Scaling plot to stand-level lidar to province in a hierarchical approach to map forest biomass in Nova Scotia

Chris Hopkinson, David Coville, Danik Bourdeau, Suzanne Monette & Robert Maher

Applied Geomatics Research Group,
Centre of Geographic Sciences,
Lawrencetown, Nova Scotia,
B0S 1M0, Canada
Chris.hopkinson@nscc.ca

1. Introduction

The amount and range of biomass stored within a forested stand is an indicator of its status and ecosystem functioning (Brown, 2002). In Atlantic Canada, revenues from sawlog and pulp wood forestry products, critically important to the rural economy, have been in a steady decline in recent years (APEC, 2008). At the same time, public energy utilities have been rising to the dual challenge of meeting growing energy demands while attempting to reduce greenhouse gas emissions. The Province of Nova Scotia, for example, has committed to 25% renewable energy supply by 2015 and 40% by 2020, and biomass is seen as a potential viable source of long-term carbon-neutral alternative energy to supplement more traditional sources (NSDE, 2010). For these reasons, the ability to map forest biomass in Nova Scotia at a scale appropriate for land management has economic, ecological, environmental value.

This paper describes the preliminary results of a study that used lidar sampling data to calibrate and extrapolate above ground forest biomass from field plot to forest stand to provincial scale.

2. Data sources

The Province of Nova Scotia is a little over 50,000 km^2 and of this area >80% is forested. The forests of Nova Scotia are catalogued and monitored by the NS Department of Natural Resources (DNR). There are two publicly available and spatially explicit datasets that describe these resources and have been used as the basis for modelling in this project: a Permanent Sample Plot (PSP) database and a Forest Resource Inventory (FRI) GIS database. The PSP database details the attributes of all trees with a stem diameter at breast height (DBH) > 9.1 cm within 3250 plots of 11.3 m radius. The PSPs are randomly distributed throughout the forests of Nova Scotia and each one covers a 400 m^2 circular area. About half of the 3250 plots were established from 1965 to 1970, while the rest were established between 1998 and 2002. About 650 plots are revisited every year to ensure a five year rotation for each plot. Within each plot, living and dead trees are numbered and several attributes are recorded for each tree including height, DBH, species, signs of disease, cause of death etc.

The FRI database describing the total forest coverage within Nova Scotia contains approximately 1.1 million stand polygons that are delineated from aerial photographs and intended to describe regions of stand similarity within contiguous parcels of land. Mean stand area is ~ 3.6 Ha but the size is highly variable. The FRI database is primarily updated from air photos collected on a ten year rotation. Using stereo pair photographs, interpreters delineate homogeneous stands of trees to extract crown closure, stand height, species and land capability information. Satellite imagery is used in between aerial photo missions to update the FRI for noticeable change, such as clear-cutting. Both PSPs and FRI stands are continuously updated on a revolving basis as opposed to updating the whole Province at one time.

While the DNR GIS databases contain extensive stand-level inventory data covering the Province, the conversion of these data into meaningful estimates of available and sustainable biomass energy requires calibration (Townsend, 2008). At the local scale, PSP data collected in the field allow reasonably accurate calculation of biomass over small areas (Lambert *et al.* 2005), which have been used to provide coarse estimates of biomass for Nova Scotia down to the ecoregion scale (Townsend, 2008). However, to derive more spatially explicit estimates of biomass at the typical management unit or stand-scale is challenged by the heterogeneity displayed by Acadian mixed wood forests.

