A STRATEGY TO MAP BIOMASS OF CANADIAN NORTHERN BOREAL FORESTS USING MULTI-SENSOR / MULTI-RESOLUTION IMAGERY: APPROACH, METHODS AND PRELIMINARY RESULTS

André BEAUDOIN[a], Luc GUINDON[a], Ronald J. HALL[b],
Joan E. LUTHER[c], Richard FOURNIER[d]

[a] Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, 1055 rue du P.E.P.S., C.P. 10380 succ. Sainte Foy, QC, Canada, G1V 4C7, E-mail: andre.beaudoin@nrcan.gc.ca
[b] Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Edmonton, AB, Canada
[c] Natural Resources Canada, Canadian Forest Service, Atlantic Forestry Centre, Corner Brook, NL, Canada
[d] Département de Géomatique Appliquée, Université de Sherbrooke, Canada, Sherbrooke, QC, Canada

ABSTRACT

The mapping of aboveground forest biomass is necessary for reporting on the state of Canada's forests. Undertaking this task over the non-inventoried regions of Canada's northern boreal forests is challenging due to its remoteness, vast area, and relative lack of available inventory information. To address this problem, we developed a biomass mapping strategy that was predicated on the integration of field and multi-sensor/multi-resolution satellite data. The strategy entails generating information about stand attributes and biomass from small numbers of ground sample plots that are subsequently related to the shadow fraction generated from high spatial resolution images. This process results in an image map from which satellite sample plots can be derived and scaled to coarser resolution imagery such as Landsat Thematic Mapper using a kNN process. This strategy was applied to pilot regions in 3 ecozones located in the Northwest Territories, Quebec, and Newfoundland and Labrador. Shadow fraction models of forest biomass were relatively similar across the 3 pilot regions and independent validations reported RMSE that ranged from 18 to 21 t/ha with bias less than 10 t/ha. These values are comparable or even less than those obtained from the combined use of coarser resolution Landsat and inventory data. Ongoing work is testing method applicability with other sensors and extending application of the strategy over additional species and larger geographic areas.

Keywords: biomass mapping, boreal forest, QuickBird, Landsat, shadow fraction, kNN.

1 INTRODUCTION

The need for aboveground biomass information is driven by requirements for reporting on the state of Canada's forests, a function met by national programs directed at assessing forest carbon stocks and execution of the National Forest Inventory (NFI). This information need is being met, in part, through the recent completion of a satellite land cover map over the forested areas of Canada, an initiative of the Earth Observation for Sustainable Development of Forests (EOSD) (Wulder *et al.*, 2003), a project that was undertaken with the support of the Canadian Space Agency. Within the framework of the land cover map, a biomass mapping strategy was developed for non-inventoried northern boreal forests, and is being applied to three large northern pilot regions located in the Northwest Territories, Québec, and Newfoundland and Labrador. The purpose of this paper is to i) describe the biomass mapping strategy and related approaches and methods, and ii) present initial results from application of this strategy over the three pilot regions.

2 STRATEGY AND APPROACHES

The biomass mapping strategy is based on the integration of field and multi-sensor / multi-resolution satellite data. The strategy starts with a relatively small number of ground sampling plots (GSP) from which biomass estimates are derived using field measurements. At these plot locations, field-based biomass estimates are related to attributes of high spatial resolution imagery (HSRI, < 1m) to derive a model to map biomass at the local scale. This map is sampled to provide satellite sampling plots (SSP) as surrogates to the lack of GSP in non-inventoried northern boreal forests. Biomass SSPs are then used as calibration data to map biomass at the regional scale across a mosaic of medium spatial resolution imagery (MSRI, < 100m) with the k-Nearest Neighbour algorithm (Guindon *et al*., 2005). We describe in the following sub-sections the adopted mapping approaches according to this strategy.

2.1 LOCAL BIOMASS MAPPING APPROACH USING HRSI

Recent research (Leboeuf *et al*., 2007) has reported that tree shadow fraction (SF) obtained from panchromatic QuickBird HRSI is an efficient predictor of biomass for black spruce species largely dominating the Canadian northern boreal forests. We adopted the SF approach and undertook further evaluation with regional GSP data sets and additional species. This paper summarizes initial results from the local estimation of biomass for conifers using panchromatic QuickBird images through regional equations used to generate estimates at SSP locations. Future work entails developing a national set of equations.

