Can. J. Remote Sensing, Vol. 29, No. 5, pp. 536–543, 2003

Forest inventory height update through the integration of lidar data with segmented Landsat imagery

Michael A. Wulder and David Seemann

Abstract. Estimates of stand height are an integral component of forest inventories. Lidar has been demonstrated as a tool for remotely sensing information on the vertical structure of forests, such as height. The ability to remotely sense height information for forest inventory purposes may allow for procedures such as update, audit, calibration, and validation. With current technologies, lidar data collection and processing are a resource-intensive undertaking. The ability to use a regression model to spatially extend a lidar survey from a sample to a larger area would act to decrease costs while allowing for the characterization of a larger area. In this study we address the ability to extend lidar estimates of height from sample flight lines to a greater area using segmented Landsat-5 thematic mapper (TM) data. Based upon empirical relationships between lidar-estimated height and within segment digital numbers, height is estimated for an entire landscape from a 0.48% sample. To conform to current polygon-based forest management practices, the within polygon segment-based height estimates are combined to create an updated height attribute for each polygon. The empirical relationships between lidar data and forest inventory polygon attributes (coefficient of determination (r^2) = 0.23; standard error (SE) = 4.15) and within polygon spectral values (r^2 = 0.26; SE = 4.06) indicated a need to develop more representative models. To this end, we developed a regression model to produce a relationship between quantile-based estimates of mean canopy top height for the segments with lidar hits (r^2 = 0.61; SE = 3.15). This segment/height empirical relationship allowed us to extend the height estimates to polygons that have no lidar information using the image digital numbers. The segment/lidar estimates of height generally form a range centered on zero (no difference) to ±6 m of the ground measured height for over 80% of the available validation plots, with a r^2 of 0.67 and a SE of 3.30 m.

Résumé. Les estimations de hauteur de peuplement constituent une partie intégrale des inventaires forestiers. Le lidar a démontré son utilité comme outil pour la télédétection d'information sur la structure verticale des forêts comme la hauteur. La capacité de détecter à distance une information sur la hauteur dans le contexte des inventaires forestiers peut permettre la mise au point de procédures comme la mise à jour, l'évaluation, l'étalonnage et la validation. Avec les technologies actuelles, l'acquisition et le traitement des données lidar constituent une tâche considérable en termes de ressources. La possibilité d'utiliser un modèle de régression pour étendre spatialement un relevé lidar d'un échantillonnage donné en fonction d'une surface plus grande permet de diminuer les coûts tout en permettant la caractérisation d'une surface plus grande. Dans cette recherche, nous examinons la possibilité d'étendre les estimations de hauteur lidar acquises le long de lignes de vol échantillons en fonction d'une surface plus grande à l'aide de données Landsat-5 segmentées. Se basant sur les relations empiriques entre la hauteur estimée par lidar et les valeurs des comptes numériques à l'intérieur des segments, la hauteur est estimée pour un paysage en entier à partir d'un échantillonnage de 0,48 %. Conformément aux pratiques actuelles en matière de gestion forestière axées sur les polygones, les estimations de hauteur basées sur un segment à l'intérieur d'un polygone sont combinées pour créer un attribut de hauteur mis à jour pour chaque polygone. Les relations empiriques entre les données lidar et les attributs des polygones d'inventaire forestier (r^2 = 0,23; écart type (ET) = 4,15) et les valeurs spectrales à l'intérieur des polygones (r^2 = 0,26; ET = 4,06) ont mis en évidence le besoin de développer des modèles plus représentatifs. À cette fin, nous avons développé un modèle de régression pour produire une relation entre les estimations basées sur les quantiles de hauteur moyenne au sommet du couvert pour les segments avec les échos lidar (r^2 = 0,61; ET = 3.15). Cette relation empirique segment/hauteur nous a permis d'étendre les estimations de hauteur aux polygones qui n'ont pas d'information lidar à l'aide des valeurs des comptes numériques de l'image. Les estimations de hauteur segment/lidar sont généralement regroupées autour du zéro (aucune différence) jusqu'à ±6 m de la hauteur mesurée au sol pour plus de 80 % des parcelles de validation disponibles, avec une valeur de r^2 de 0,67 et de ET de 3,30 m.
[Traduit par la Rédaction]

Introduction

Height estimation is an expensive and time-consuming process with current forest inventory methods based upon air photo interpretation. Light detection and ranging (lidar) data may be processed to provide accurate estimates of tree or canopy heights (Lim et al., 2002; Lefsky et al., 2002). Yet, lidar

Received 12 September 2002. Accepted 1 May 2003.