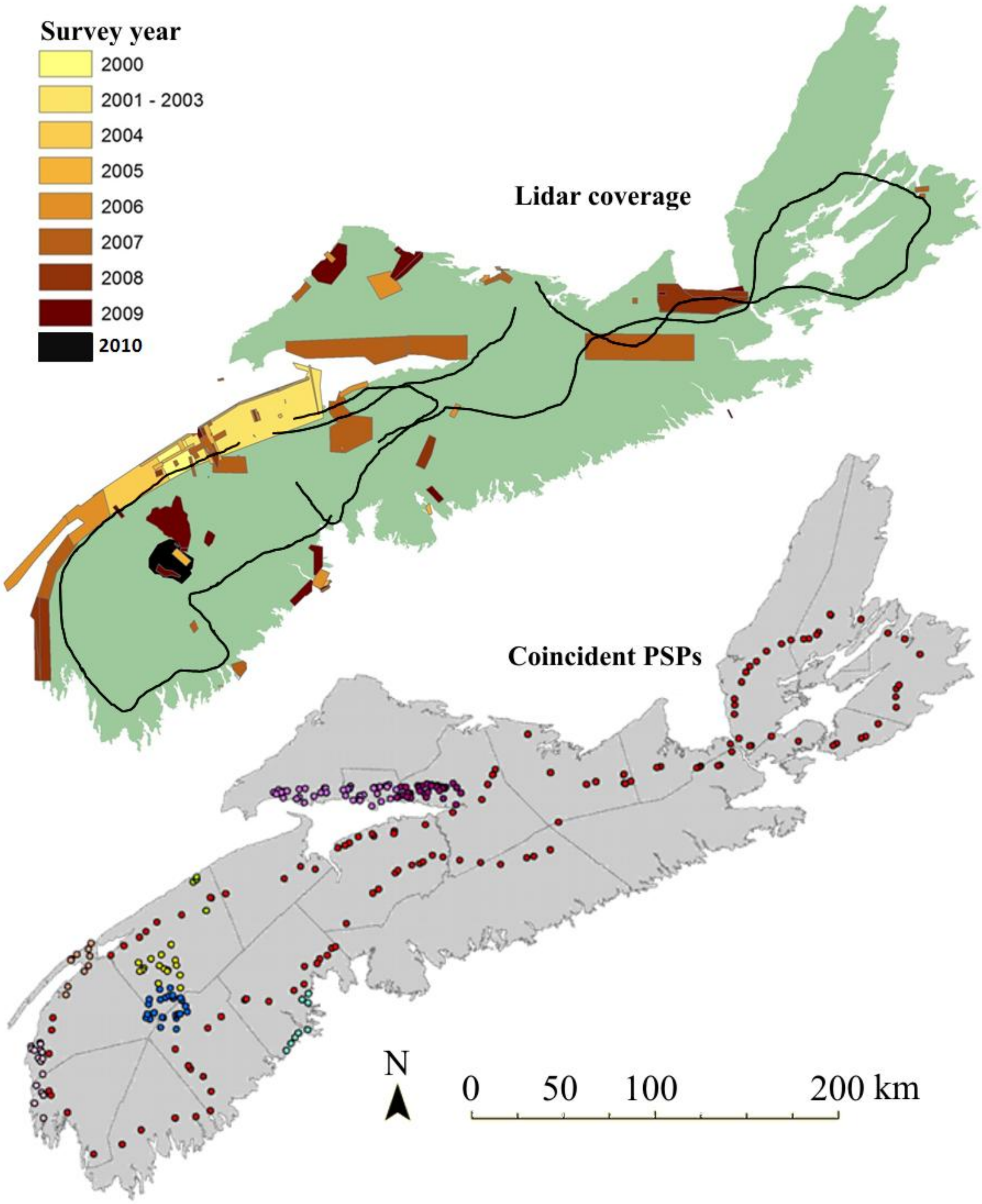

Figure 1: Top - AGRG Lidar survey polygons and sampling transects from 2000 to 2010 within Nova Scotia. Bottom – spatially coincident permanent sample plots that are within two years of a lidar survey.

To scale between PSP and FRI data layers to develop a spatial model of biomass representing both spatial domains requires a data source that can sample the canopy structure within a plot and allow for effective aggregation at the stand scale. Airborne lidar data have been shown time and again to be ideally suited to the task of plot- and stand-level canopy structure and biomass modelling (Lim *et al.* 2003). Lidar data have been collected across Nova Scotia by the Applied Geomatics Research Group since 2000 (Figure 1). Several polygons and sample transects covering ~ 10,000 km^2 or ~20% of the area of Nova Scotia have been mapped using an Airborne Laser Terrain Mapper (ALTM) 3100C (Optech Inc. Toronto, Canada). Only ~ 50% of the data were suitable for this study, as a threshold of ~ 1point/m^2 was applied to ensure a high density of data for subsequent model generation. Of the 3250 provincial PSPs, 281 were spatially coincident with lidar cover, and of these 99 were culled after applying a 2 year temporal buffer. The data sources and associated modelling approaches are described in Table 1.

