there are difficulties.
- If observation agrees with the ~~model~~ RS estimates, everyone is happy. If not what then? Do we conclude that the ~~model~~ RS inventory is incorrect? That the sample is inadequate (it likely is)?

Ideally a test is conducted, at least partially, where the target population is known



13.1 ha study area in Warren County, GA.

All stems are mapped

How can we move remote sensing mensuration in the private sector from pilot studies to operational status

- I would like to see this problem addressed at a national level, since it has a large impact on private forest management nationwide.
- We must bridge the gap in inventory between remote sensing interpretation and well-established sampling estimators.
- I believe that an industry-university (and perhaps Forest Service) partnership is the best path for success. This could become a major emphasis of Phase III of the National Science Foundation, Center for Advanced Forestry Systems. Up until now however, the response from vendors has been tepid.

Venturing an opinion – Unmanned Aerial System (UAS) is potentially a great fit for forest inventory

- Inexpensive, simple to fly
- Can operate in confined spaces
- Low flying speed which results in less blur for a given shutter speed
- Suitable for stereo photogrammetry, enabling 3D point cloud reconstruction

- The drone is well suited for timber cruises which are scheduled by age or size class, typically before an intermediate or final harvest. In a compartment or forest district, perhaps only a handful of stands are scheduled to be cruised.

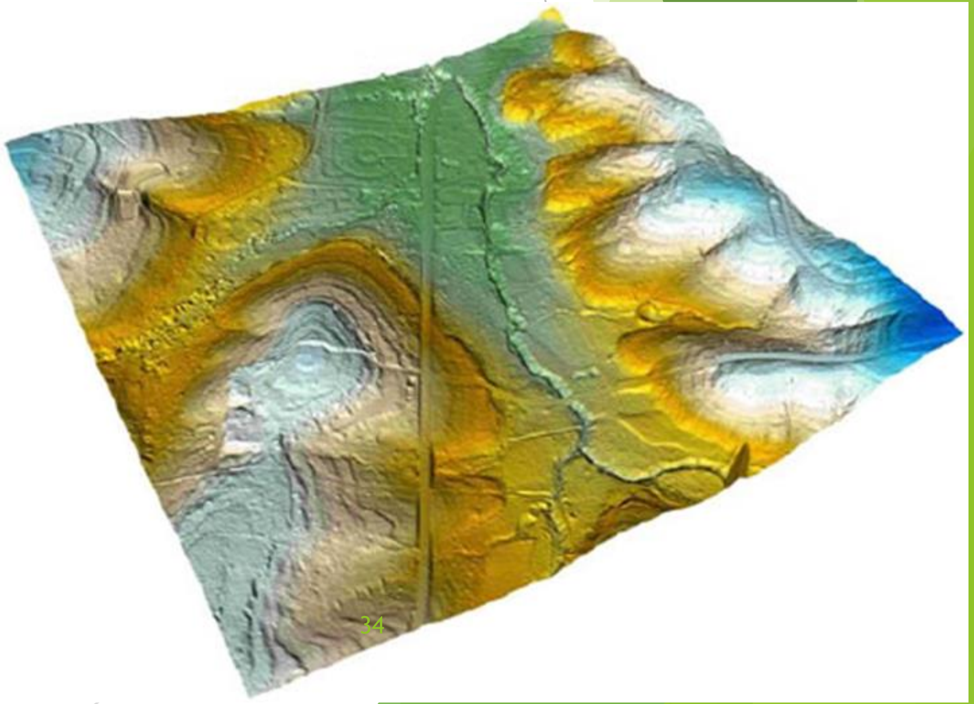
Back to the future – combining active and passive remote sensing

- This idea is advocated by Dr. Randy Wynne of Virginia Tech.

- Big picture

1. Acquire good DEM data for area of target population with low pulse LiDAR (≈2 pulses/m^2)
2. Acquire high overlap aerial photography to generate high resolution 3D point clouds