M.A. Wulder[1] and D. Seemann. Pacific Forestry Centre, Canadian Forest Service, Natural Resources Canada, 506 West Burnside Road, Victoria, BC V8Z 1M5, Canada.

[1]Corresponding author (e-mail: mwulder@nrcan.gc.ca).

information can be costly. Approaches that apply a hierarchical sampling strategy to the use of lidar in combination with forest inventory data may allow for a relatively cost effective and accurate means of characterization heights over large areas.

Optical remotely sensed data represent well the vegetative content and horizontal structure of forests, but do not capture well the vertical elements of forest structure (Wulder, 1998a). Lidar data provide rich information on the vertical structure of forests (Lefsky et al., 2001; 2002), but generally over small areas only and often at a high cost. The merging of these two unique data-types provides opportunities to capitalize upon the unique characteristics of both. While there are examples of lidar data integration with high spatial resolution data (Blackburn, 2002; Gougeon et al., 2001), few studies have integrated lidar with Landsat imagery. Several methods for estimating heights through the integration of lidar and Landsat-7 enhanced thematic mapper plus (ETM+) data are explored in Hudak et al. (2002). Forest inventory update is one application that could be assisted by a successful integration of lidar with Landsat data. Forest inventories are generally updated on a 10-year cycle in Canada (Gillis and Leckie, 1993) with procedures varying by jurisdiction (Leckie and Gillis, 1995). In general, aerial photographs are obtained, manually interpreted, and stored in a geographical information system (GIS) for analysis, mapping, and information retrieval. Within this inventory cycle, lidar and multispectral data may be used to verify the current conditions with those represented in forest inventory GIS databases and to update attributes such as height.

Building relationships between forest inventory, multispectral, and lidar data is a complex proposition. Forest inventory data are a vector data-type with multiple categorical attributes per spatial unit compiled from a range of data collection dates. Forest inventory spatial units are generally larger than a minimum size of 3 ha for forested polygons, with the minimum mappable unit varying by jurisdiction. Landsat multispectral data are raster in nature, commonly with integer digital numbers (DNs) representing the surface conditions at a given location. The Landsat spatial resolution of 30 m × 30 m is the equivalent of 0.09 ha. Image data provide large area coverage representing a single date. Lidar data, depending on the sensor used, represent the interception of a signal in the vertical plane. Canopy heights are generally computed from the difference between the lidar-derived canopy surface and ground surface. The spatial resolution of lidar data also varies by sensor used and may range from a footprint of several decimetres to 10 m (Lim et al., 2002). As a result, to integrate these data, one must consider several differing scales of information and a range of data-types. Further, the calibration and validation stages of the integration projects often utilize point data from field-based measurements — an additional data-type, representing a differing scale. An additional consideration is the actual information content represented by each data-type. From a sensor perspective, the differing spatial, spectral, temporal, and radiometric characteristics combine to produce a unique measurement of the phenomena of interest. The selection of remotely sensed data-type with desired information content is required to ensure that the data to be integrated are compatible (Lefsky and Cohen, 2003).

Forest inventory polygons are generally developed from the manual interpretation of aerial photographs. The decision rules used to delineate the forest stand units are based upon a desire to group a landscape into units with similar characteristics, such as age, height, species, and crown closure. To capture the conditions that are common outside of plantation forestry, a range of stand conditions is often acceptable within a particular polygon. The delineation into polygons is also influenced by nonvisible elements, such as management units and map sheet boundaries.