Table 1: Data sources and attributes used, domains of spatial representation and notes on how the data were used in the Provincial biomass modeling approach

Raw data [number units/ total db units]	Spatial model scale [unit area] (Total db area)	Data attributes	Modeling approach	Modeling purpose
PSP attribute db [258 / 3250] 8%	Plot [400 m^2] (1.3 km^2)	Height, DBH, species, stem count	Based on Lambert et al. (2005). Species divided into hardwood / softwood	Generate biomass ‘ground truth’ at plot scale
Lidar point cloud [1000km^2/50,000km^2] >2%	Lidar survey coverage [~1 m point sampling to 500 km^2 polygon] (~ 10,000 km^2)	Height percentiles & vertical distribution ratios	Linear, quadratic and JGLS regression models with < 2 variables to predict bole / whole tree biomass	Calibrate FRI stand data by extrapolating PSP-based lidar model
GIS stand polygon [2639 / 1.1 million] 2.4%	FRI stand [~0.01 km^2 to 10 km^2] (~ 42,000 km^2)	Mean canopy height & closure	Linear, quadratic and JGLS regression models of bole / whole tree biomass	Simulate stand level biomass and aggregate up to Province

3. Summary of methods

After initial quality control of the coincident lidar and PSP data, there were 182 PSPs between the years 2005 and 2010 available to train and test a lidar-based model of biomass. PSP data were used to derive ground truth estimates of bole and whole tree dry biomass through the application of a robust individual tree biomass model that was constructed from plot-level sample data collected across Canada (Lambert *et al.* 2005). Lidar data metrics collected over 13 different survey missions using the same ALTM 3100C sensor were extracted for each of the PSPs. Summary statistics extracted using FUSION (McGaughey, 2010) describing the vertical within-plot lidar frequency distributions and point cloud ratios were used to describe canopy height and cover attributes. These lidar ‘metrics’ were then correlated with the associated PSP biomass estimates to construct predictive models of biomass. The lidar-based biomass data were then used to train an FRI stand-based model using canopy height and closure attributes from 1873 stand polygons. The lidar-trained FRI model of biomass was then applied to all 1.1 million stand polygons within Nova Scotia to generate a spatially explicit Provincial total biomass estimate. This overall approach is here referred to as plot- lidar-stand (PLS) model calibration.

4. Results & discussion

4.1 Plot-level biomass

Statistical descriptions of each PSP lidar point cloud dataset generated in the FUSION software (McGaughey, 2010) were tested for inter-correlation and suitability for use in multivariate biomass modelling (Table 2). As expected, all height-based frequency distribution metrics demonstrated high inter-correlation, as did most ratio-based metrics. While other derivatives of the frequency distribution are possible, it was decided to keep the PSP biomass modelling approach simple to allow for maximum transferability across diverse lidar datasets. Consequently, models tests were limited to two variables; one height-based and one ratio-based metric, as these demonstrated the least inter-correlation (Table 2). Furthermore, height metrics are an index of canopy height (e.g. Naesset, 1997) while ratio metrics are an index of canopy cover (e.g. Hopkinson and Chasmer, 2009). These two attributes are logical indices of the two physical dimensions (height and width) that are fundamental to volume, and therefore, biomass calculations.

Table 2: Correlation matrix of selected PSP lidar point cloud frequency distribution attributes. Shaded cells denote correlations between height and ratio metrics. **Bold** values illustrate weakest inter-correlation (Pearson's r < 0.5) and therefore increased suitability for multivariate modeling.