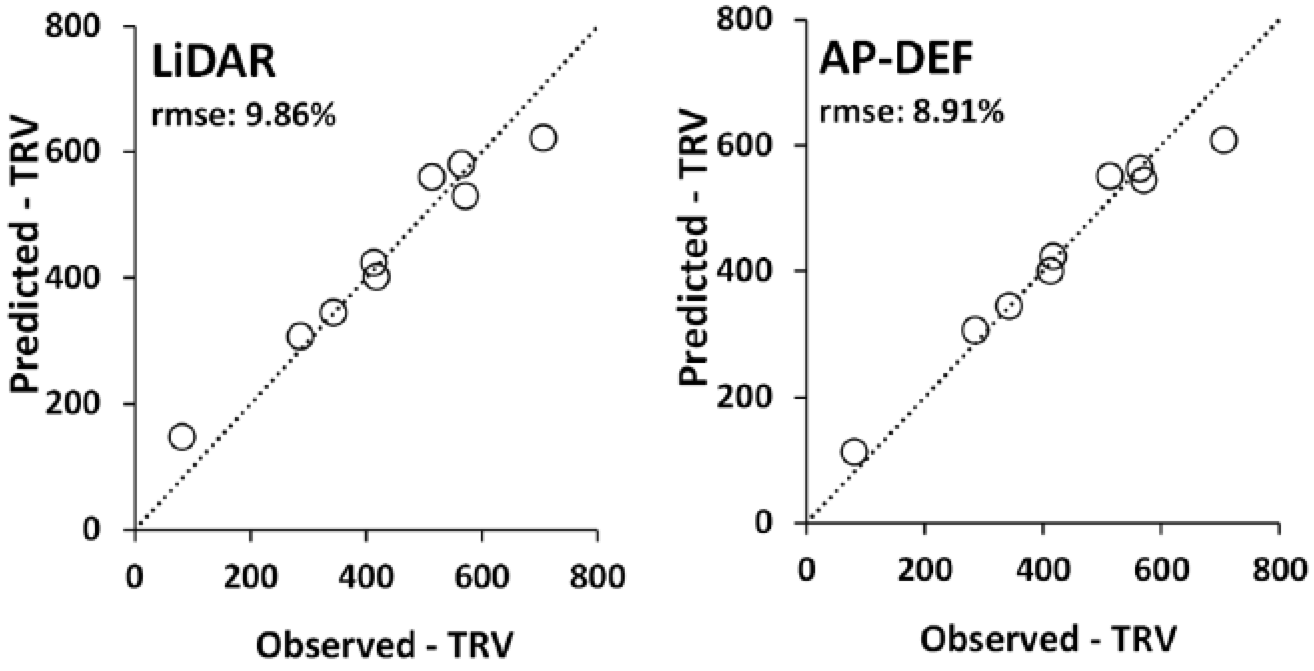
Digital Elevation Model (DEM)—DEMs are digital cartographic representations of the elevation of the land at regularly spaced intervals in *x* and *y* directions, using *z*-values as a common vertical reference.



Structure from motion multi-view stereo (SfM-MVS)

- Key advantage – aerial photographs are less expensive to acquire than LiDAR
- Metrics are easily extracted from stereo photos with a process that is commonly applied to LiDAR Canopy height models (CHM)
- Aerial photograph (phodar) data compare favorably with estimates derived from LiDAR

Comparisons using *P. radiata* in Tasmania



DEF =Default settings

Figure 8. Scatterplots of observed total recoverable volume (TRV, m^3/ha) averaged at stand level on predicted TRV using LiDAR and aerial photogrammetric -DEF (AP-DEF) models. Each observation represents a single stand (n=9).

Comparison (continued)