2.2 REGIONAL BIOMASS MAPPING APPROACH USING MRSI

The regional mapping approach uses SSP with biomass estimates to map biomass across normalized mosaics of spaceborne MRSI. The scaling of biomass from the SSP to the MRSI is accomplished with k-Nearest Neighbour (kNN), a non-parametric algorithm that has been used operationally for inventory across the Scandinavian boreal forest (Fazakas *et al*., 1999; Reese *et al*., 2002). Similar to other methods, kNN requires GSP representing the full range of forest conditions found within the intended area for mapping. In the absence of a consistent network of GSP in northern regions of Canada, our biomass strategy promotes the use of surrogate SSP as an alternative to GSP. This approach was successfully tested for black spruce coniferous stands at the Landsat scene level in Northern Quebec and Labrador (Guindon *et al*., 2005). Our current work generalized this approach to larger regions represented by Landsat mosaics and to broader boreal forest types.

3 PILOT REGIONS AND DATA SETS

The biomass mapping strategy was applied to three large pilot regions (Figure 1) located in the Taiga Plains, Eastern Taiga Shield and Boreal Shield Ecozones, respectively. These pilot regions represent a wide range of forest conditions (stands composed of black spruce, jack pine and white spruce stands) and biomass of northern boreal forests in Canada. In addition, they are of particular interest to i) the NFI for filling large information gaps in Northern Canada and ii) partnering forest management agencies as they offer an opportunity to explore new inventory approaches that extend beyond the range of their current forest inventory.

The Northwest Territories/Fort Providence region (NT/FP) falls within the Taiga Plains Ecozone where inventory data is sparse and commercial interests for use of the land base is increasing. Dominant species include black spruce, jack pine, white spruce and trembling aspen. The Quebec/Mistassini region (QC/MS) is a northern extension of the contiguous inventoried boreal forests within the Boreal Shield Ecozone that extends into the non-inventoried forests of the Eastern Taiga Shield Ecozone. Dominant species include black spruce and jack pine associated with lichen or moss background. The Newfoundland and Labrador/Goose Bay region (NL/GB) largely consists of black spruce located on more rugged terrain conditions, and is under similar management and exploitation issues as in Quebec.

Each pilot region extent is defined by a mosaic of Landsat scenes (Landsat-7 ETM or Landsat-5 TM) (Table 1). Each Landsat scene was orthorectified with a resulting positional RMSE in the order of 15-20m and transformed into top-of-atmosphere reflectance. In addition, the scenes were normalized using a MODIS-based cross-calibration approach adapted from Olthof *et al*. (2005) and subsequently mosaicked.

Portions of each pilot region were imaged by a small number of QuickBird images (Table 1). These images were acquired under clear-sky conditions during the growing season and as close as possible to the Landsat acquisitions to minimize differences in vegetative phenology. The overall HRSI coverage was selected to represent a similar range of forest stands as those found within the Landsat mosaic. Each HRSI was geometrically corrected with a 1st

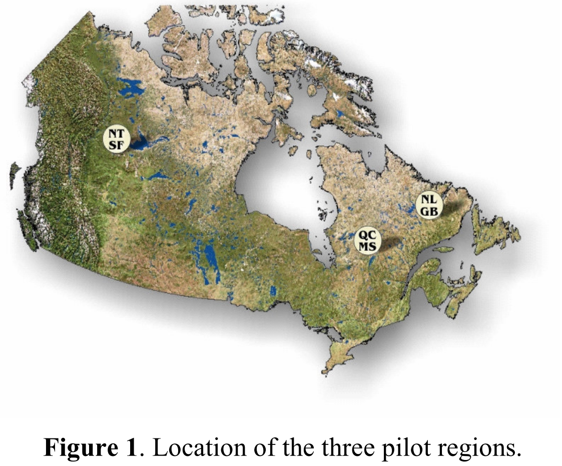

Figure 1. Location of the three pilot regions.

Table 1. Data set summary for the pilot regions.

