Mapping tree biomass of northern boreal forest using shadow fraction from Quickbird imagery

A. Leboeuf[1], A. Beaudoin[2], R.A. Fournier[1] L. Guindon[2], J.E. Luther[3] & M.-C. Lambert[2]

[1] Département de Géomatique Appliquée, Université de Sherbrooke, Canada
e-mail: antoine.leboeuf@usherbrooke.ca
[2] Canadian Forest Service - Laurentian Forestry Centre
[3] Canadian Forest Service - Atlantic Forestry Centre

ABSTRACT

Forest mapping from satellite remote sensing images is a convenient approach for regions with limited or absent forest inventories. We developed and tested a method to map above-ground biomass of black spruce (*Picea mariana)* stands in northeastern boreal regions of Canada using high resolution satellite images. Development of the method involved: 1) calculating shadow fraction (SF) using either classification (pixel-based) or segmentation and classification (object-based) algorithms, (ii) generating linear regression relationships between SF and biomass from ground sample plots using several combinations of method parameters towards defining the best options, (iii) calculating a global linear regression applicable for all sites using the best options, and (iv) mapping biomass as a grid layer for each site using the global regression. The linear relationships were calibrated using biomass estimates of 108 ground sample plots and the shadow fraction of tree crowns calculated from QuickBird images representing three test sites. The global regression relationship produced R^2, RMSE and bias in the range of 85 to 88% (except one case at 44%), 14 to 18 t/ha and -3 to 8 t/ha, respectively. The results suggest that the method may be an efficient means of mapping biomass of black spruce stands in northern Canada.[1]

Keywords: Biomass, boreal forest, black spruce, QuickBird, shadow fraction.

1 INTRODUCTION

The increasing importance of modelling carbon balance and monitoring of changing climate, combined with the economical interest to increase forest logging at the northern margin of the merchantable forests all require improved maps of forest attributes such as total volume and above-ground tree biomass. The remoteness and lack of infrastructure in these regions has resulted in a very limited forest inventory, which means that only a very small number of ground sample plots (GSP) and stand maps are available. Yet, the boreal and subarctic regions represent 60% of Canadian territory for a total of 4M km^2. This context requires the development and use of satellite remote sensing and modelling methods to estimate forest biomass, defined here as the total dry weight of above-ground trees per unit area (t/ha).

Segmentation of high-resolution satellite images (HRSI) has been used successfully to assess timber volume [1] and to identify land cover units [2]. Seed and King [3] found statistical relationships between tree shadow fraction (SF) or shadow brightness and leaf area index (LAI). A significant relationship between SF and different forest attributes such as biomass and LAI can also be calculated from a spectral mixture analysis method ([4] and [5]). These results suggest that tree shadow area, an obvious image attribute in HRSI imagery of northern open coniferous forests, can be identified through a combination of an object-based segmentation and a spectral classification to calculate tree shadow fraction, which in turn may be used to estimate biomass through functional relationships.

In this paper, we report on the development of a biomass mapping method based on SF of QuickBird (QB) HRSI. Application of the method uses a global relationship between biomass estimated at GSP and SF from QB images to map biomass of northern boreal black spruce stands.

Paper No. 57, Proceedings of the 26th Canadian Symposium on Remote Sensing, Wolfville, N.S., June 14-16, 2005

2 TEST SITES AND DATA

Three test sites were selected in Northeastern Canada close to the towns of Chibougamau (CH) and Radisson (RA) in Quebec, and Wabush (WA) in Labrador (Fig. 1). These sites were representative of forest conditions found in northeastern boreal forest: i) various density stands dominated by black spruce, ii) flat or gently rolling topography, and iii) understory composed of lichen, moss and shrub in various proportions.

Each test site was located within the extents of one of the three QuickBird images, which provided areal coverage ranging from 90 to 150 km^2. QB panchromatic (PAN) and multispectral (MS) images were acquired during mid-summer, under clear-sky conditions and with similar sun/sensor-viewing geometry (Table 1). The MS image (blue, green, red and near-infrared bands) with 2.4 m resolution was fused with the PAN image with 0.6m using the *Pansharpening* algorithm implemented in PCI Geomatica V8.2 [6] to create the pan-sharpened multi-spectral (PSMS) images. PAN and PSMS images were geometrically corrected with a first order polynomial using at least 12 ground control points with position accuracy within 5m, resulting in an average positional RMSE of 10m.