Polygon decomposition is an approach that enables the boundaries of one spatial data source to form the basis for the extraction of another spatial data source. For instance, polygon boundaries may be used in this context for the extraction of DN from co-georegistered image data. The spectral values found within a particular polygon are unique to that polygon and may be used to develop empirical relationships (Wulder and Franklin, 2001), among other applications. The extraction of all DN within a polygon allows for summarization of the DN to indicate conditions present within each polygon. Polygon decomposition can be applied to extract the class given to each pixel or the actual DN values. In this study, the actual DN values are extracted. When extracting all DN values from within a polygon, a summarization of the DN may be placed back into the GIS using the polygon identification number as the relational field. The remotely sensed data, with a higher spatial resolution and frequent revisit rate, enable unique information to be developed for each forest stand polygon. Data integration attempts have been made to utilize within polygon spectral information to estimate forest attributes (Magnussen et al., 2000; Wulder, 1998b). When combining forest inventory and satellite imagery, the spectral contents found within the co-registered polygons are often poorly related to the polygon attributes. For instance, Wulder et al. (1999) found, when using polygon cover-type information from forest inventory data as the dependent variable in an analysis of variance study with within polygon DN that the within class variability was similar to the between class variability. The weak relationship between the polygon cover-types and associated DN precluded the creation of predictive models. To enable the maintenance of forest inventory polygons as a management unit, homogeneous units within each polygon may be generated. Wulder (1998b) improved leaf area index (LAI) estimates based upon within polygon DN through stratification of the polygon with the classes generated from an image classification. Similarly, an image segmentation procedure may be applied to group the DN into homogeneous units suitable as a within polygon stratifier. Image segmentation is a means by which remotely sensed data can be placed into meaningful groups based upon both spatial and spectral characteristics (Pal and Pal, 1993). DNs extracted to represent the interior of a polygon often do not capture a unique spectral signature representative of a particular suite of attributes. As a result, the development of empirical equations from within polygon spectral values is problematic. Conversely, within segment DNs are, by nature of the segmentation process, related to one another, allowing for the improved development of empirical equations. In this study, polygon decomposition allows the development of relationships between within polygon and within segment DN to the co-georegistered lidar height estimates. Polygon decomposition also allows segment-based estimates of height to be prorated, based upon the area represented in that polygon by each segment, back into the forest inventory polygons as a new attribute.

The scanning lidar imager of canopies by echo recovery (SLICER) records data on canopy height, vertical structure, and ground elevation. In this research we address the ability to extend SLICER estimates of height from sample flight lines to a greater area using Landsat-5 TM data. The hypothesis of this study is that height estimation for an entire inventory area, based upon a subsample of lidar data, is enabled through the integration of forest inventory data, segmented Landsat imagery, and a sample of lidar data. The relationships between lidar-estimated heights from categorical polygon attributes and DNs representing individual polygons provide a benchmark to compare to heights estimated from within polygon segment-based DNs. We illustrate that segmented remotely sensed data may be used to spatially extend SLICER lidar height estimates from a sample to represent a larger area and be made available in a traditional forest management environment.

Methods

The methodological steps of this research broadly are data collection and preparation, image processing, lidar processing, height estimation, height extension, and validation. These research elements are described in detail in the following subsections. From the Boreal Ecosystem–Atmosphere Study (BOREAS) (Sellers et al., 1995), we have access to lidar (SLICER), Landsat-5 TM imagery, GIS forest inventory, and plot data.