Data type		NT/FP	QC/MS	NL/GB
MRSI Landsat	Years	2004	2004-2007	1999-2001
	Number	3*	9	4
	Area (km^2)	85 000	200 693	91 969
HRSI QuickBird	Years	2003	1999-2005	2002
	Number	1	6	1
	Area (km^2)	320	872	238
Biomass GSP	Provincial	46	-	128
	CFS	-	141	-
	species**	BS,WS,JP	BS, JP	BS
	min (t/ha)	20.4	4.8	21.1
	max (t/ha)	156.8	163.5	128.8
	mean(t/ha)	78.7	54.1	74.2

* Results from one scene reported in this paper
** BS = black spruce, WS= white spruce, JP=jack pine

order polynomial and a resulting average positional RMSE (root mean square error) of 10m. However, original delivered images with 20m RMSE positional error had to be used in some areas where accurate tie-points were unavailable (roads, water bodies). The HRSI images were transformed into SF images with 25 or 30m grid cells according to the approach reported in Leboeuf *et al*. (2007).

GSP were established and measured in the field or obtained from provincial sources (Table 1). Black spruce plots dominated the QC/MS (126/141) and NL/GB regions (128/128), whereas a variable mixture of black spruce, white spruce and jack pine species occurred in GSPs of the NT/FP region. For each GSP, tree-level biomass was estimated from a set of national biomass equations using DBH as the predictor variable (Lambert *et al.*, 2005). Biomass for each tree was consequently summed to estimate biomass at the plot-level. The range of biomass was largest for QC/MS and smallest in the NT/FP region (Table 1).

4. METHODS

The biomass mapping strategy is primarily a two-step process entailing an initial estimation of biomass from HRSIs and scaling to a Landsat mosaic using the kNN method.

4.1 LOCAL BIOMASS USING SF FROM QUICKBIRD

Biomass was estimated with linear regression equations based on tree shadow fraction (SF) as a single predictor variable similar to that employed by Leboeuf *et al*. (2007):

$$BIOsf = a + b*SF \qquad (1)$$

where SF was extracted from a 25 or 30m grid cell on the HRSI image and BIOsf is the SF-based biomass estimate. Regional equations with specific a and b values were developed for each pilot region using 70% of the GSP for calibration (GSPcal) with the remaining GSP for validation (GSPval), from which RMSE and bias were computed. The regional equations were applied to each HRSI-derived SF image to produce local biomass maps within a coniferous mask provided by the EOSD land cover map. The HRSI-derived biomass maps were randomly sampled to create a set of N SSP.

4.2 REGIONAL BIOMASS USING kNN AND LANDSAT

For each Landsat mosaic, N SSP (SSP_{cal}) were used to calibrate the following kNN equation:

$$BIOknn_i = \sum_{k=1}^{K} W_k BIOssp_k \Big/ \sum_{k=1}^{K} W_k \qquad (2)$$

$$\text{for } k \neq i \qquad \text{where } W_k = \frac{1}{d_k^j}$$

where $BIOknn_i$ is the kNN biomass estimate for the i^{th} Landsat pixel, k is the k^{th} out of K nearest spectral neighbours, $BIOssp_k$ is SSP biomass, W_k is a weighting coefficient, d^j_k is the spectral Euclidian distance to which is applied a j power value. Biomass estimated from the kNN method is a weighted average of biomass values from the SSP that are spectrally the nearest from each Landsat pixel, where j = 0 results in a simple average. Various combinations of N, k and j were tested. For consistency across the pilot regions, a single combination (k=10, j = 1, N =300) was used in Eq. (2) and applied to all coniferous pixels within the Landsat mosaic. The output was a biomass raster layer generated by the kNN method. RMSE and bias of the BIOknn estimates were calculated using (i) the SSPcal set through cross-validation (QC/MS) or from an independent SSPval set (NL/GB) to estimate the scaling error from application of the local biomass maps to the regional level; and ii) GSPval data set to compute the prediction error when compared to biomass that was derived from field measurements.

5. RESULTS AND DISCUSSION

5.1 LOCAL BIOMASS USING SF FROM QUICKBIRD

Although the strength and nature of the relationship between GSP biomass and SF varied across the pilot regions (Figure 2), the RMSE and bias, values were relatively similar across regions (Table 2). For NT/FP, where GSP were composed of three conifer species (Table 1), the regression line was comparable to the one for QC/MS with a similar goodness of fitness but a slightly higher RMSE. For the QC/MS pilot region, the linear relationship developed for black spruce (Leboeuf *et al.*, 2007) provided a high goodness of fit and the lowest RMSE. For the NL/GB pilot region, where GSP were also composed of black spruce, its narrower biomass range resulted in a shift in the intercept, lower goodness of fit but similar RMSE. Differences in relationships among the regions can be explained partly by the different biomass ranges and species composition, variability in the parameters for the SF extraction (visual-based threshold, normalization coefficient) and residual variance not accounted for by using SF as a single predictor of biomass in Eq. (1). Preliminary results suggest that the relationship is not strongly species-dependant although factors such as terrain slope, stand density, timing of image acquisition, and geographic latitude may influence the nature of the model relationships.