A total of 108 circular GSP (diameter = 22.56m; area=400m^2) were established across the three test sites using a stratified semi-random sampling design (31 in CH, 49 in RA, and 28 in WA). GSP centre points were localized using differential GPS with a precision better than 5m. GSP were selected with the criterion that black spruce and deciduous trees represented respectively more than 50% and less than 20% of the plot basal area. For each GSP, diameter at breast height (DBH) was measured for all trees (DBH > 5cm) and in a 4m sub-plot for seedlings (DBH < 5cm) following the procedures suggested by the Canadian National Forest Inventory [7]. Oven-dry biomass (kg) of each tree of the plots was calculated from DBH using allometric equations [8] and was summed to obtain plot-level biomass (t/ha). GSP biomass ranged from 4.8 t/ha to 163.0 t/ha with a mean value of 50.0 t/ha. Crown closure and height ranged respectively from 5 to 85% and 0 to 18m.

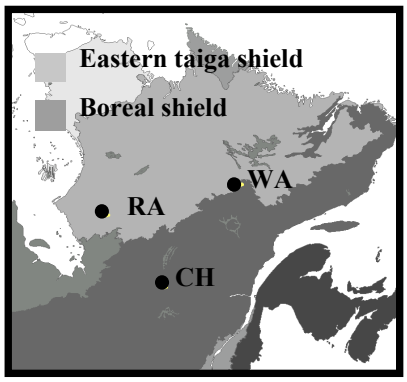

Figure 1: Location of the three test sites within two ecozones of Northeastern Canada

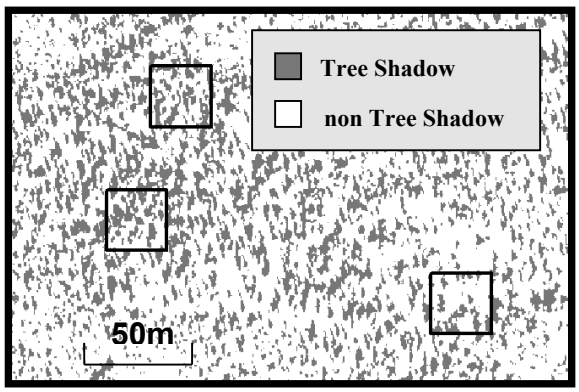

Figure 2: Reference squares used to estimate SF, overlaid on the tree shadow bitmap layer

Table 1. QB image acquisition parameters for each test site

	CH	RA	WA
Date, local time	07/10/2003, 10:43	08/12/2003, 10:57	08/10/2002, 10:35
Sun elevation / azimuth (°)	58.6 / 143.2	49.0 /152.9	51.1 / 157.7
Sensor elevation / azimuth (°)	78.5 / 96.4	76.7 / 172.3	80.4 / 225.6

3 METHOD

The development of a biomass mapping method based on SF was divided into four separate sets of procedures that: (1) calculated SF using either classification (pixel-based) or segmentation and classification (object-based) algorithms, (ii) generated, for each study site, a linear regression between SF and GSP biomass using several combinations of method parameters towards defining the best options, (iii) calculated a global linear regression applicable for all sites using the best options, and (iv) mapped biomass as a grid layer for each site using the global regression.

First, eCognition software [9] was applied to create spatial objects including tree shadows (TS). Five segmentation scale factors were tested: 0, 5, 10, 20, and 40, where a scale of 0 implies no segmentation (pixel-based) and an increasing scale factor produced objects of larger dimension. Binary objects (TS/ non-TS) were generated from the segmented HSRI by applying a threshold on the mean intensity value. The suitable threshold to generate TS was established visually by an analyst. Reference squares were then aligned with the center of all 108 GSPs and provided a reference area from which SF (shadow area / reference square area) was calculated using the TS bitmap layer (Fig. 2). Several reference square sizes were tested (10, 30, 60, and 90m). SF values were normalized to a common sun-terrain-sensor viewing geometry to account for differences in topography, as well as sun illumination and sensor viewing angles.