Study area

The Saskatchewan study area was selected to represent a variety of boreal forest species types and forest conditions. The study area is located near the Southern Study Area established as a component of BOREAS (Sellers et al., 1995). The study area is largely within the boreal plains ecoregion with a tree species transition from deciduous-dominated mixed woods in the southern portion to coniferous-dominated mixed woods in the northern portion (Rowe, 1977; Lowe et al., 1996). The study area is classified as mixed boreal forest consisting of mixed woods composed of aspen and white spruce, which are common where the sites are well drained; whereas, jack pine (*Pinus banksiana* Lamb.) and black spruce (*Picea mariana* (Mill.) BSP) are found with pure stands of jack pine on dry sites composed of coarse textured soils. In the poorly drained areas throughout the pilot region, bogs support black spruce and small proportions of tamarack (*Larix laricina* (Du Roi) K. Koch). Also present are fen areas, which are composed mostly of sedge vegetation with discontinuous cover of tree species such as tamarack. Forest disturbance is largely the result of localized logging operations and fire. Recent fires have generally been limited in areal extent and frequency through a comprehensive forest fire suppression program (Sellers et al., 1995). The study area is approximately 115 km (east–west) by 65 km (north–south) (**Figure 1**), for a total area of 7073 km^2. The topography of the study area is of gentle relief, with elevations ranging from 400 to 700 m. The dominant landforms consist of glacial till plains, and rolling or hilly moraines.

Field data

As a component of BOREAS, mensuration data were collected during the summer of 1994. Published data describe in detail the plot locations (Halliwell and Apps, 1997a), measurement methods, and overstory and understory characteristics (Halliwell and Apps, 1997b). The field site selection was based upon homogeneity of species composition and age, with a minimum stand size of 100 m × 100 m (Halliwell and Apps, 1997a). Within the confines of the study region, sample plot data were available for validation purposes; a total of 93 points were available, with 41 plots located within the edge buffered polygons, and 49 plots located within the edge buffered segments (**Table 1**). In some cases, plot measurements made to represent a particular species composition and structure, as determined in the field, were located in more than one forest inventory polygon or segment when plotted spatially (**Table 1**). The number of plots located within forest inventory polygons and segments varied owing to the size and shape of the units. As will be described in a later subsection, the polygons and segments were buffered to reduce edge effects. Plots falling within the buffered areas were removed. The within polygon and within segment plot values were utilized for model validation of the lidar-based height estimates.

Forest inventory data (GIS)

The forest inventory system in Saskatchewan is based on interpretation and digitization of air photos on an approximate 15-year completion cycle (Gillis and Leckie, 1993). Inventory validation is undertaken through field visits and the establishment of temporary sample plots. As a function of the on-going inventory update, the forest inventory data provided for this study are of variable vintage, with 82.7% of the inventory compiled in 1984, 3.8% compiled before 1984, and 13.5% after 1984. From these variable vintages, the entire inventory was processed by Saskatchewan Parks and Natural Resources to represent 1994 conditions. The GIS data provide the forest inventory polygon context within which we consider the consistency lidar observations from differing lidar data collection flight lines.

SLICER lidar data

The SLICER was developed at the NASA Goddard Space Flight Center as a scanning modification of a profiling laser