	ElevMean	**ElevStdDev**	**ElevP70**	**ElevP75**	**ElevP90**	**ElevP95**	**ElevP99**	**Perc1stReturns>Mean**	**PercAllReturns>Mean**	**AllReturns>Mean/ Total1stReturns*100**	**Perc1stReturns>1.50**	**PercAllReturns>1.50**	**AllReturns>1.50/ Total1stReturns *100**
ElevMean	1												
ElevStdDev	0.81	1											
ElevP70	0.97	0.89	1										
ElevP75	0.97	0.91	0.97	1									
ElevP90	0.93	0.95	0.95	0.97	1								
ElevP95	0.91	0.95	0.93	0.95	0.99	1							
ElevP99	0.86	0.93	0.87	0.90	0.96	0.98	1						
Perc1stReturns>Mean	0.76	0.58	0.78	0.75	0.62	0.58	0.51	1					
PercAllReturns>Mean	0.72	**0.45**	0.71	0.67	0.53	**0.49**	**0.41**	0.93	1				
AllReturns>Mean/Total1stReturns*100	0.80	0.72	0.85	0.83	0.74	0.70	0.64	0.97	0.86	1			
Perc1stReturns>1.50	0.71	**0.46**	0.69	0.67	0.58	0.56	**0.49**	0.83	0.84	0.79	1		
PercAllReturns>1.50	0.70	**0.34**	0.63	0.61	0.53	0.51	**0.45**	0.64	0.73	0.59	0.91	1	
AllReturns>1.50/Total1stReturns*100	0.86	0.66	0.85	0.84	0.78	0.76	0.70	0.82	0.75	0.86	0.91	0.85	1

Using PSP data and the biomass model of Lambert *et al.* (2005), several lidar biomass models were trained and tested for whole tree and bole. For the lidar model, species information was ignored but the ratio of softwood to hardwood stems was considered in PSP model training. Root mean square error (RMSE) for the best polynomial regression whole tree and bole lidar model remained approximately the same in both test datasets at ~ 26%. The explanation of variance in the test data was greater for bole biomass estimates at 75%, than whole tree biomass at 63%. All models were significant at the 99% level of confidence. These results indicated that the lidar models were robust enough to be applied at the stand-level.

4.2 Stand-level biomass

A challenge that became immediately apparent during the process of relating lidar biomass estimates to FRI stand-level attributes, was the temporal latency between the two datasets (FRI data dating back to 1990s in extreme cases, while the lidar data used in this study ranging from 2005 to 2010). This latency was most evident when comparing stand-level FRI mean tree height with the mean maximum of the lidar data aggregated into 25 m grid cells (Figure 2). Using canopy height as an indicator, quality control procedures were put in place to systematically remove the most obvious outliers (due to growth and clear cutting) using objective height, stand age and date criteria. However, even after this quality control process, the latency between lidar and FRI still has the potential to propagate uncertainty into the model. The nature of this error is such that any stand growth, decay, thinning or clear cut occurring following the last FRI stand update and preceding the associated lidar acquisition will lead to divergence between the lidar and FRI attributes. As long as the forests are in a state of dynamic equilibrium (i.e. the Provincial forest resource as a whole is neither increasing nor decreasing appreciably), then any stand-level biases will not necessarily lead to a systematic bias in the overall population statistics. In practical terms, this means that we expect the model to display a high level of variance at the stand population scale but when aggregating biomass estimates to larger and larger spatial domains, there should be a level of compensation between high and low model predictions.