Second, linear regressions between normalized SF and GSP biomass were calculated for each of the three test sites. Several options were tested for a total of 20 possible combinations: 5 segmentation scales and 4 reference square sizes. The regressions were calculated using a 70% random selection of GSP, (GSP_{cal}). The remaining 30% of GSP (GSP_{val}) was used for validation. The most suited combination of option was identified as the one for which the regression maximised R_ value and minimized RMSE and bias values. A comparison of the R_ values obtained using respectively QB PAN and PSMS images was also done for the three test sites.

Third, a global linear regression was calculated from GSP_{cal} of the three sites. Several statistical tests were applied to evaluate if significant differences existed between the three local regressions and the global linear regression. The analysis of covariance was performed to assess if the local linear regressions were significantly different. The normality of residuals was tested with the Shapiro-Wilk test and the homogeneity of variance was tested visually using a graph of residuals. In addition, the Fisher test was used to evaluate the coincidence and parallelism hypothesis of the regression equations with a threshold of $\forall$=0.05 [10] using GSP_{cal} data set. Finally, overall RMSE and bias values were calculated using the pooled set of GSP_{val}.

Fourth, the global regression using the most suited options was used to map biomass over the extent of the three test sites. Post-classification was used to remove false TS such as water bodies, which can have similar intensity values. Then, for each cell of a grid layer (size given by the most suited reference area), we calculated SF, normalized it and estimated biomass using the global regression equation. We finally assessed the overall precision of the three maps using the pooled GSP_{val} set. The resulting map was a grid layer where each cell, except water bodies, had a biomass value.

4 RESULTS AND DISCUSSION

Similar R_ results were obtained using either QB PAN or QB PSMS images and QB PAN images were used and are presented in the following steps. In terms of method parameters, the best results were found with the smaller scale factors (0 to 10) and reference square sizes (10 to 30m) (Table 2a, b and c). Such reference square size corresponds with the approximate size of the GSP. Considering practical issues, a scale factor of 0 (pixel-based approach) and a reference square size of 30m (Landsat resolution) were selected as the most suited options for input parameters. The three site-specific linear regressions (Bio = a + b*SF) using these input parameters are shown in Fig. 3a (dotted lines) with the related statistics reported in Table 3. High R^2 values were obtained (85-88%) except for the WA site (44%). The poorer result for the WA site was partly explained by the predominance of GSP with low biomass values, the uneven-aged forest stand structure, and the more rugged relief resulting in higher GSP-HRSI co-registration errors and additional variance of the SF values. Error statistics were similar between the three sites: RMSE from 14 to 18 t/ha and all absolute bias values were below 7.5 t/ha.

All statistical tests confirmed that the three linear regressions were similar to the global regression with a p-value=0.41< F=0.89 for the slope and p-value=0.12 < F=2.14 for the intercept. The global linear regression using a pooled dataset of all GSP_{cal} (Fig. 3a) resulted in a RMSE value of 15.29 t/ha with a relative error around 30% considering a mean biomass of 50 t/ha, and a low bias of 4.18 t/ha. Biomass estimated from the global regression corresponded well with biomass measured at GSP_{val} for the full range of conditions sampled (Fig. 3b).

Error assessment of the mapped biomass for the three test sites gave an overall RMSE of 9.97 t/ha and a low bias of 4.66 t/ha. Despite the low number of GSP used for validation, the mapped biomass was spatially consistent: high biomass in valleys where the soil is richer, low biomass in wetlands or on dry hill tops exposed to strong winds (Fig. 4).

Several factors affecting the accuracy of the biomass map were explored. First, the sensitivity to the image threshold was tested to assess if significant changes depended on the visual interpretation of an operator. We found that a large range of threshold values could be applied without changing significantly the regression statistics, thus leaving comfortable margins for the manual thresholding step. We also tested if separating the calculations for the two main understory types would improve the regressions. Results did not show an improvement so no separation of data by understory type was necessary. However, the normalization for sun-terrain-sensor viewing geometry improved the results which helped to generalize the

Table 2 A, B and C. Statistics from the linear regression (R^2) and the comparison with GCP_{val} (RMSE and bias) for the potential values of scale factor and reference square sizes tested for the CH, RA and WA test sites.