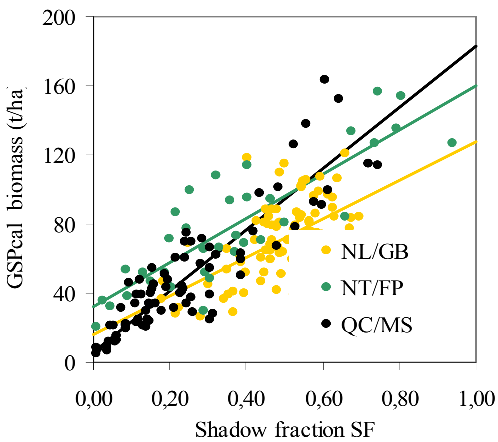

Figure 2. GSP biomass as a function of SF with regression lines for the three pilot regions.

Table 2. Linear regressions over three pilot regions according to Eq. 1 for conifers.

	Nb GSPcal	Adj. R^2	Nb GSPval	RMSE t/ha	Bias t/ha
NT/FP	35	0.72	11	24.34	1.99
QC/MS	76	0.82	32	15.29	4.18
NL/GB	90	0.35	38	18.52	1.32

5.2 REGIONAL BIOMASS USING kNN AND LANDSAT

The predicted biomass (BIOknn) and observed plot biomass (GSPval) were comparable across the pilot regions (Figure 3). Error estimates obtained using either SSPval or GSPval sets report bias with less than 10 t/ha with RMSE values that ranged between 17 and 21 t/ha. These results are encouraging for large scale implementation as the estimation error did not degrade significantly when applying the kNN approach in Quebec over a large Landsat mosaïc (9 scenes) compared to single Landsat scenes as initially reported in Guindon *et al.* (2005). Overall, estimation errors were comparable or less than those obtained from the combined use of Landsat and inventory data (Hall *et al.*, 2006; Luther *et al.*, 2006). While the application conditions (sites, forest conditions) were different compared to those reported in this paper, the kNN method driven by SSP as substitutes to GSP does appear to be a promising approach for mapping the biomass of non-inventoried regions.

6. CONCLUSION AND PERSPECTIVES

We propose and demonstrate a biomass mapping strategy for northern non-inventoried boreal forests of Canada. The strategy is based on multi-resolution/multi-sensor image data from which local biomass maps and related SSP are derived. Biomass estimates from the SSP are then scaled across a mosaic of MRSI to map biomass with the kNN

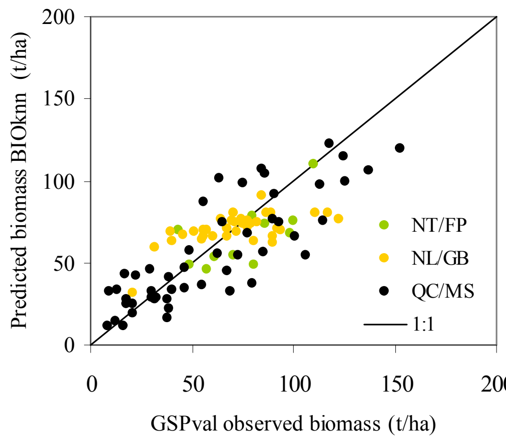

Figure 3. BIOknn biomass estimates vs GSP biomass.

Table 3. Errors for BIOknn using GSP_{val} and SSP_{val} datasets.

	Nb SSPval	RMSE (t/ha)	Bias (t/ha)	Nb GSPval	RMSE (t/ha)	Bias (t/ha)
NT/FP	na	na	na	11	19.5	9.79
QC/MS	300	22.6	-0.1	86	21.0	4.97
NL/GB	300	12.9	-0.7	35	17.9	-1.2

algorithm. This paper presented results from application of this strategy across 3 pilot regions.