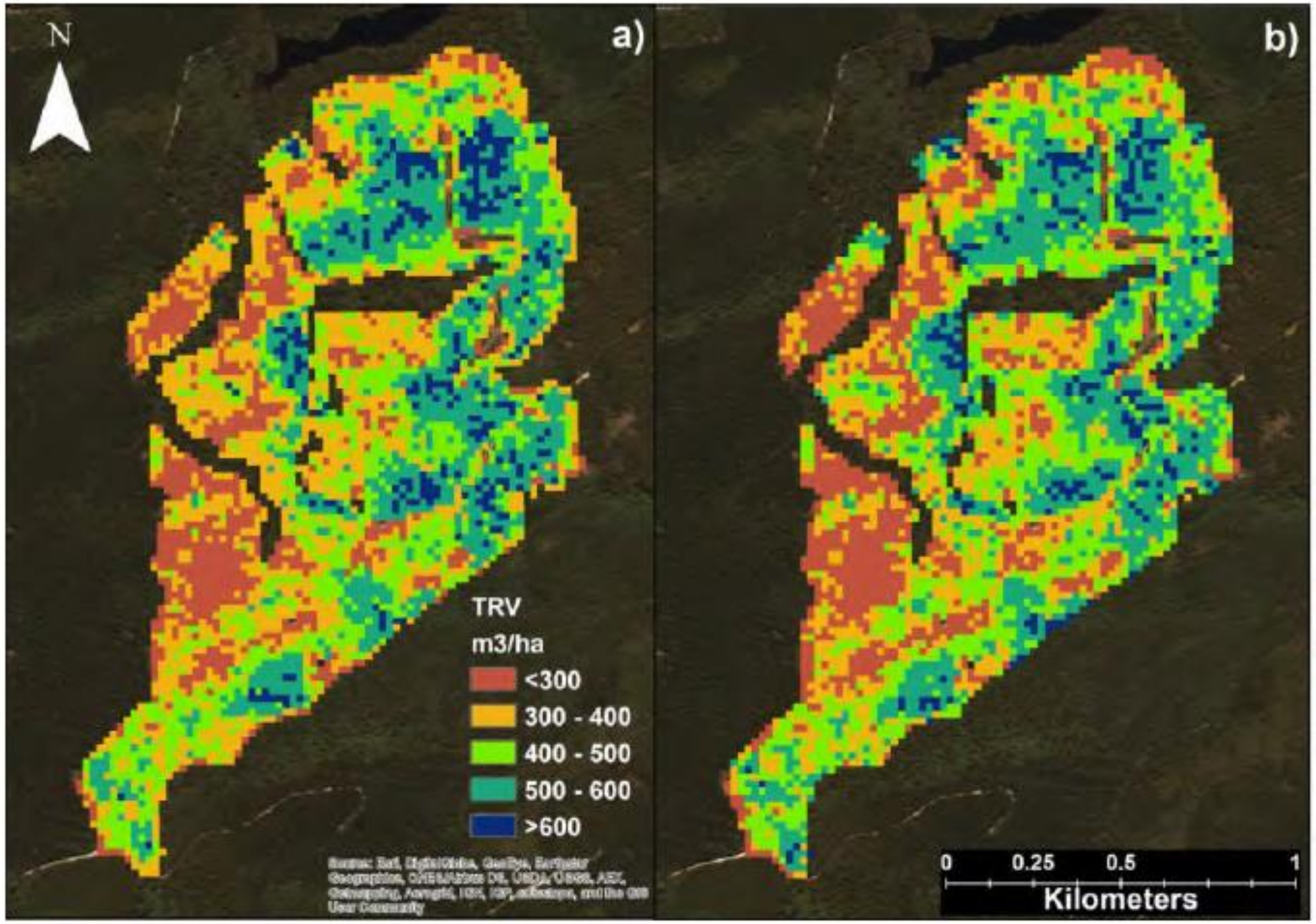


Figure 9. Total recoverable volume (TRV, m³/ha) grids generated using (a) LiDAR and (b) aerial photogrammetric -DEF models.

Sun angle (shade), wind, and % overlap of adjacent photos have a large impact of results of the aerial photograph

$R^2 = 0.88$

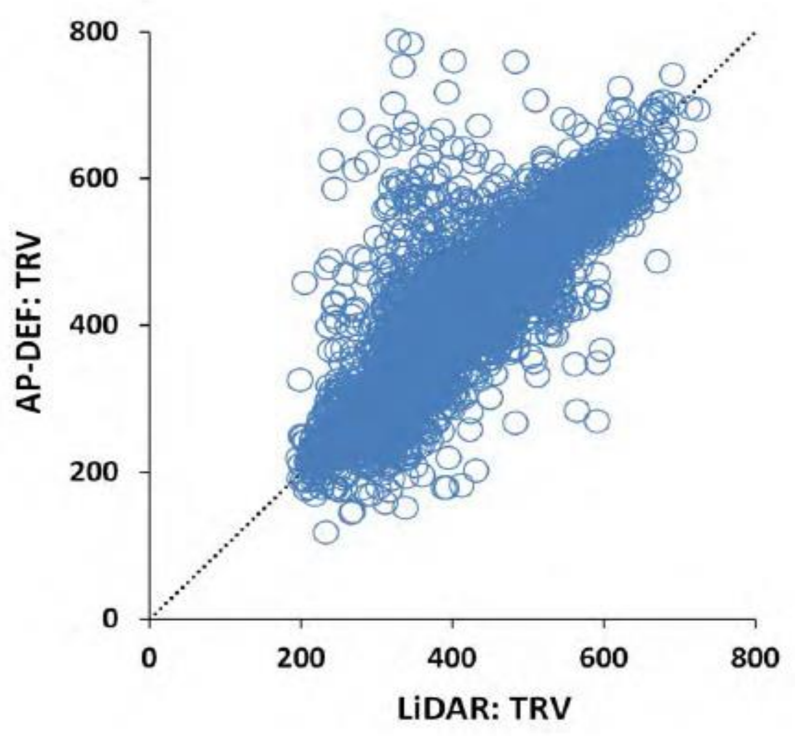


Figure 10. Scatterplot of total recoverable volume (TRV, m^3/ha) values predicted from (a) LiDAR and (b) aerial photogrammetric -DEF models.

Conclusion

- The private forest-based sector is ready to adopt remote sensing mensuration but is cautious.

- Vendors offer scant to no information about their interpretation techniques and in most cases, but not all, make no effort to integrate interpretation with probability-based sampling techniques.

- A research cooperative university-industry alliance represents an ideal mechanism for addressing the need on the integration of remote sensing interpretation and ground-based sampling. Together with the NSF-CAFS, this research problem can garner national attention and additional funding.