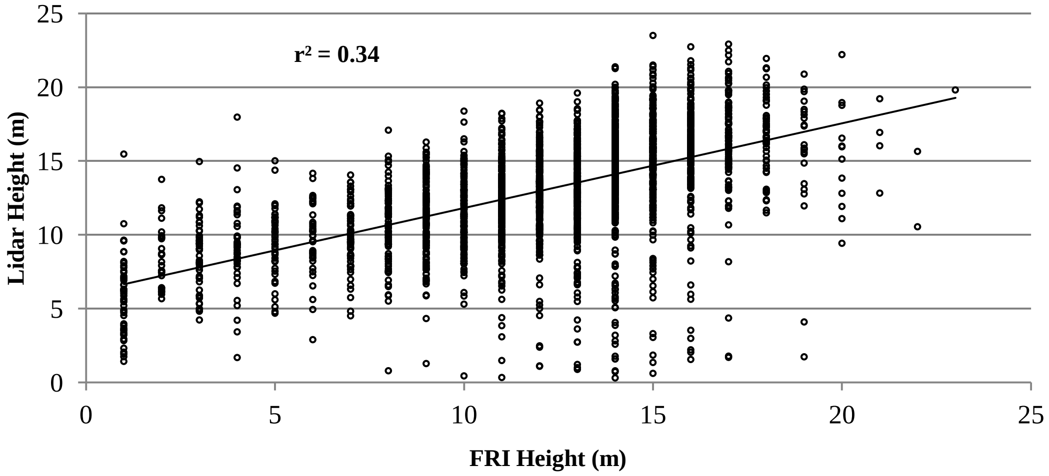

Figure 2: Mean maximum grid level lidar canopy height vs. FRI mean canopy height for stands completely covered by lidar data (n = 2639).

Given the accepted level of uncertainty in the models at this stage of the analysis, no attempt was made to develop highly sophisticated or complex regression models between stand-level lidar biomass and FRI stand attributes. Crown closure and mean canopy height were chosen as the FRI attributes to be used in a simple bivariate regression model as they most closely resembled the lidar metrics used in the previous modelling step. Similar to the PSP results, the RMSE in stand biomass approximated 27% both for whole tree and bole. However, the explanation of variance dropped to 41% and 43%, respectively, most likely a large function of the temporal latency issues described above.

4.3 Nova Scotia's biomass

Using the PLS modelling approach summarised above, we derived six estimates of total provincial biomass; three for bole (stem wood) and three for whole tree. The three modelling approaches did not differ in terms of the data used at each stage of model development, rather the differences are simply in terms of the algorithm construction; ranging from simple single variable linear regression to bivariate quadratic and a further model that mimicked the joint generalised least squares structure of the model proposed by Lambert *et al.* (2005). Given the Lambert *et al*. (2005) model was used to derive the 'ground truth' plot-level dry biomass estimates from which the rest of the lidar and stand-level predictions are based, the model results are to be considered more reliable if expressed as dry biomass. The values for total bole biomass within the Province ranged from 253 – 260 x 10^6 dry tonnes, while whole tree biomass ranged from 365 – 373 x 10^6 dry tonnes (Figure 3). [Bole biomass is the number to refer to if only stems are to be extracted for further processing, while leaving tree tops in situ for nutrient recycling.] For the sake of comparison, the Nova Scotia Department of Natural Resources (DNR) PSP-based estimate of total living merchantable whole tree biomass in the Province is 309 x 10^6 dry tonnes (Townsend, 2008). The Canadian Forest Service (CFS) has also developed a largely satellite image-based 1km grid-cell resolution model (Hall *et al.* 2010) that produces an estimate of 362 x 10^6 dry tonnes for total above ground biomass within Nova Scotia.