CHIBOUGAMAU (31 GSP)

A

Reference square size	Scale factor 0	5	10	20	40
10	**0,68** *(18,2)* [-0,8]	**0,72** *(20,3)* [11,3]	**0,71** *(21,7)* [13,2]	**0,67** *(21,8)* [3,5]	**0,36** *(45,1)* [-9,3]
30	**0,72** *(18,3)* [3,2]	**0,73** *(26,8)* [19,2]	**0,72** *(28,1)* [19,5]	**0,73** *(28,4)* [17,4]	**0,48** *(35,2)* [21,2]
60	**0,73** *(31,9)* [24,5]	**0,72** *(32,0)* [24,5]	**0,72** *(32,4)* [25,0]	**0,66** *(35,4)* [25,4]	**0,53** *(36,9)* [8,4]
90	**0,73** *(30,6)* [21,2]	**0,67** *(30,9)* [21,6]	**0,67** *(31,2)* [22,0]	**0,63** *(33,3)* [22,5]	**0,54** *(35,9)* [25,6]

RADISSON (49 GSP)

B

Reference square size	Scale factor 0	5	10	20	40
10	**0,69** *(14,5)* [3,3]	**0,69** *(10,0)* [-3,3]	**0,66** *(9,8)* [-0,8]	**0,65** *(12,0)* [-3,5]	**0,57** *(13,3)* [-1,6]
30	**0,75** *(14,2)* [4,2]	**0,75** *(10,6)* [-4,4]	**0,73** *(10,6)* [-3,9]	**0,74** *(11,8)* [-3,8]	**0,70** *(11,7)* [-1,8]
60	**0,69** *(10,0)* [-5,0]	**0,70** *(9,3)* [-4,4]	**0,70** *(9,7)* [-4,5]	**0,73** *(9,9)* [-4,0]	**0,66** *(10,5)* [-0,7]
90	**0,69** *(11,3)* [-6,2]	**0,67** *(11,3)* [-6,1]	**0,67** *(10,9)* [-5,4]	**0,70** *(10,6)* [-4,1]	**0,66** *(10,3)* [-2,0]

WABUSH (28 GSP)

C

Reference square size	Scale factor 0	5	10	20	40
10	**0,42** *(9,7)* [4,1]	**0,44** *(8,7)* [1,8]	**0,35** *(10,8)* [0,0]	**0,24** *(10,1)* [1,6]	**0,27** *(12,5)* [-1,1]
30	**0,42** *(10,4)* [0,6]	**0,42** *(11,1)* [-2,2]	**0,42** *(11,4)* [-2,0]	**0,37** *(9,2)* [-2,6]	**0,27** *(15,9)* [-3,6]
60	**0,40** *(11,1)* [-2,5]	**0,40** *(11,4)* [-2,5]	**0,40** *(11,0)* [-2,3]	**0,41** *(10,2)* [-2,2]	**0,27** *13,5)* [-2,9]
90	**0,40** *(12,3)* [-1,7]	**0,39** *(12,2)* [-1,7]	**0,37** *(12,1)* [-1,9]	**0,40** *(12,1)* [-2,0]	**0,30** *(13,8)* [-2,2]

Legend: **R^2** ; *(RMSE)* ; [Bias]

Table 3. Local and global linear regressions with intersect A and slope B, and RMSE/bias statistics from GSP_{cal} and GSP_{val}

GSP set	Nb GSP_{cal}	A	B	Adj. R_	Nb GSP_{val}	RMSE	BIAS
CH	22	1.94	211.89	88 %	9	17.30	-2.50
RA	35	6.57	238.92	85 %	14	14.46	6.47
WA	19	16.19	177.40	44 %	9	14.39	7.31
Pooled	76	7.36	215.13	82 %	32	15.29	4.18