The SF-based approach is applicable to biomass estimation for coniferous species within the northern regions evaluated. Current investigations are attempting to apply the strategy to a wider variety of species and to evaluate its applicability with other high spatial resolution sensors such as IKONOS. The potential to pool, partially or totally, the various data sets will be investigated in order to assess if a generalized set of regional or national equations to estimate biomass at the cover type level can be developed. Other approaches are also under investigation to address the limitation of the SF-method for application to deciduous and mixedwood species, and to determine if additional attributes derived from individual tree crown delineation and lidar-derived height can be used as predictor variables for improved estimation of biomass.

The kNN method driven by SSP appears to be a promising approach to estimate biomass across a mosaic of Landsat data that yielded estimates whose errors were comparable or even smaller than those reported previously from the combination of Landsat satellite data and inventory data.

Overall, the biomass mapping strategy developed and tested across 3 pilot regions located in northern non-inventoried boreal forests offers a potential solution to map forest biomass in areas where existing inventory information is sparse or non-existent. Investigations underway will further evaluate its applicability to map forest biomass across a broader range of biomass and forest types that are reflective of the diversity of northern boreal forest ecosystems, in support of reporting and monitoring the state of Canada's forests.

ACKNOWLEDGMENTS

We acknowledge support for this study provided by the Canadian Space Agency's Government Related Initiatives Program and the Earth Observation for Sustainable Development Project. We particularly thank the various provincial inventory services involved in the project who did provide GSP and other data sets: Ministère des ressources naturelles et de la faune du Québec, Direction des inventaires forestiers; Government of Newfoundland and Labrador, Department of Natural Resources; and finally the Government of Northwest Territories, Department of Environment and Natural Resources. We thank R. Skakun, E. Arsenault, P. Villemaire, M. Houle and D. Piercey for processing and analyzing the various data sets. Finally, we thank our field crew member: S. Dagnault, J.-P. Bérubé, P. Villemaire, S. Côté, G. Simard, G. Strickland and D. Sweetapple.

REFERENCES

.

Fazakas, Z., Nilsson, M. and Olsson, H., 1999. Regional forest biomass and wood volume estimation using satellite data and ancillary data, *Agricultural and Forest Meteorology*, Vol 98-99, pp. 417-425.

Guindon, L. Beaudoin, A., Leboeuf, A., Ung, C.-H., Luther, J.E., Côté, S. and Lambert, M.C., 2005. Regional mapping of canadian subarctic forest biomass using a scaling-up method combining QuickBird and Landsat imagery, *In Proc. FORESTSAT 2005 workshop*, 31 May – 1 June 2005, Borås, Sweden.

Hall, R.J., Skakun, R.S., Aresenault, E.J., and Case, B.S. 2006. Modeling forest stand structure attributes using Landsat ETM+ data: application to mapping of aboveground biomass and stand volume. *Forest Ecology and Management*, 225:378-390.

Lambert, M.-C., Ung, C.-H., and Raulier, F. 2005. Canadian national tree aboveground biomass equations. *Canadian Journal of Forest Research*, 35(8): 1996-2018.

Leboeuf, A., Beaudoin, A., Fournier, R., Guindon, L., Luther, J. and Lambert, M.C., 2007. A shadow fraction method for mapping biomass of northern boreal black spruce forests using QuickBird imagery, *Remote Sensing of Environment*, 110(4): 488-500.

Luther, J.E., Fournier, R.A., Piercey, D.E., Guindon, L., and Hall, R.J. 2006. Biomass mapping using forest type and structure derived from Landsat TM imagery. *International Journal of Applied Earth Observation and Geoinformation*, 8:173-187.

Olthof, I., Pouliot, D., Fernandes, R. and Latifovic, R. 2005. Landsat-7 ETM+ radiometric normalization comparison for northern mapping applications. *Remote sensing of Environment,* 95(3), pp. 388-398.

Reese, H., Nilsson, M., Sandström, P., and Olsson, H., 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Comp. Electr. Agric*., 37: 37–55.

Wulder, M.A., Dechka, J.A., Gillis, M.A., Luther, J.E., Hall, R.J., Beaudoin, A., and Franklin, S.E. 2003. Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program. *The Forestry Chronicle*, 79(6):1075-1083.