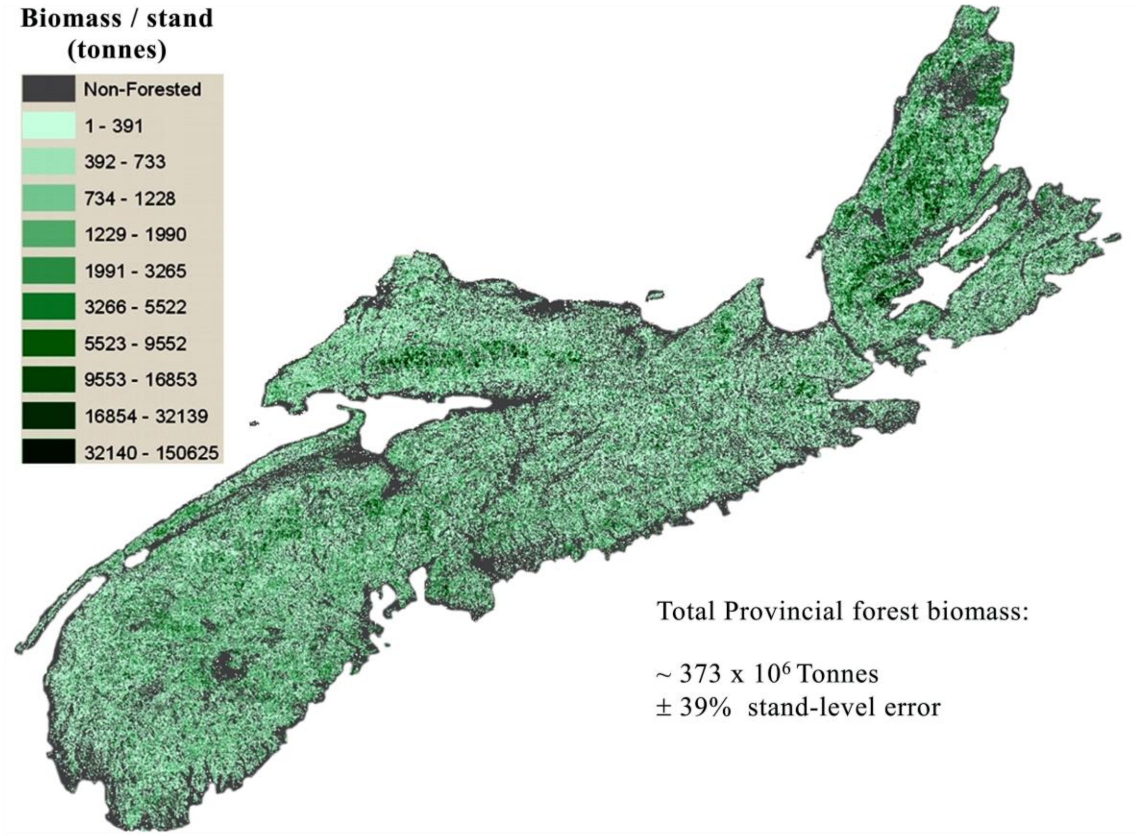

Figure 3: Map of predicted forest stand-level biomass across Nova Scotia

Both the DNR and the CFS estimates of total above ground biomass for the Province of Nova Scotia are lower (by 17% and 3%, respectively) than that generated using the PLS scaling approach described here. The DNR biomass value refers to living whole tree dry biomass but it should be noted that this only considers stems with DBH > 9.1 cm. The whole tree biomass estimate using PLS includes standing dead stems and given lidar cannot differentiate stems based on DBH it is possible the estimate was further inflated relative to the DNR estimate by inclusion of small and immature tree stems. The primary difference between PLS and the DNR and CFS approaches is that the approach described here allows calculation of biomass at the stand scale and is useful for operational planning and decision making. Based on a conversion ratio of 140% for dry to green biomass for typical Acadian mixed wood species (e.g. Shelton and Shapiro, 1976), the estimates above provide values of around 357 x 10^6 green tonnes for total bole biomass and 514 x 10^6 green tonnes for whole tree biomass.

Given the PLS approach uses three modelling steps, each building on the previous, uncertainty will propagate throughout. The RMSE values observed in the model results at each stage, demonstrated errors in the 15% to 30% range. Propagating the RMSE at the PSP, lidar and FRI stand modelling steps in quadrature compounds to an overall stand- and Province-level error of ~ 39%. This assumes that all errors are random and there is no significant bias.

5. Conclusions

For the time period from 2005 and 2010, 182 PSPs were used to train and test a lidar forest biomass model. The products of this model were then used to train a stand-level model from 1869 coincident FRI stands. These results were then aggregated across all 1.1 million stands in Nova Scotia to arrive at a total above ground forest biomass estimate for the Province. This biomass estimate can be expressed several ways but the whole tree dry biomass estimate is ~ 373×10^6 tonnes ±39%. Where lidar data are available in the Province (about 20% of the land surface area) a spatially explicit estimate of biomass can be generated at the 25 m grid cell resolution. In other areas, biomass can be estimated at the stand-level. The spatial resolution of these estimates constitutes an improvement over previous biomass estimates that were available at the ecoregion (DNR; Townsend, 2008) or 1 km pixel (CFS; Hall et al, 2010) resolutions. These results, therefore, can be used to aid in either stand- or within stand-level forest management practices and in informing forest biomass energy policy in Nova Scotia.