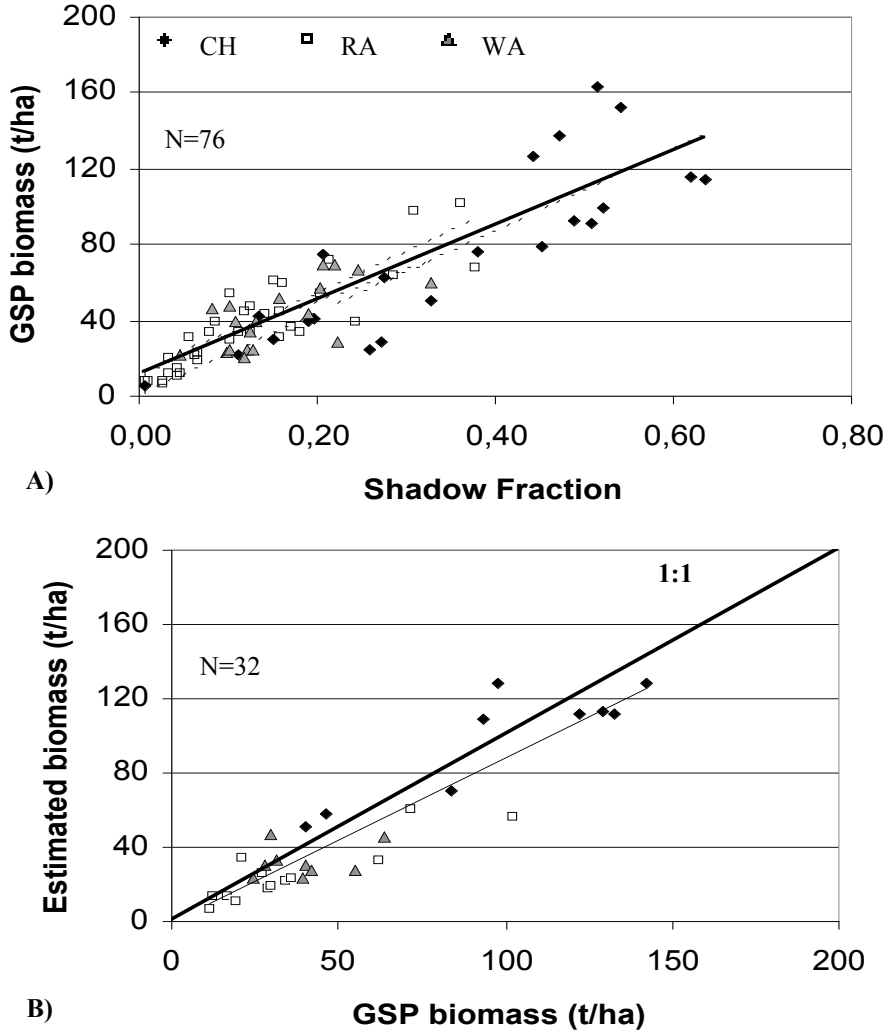

Figure 3. A) Linear regressions calibrated using 76 GSP_{cal} (dotted lines: site-specific regressions; thick line: global regression); B) estimated vs GSP biomass using remaining 32 $GSP_{val.}$

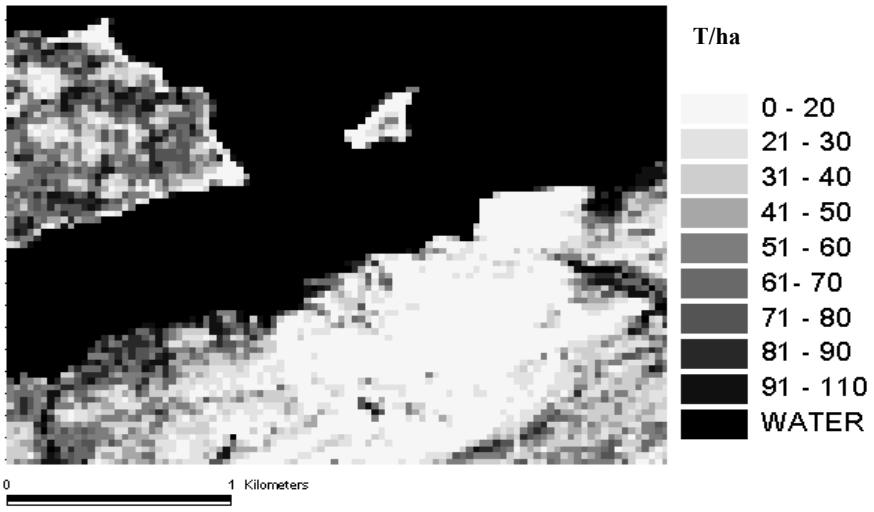

Figure 4. Zoom of the biomass map derived from shadow fraction for RA test-site.

method across the three test sites, especially for the sun elevation angle affecting the length of tree shadows. Another factor, the co-localization errors between the GSP and QB image, could affect significantly the results. Several random shifts of the GSP center position around the real position were applied and the results from the shifted GSP showed similar variance to the original analysis. Finally, an Ikonos image (PAN band with 1m resolution) was tested over an overlapping portion of the WA test site. All the procedures of the method were applied and led to results comparable to those obtained with QB but with weaker fits for the regression and higher RMSE and bias.