Modeling biomass over such a large area is not without challenges. Greatest of these is obtaining useable model calibration and validation data. In this study, DNR PSP and FRI data were all that were available at the scale required. Both data sources were limited in terms of temporal compatibility with the lidar data that were used to scale between the two. Up to two years of latency in the PSP data is less than ideal given forests grow, die and are managed. However, this was less problematic than the > 10 years of latency for some of the FRI stands. The time discrepancy will introduce larger errors for younger stands and for those that have been clear cut. While objective criteria were used to mitigate such occurrences it is impossible to remove all such instances without manual selection and verification of each stand. Such an approach is not practical or even feasible at the provincial scale so a substantial amount of model uncertainty remains. However, it is assumed that temporal discrepancies will cause both over- and under-estimation of stand-level biomass, such that there will be some compensation.

Acknowledgements

Many thanks to Allyson Fox and Laura Chasmer for their dedication to flying long airborne survey missions and assisting with data processing. Nova Scotia Power Inc. is acknowledged for funding this project, and the Canada Foundation of Innovation for funding AGRG's lidar laboratory. Nova Scotia Department of Natural Resources provided the raw PSP and FRI GIS data layers, while Cory Isenor, Collin Gillis and Heather Morrison assisted with field data collection. Dr Bob McGaughey of the US Department of Agriculture is gratefully acknowledged for his support with FUSION. Thanks are also extended to Dr. Ron Hall of the Canadian Forest Service for providing a 1 km resolution map of total biomass for Nova Scotia.

References

Atlantic Provinces Economic Council. 2008. *Building competitiveness in Atlantic Canada's forest industries: A strategy for future prosperity.* Report published by APEC, Halifax, Nova Scotia. 75pp.

Brown, S. 2002. Measuring carbon in forests: current status and future challenges. *Environ. Pollut*. Vol. 116: 363-372

Hall, R.J. Skakun, R.S. Beaudoin, A. Wulder, M.A. Arsenault, E.J. Bernier, P.Y. Guindon, L. Luther, J.E. and Gillis. M.D. 2010. Approaches for forest biomass estimation and mapping in Canada. DVD pp. 1988-1991 in *Proc. 2010 IEEE International Geoscience and Remote Sensing Symposium*, Honolulu, Hawaii, USA. July 25-30, 2010. DVD

Hopkinson, C. and Chasmer, L.E., 2009. Testing lidar models of fractional cover across multiple forest ecozones. *Remote Sensing of Environment*. Vol. 113: 275-288.

Lambert, M.-C.; Ung, C.-H.; Raulier, F. 2005. Canadian national tree aboveground biomass equations. *Can. J. For. Res*. Vol. 35: 1996-2018.

Lim K, Treitz, P., Wulder, M.A., St-Onge, B., Flood, M. 2003. LiDAR remote sensing of forest structure. *Progress Phys Geogr.* Vol. 27: 88–106

McGaughey, R.J. 2010. Fusion/LDV: Software for LiDAR Data Analysis and Visualization. USDA Forest Service. FUSION Manual. March 2010 – FUSION Version 2.80.

Næsset, E. 1997. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing.* Vol. 52: 49-56.

Nova Scotia Department of Energy, 2010. *Renewable electricity plan; a path to good jobs, stable prices, and a cleaner environment.* Report published by the Govt of Nova Scotia, Halifax, Nova Scotia. 32pp.

Shelton, J., Shapiro, A. B. 1976. *The Woodburner's Encyclopedia.* Crossroads Press, Box 33, Waitsfield, VT 05673, U.S.A. 155 pp.

Townsend, P. 2008. *Forest Biomass of Living Merchantable Trees in Nova Scotia*. Report FOR 2008-9. Nova Scotia Department of Natural Resources. Govt. Nova Scotia publication.