5 CONCLUSIONS AND PERSPECTIVES

We developed and tested a biomass mapping method based on SF for black spruce stands which are dominant in the northeastern subarctic forests of Canada. The optimal results were obtained when a simple visual threshold was applied to the plain HRSI PAN with no segmentation (scale factor = 0) and with a reference square area of 30m to calculate the SF. A global regression used SF as an independent variable to calculate biomass with a R^2 = 82%, RMSE = 15.3 t/ha, and bias = 4.2 t/ha. The method provided consistent results for black spruce stands with various understory types, for a GSP biomass range from 0 to 170 t/ha, and for stand density and height up to 85% and 18m respectively. Future work will assess the validity domain of the method: (i) over a larger range of stand conditions (species composition, biomass range, understory type) and relief situations; (ii) using different HRSI types (such as Ikonos with 1m resolution) and related acquisition parameters; and (iii) for estimating other forest attributes of interest (e.g. stem volume, basal area, height and crown closure). Finally, the method is convenient to map above-ground tree biomass and to supply satellite sample plots (SSP) that can be used as surrogates to GSP. These two outcomes provide a base from which biomass values can be scaled up at the regional scale [11].

ACKNOWLEDGMENTS

We acknowledge support for the establishment of the test sites and GSP provided by the Canadian Forest Service Ecoleap Project. Support for the study was also provided by the Earth Observation for Sustainable Development Project (EOSD) of the Canadian Space Agency's Government Related Initiatives Program. We thank Chhun-Huor Ung for his advice related to allometry and sampling design and Ron Hall for insightful discussions. We also thank Philippe Villemaire for his support in GIS and image processing and our field crew members: L. Guindon, S. Dagnault, J.-P. Bérubé, P. Villemaire, S. Côté, J. Donnelly and G. Strickland who collected ground truth data in challenging conditions.

REFERENCES

[1] PEKKARINEN, A., 2002: Image segment-based spectral features in the estimation of timber volume. *Rem. Sens.Environ.* 82, pp. 349-359.

[2] LOBO, A., 1997: Image segmentation and discriminant analysis for the identification of land cover units in ecology. *IEEE. Trans. Geosc. Rem. Sens.* 35, pp.1136-1145.

[3] SEED, E.D. & KING, D.J., 2003: Shadow brightness and shadow fraction relations with effective leaf area index: importance of canopy closure and view angle in mixedwood boreal forest. *Can. J. Rem. Sens.* 29, pp. 324-335.

[4] PEDDLE, D.R. & JOHNSON, R.L., 2000: Spectral mixture analysis of airborne remote sensing imagery for improved prediction of leaf area index in mountainous terrain, Kananaskis Alberta. *Can. J. Rem. Sens.* 26, pp.177-188.

[5] PEDDLE, D.R., BRUNKE, S.P. & HALL, F.G., 2001: A comparison of spectral mixture analysis and ten vegetation indices for estimating boreal forest biophysical information from airborne data. *Can. J. Rem. Sens.* 27, 6, pp. 627-635.

[6] ZHANG, Y., 2002: A new automatic approach for effectively fusing Landsat-7 as well as Ikonos images. *Proc. Of IGARSS/24th CRSS symposium*, Toronto, 4 p. ROWE, J.S., 1972: Forest Regions of Canada. *Can. For. Serv., Depart. Environ.*, Publication No. 1300.

[7] GILLIS, M.D. 2001: Canada's national forest inventory (responding to current information needs). *Environ. Monit. Assess.* 67, pp. 121-129.

[8] OUELLET, D., 1983: Biomass equations for black spruce in Quebec. Environment Canada, Can. For. Serv., Laurentian Forestry Research Centre, Info. Report. 11p.

[9] DEFINIENS INC., 2002: eCognition, version 3.1.

[10] MILTON, J.S. & ARNOLD, J.C., 2003: Introduction to probability and statistics: Principles and applications for engineering and the computing sciences. 4th edition.

[11] GUINDON, L. BEAUDOIN, A., LEBOEUF, A., UNG, C.-H-., LUTHER, J.E., CÔTÉ, S. & 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.