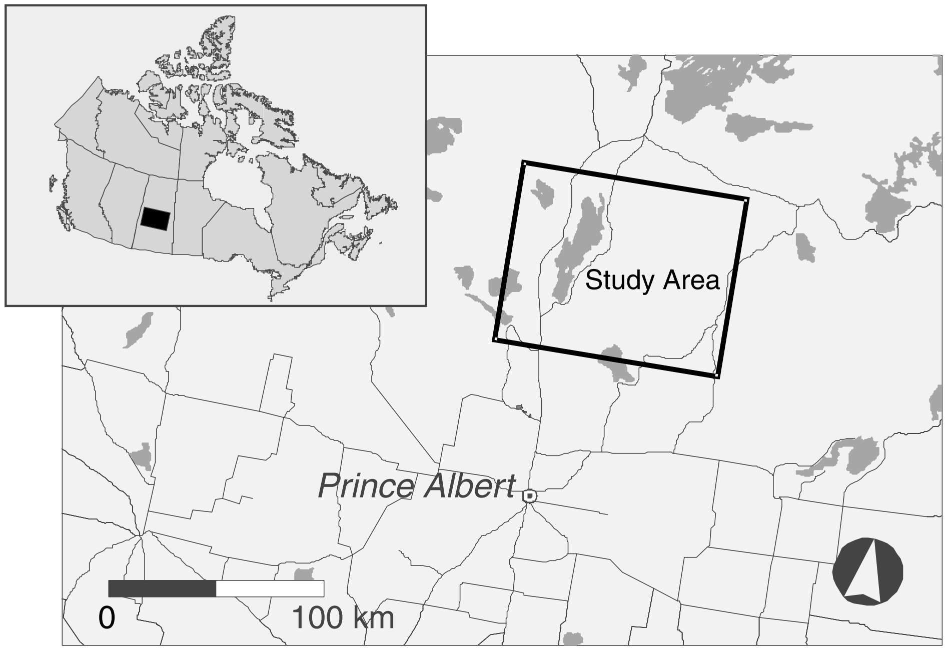

Figure 1. Location of study area in central Saskatchewan, Canada.

Table 1. Sample and population (landscape) data summary.

	Samples		Landscape	
Location	Plots	Lidar	Polygons	Segments
Study area	93 (over entire study area)	66 260 lidar hits	79 070	50 717
Forest inventory polygons	41 (geolocated within buffered forest inventory polygons)	In 2 427 polygons		
Image segments	49 (geolocated within image segments)	In 2 760 segments		

Note: Landscape values represent the entire study area; sample values represent the locations noted. The sample plot data are used as a validation data source. The sample values do not sum to 93 as not all points are eligible for inclusion. See text for more details.

altimeter (Blair et al., 1994). The SLICER is a lidar system that digitizes the backscattered return signal resulting in the capture of a full waveform representing the vertical distribution of illuminated surfaces within the laser footprint. In this study the footprint diameter was approximately 9 m, varying by approximately ±5% owing to laser divergence and changes in the distance from the aircraft to the ground.

The SLICER data utilized in this study was collected in July 1996. The lidar data were processed from the raw data into variables representing key components of the sensed waveform (Harding, 2000). Of the available SLICER variables, “ground start” was used in this study. “Ground start” is the distance between detected laser returns from the canopy top and underlying ground, hereafter called height. Details describing the determination of SLICER variables, such as the “ground start” from the full waveform data, can be found in Harding (2000). The SLICER data cover 33.9 km^2 of the study area (or 0.48% of the total study area). The lidar data were collected to capture the conditions at a subset of the BOREAS project calibration and validation sites described in the preceding field data section. The lidar also collected data while in transit to and from these sites, producing a stratified random sample of the landscape.

Image acquisition and processing

Landsat-5 TM image, path 37, row 22, July 1994, was georectified to 30 m × 30 m pixel size using a first-order polynomial rectification, resulting in an RMS error of 0.80. A top-of-atmosphere (TOA) reflectance correction based upon the procedures by Markham and Barker (1986) was applied to ETM+ bands 1, 2, 3, 4, 5, and 7. The TOA procedure corrects for variations in solar illumination, atmospheric transmission, and path radiance, and assumes a uniform atmosphere within the image. The image is transformed from raw DNs to TOA reflectance values using image calibration values, radiometric ancillary information, solar zenith angle, and Earth–Sun distance measurements.

Lidar height estimation

An initial screening of the lidar data was undertaken to ensure data quality. The data screening included ensuring that the height values were not erroneously high (related to the noise value being surpassed above the forest canopy, triggering a start to data recording). Height estimates were made following Wulder et al. (2001), which is a large footprint lidar implementation of the grid-based procedure of Næsset (1996) undertaken with data from a small footprint, discrete return lidar. Both the polygon and segment height estimates are based upon the same procedure. To reduce bias and to account for stand openness and structural variability, a 40-m grid was constructed and placed over the GIS and lidar data. In each polygon (or segment), the two highest hits within each grid cell were averaged. The selection of the two highest values within the 40-m grid allows for the representation of the highest values, or the upper canopy strata, to be included in the estimation of the polygon height. The grid allows the within polygon characteristics to be captured. All of the within polygon (or segment) grid cell averages were then averaged to provide for a height estimate. The averaging of the grid cell values provide an estimate representative of canopy top height.

Image segmentation and empirical models

The imagery was segmented using spectral and spatial characteristics to enable a stronger relationship between the polygon units and the spectral contents. Landsat-5 TM channels 1 to 5 and 7 were input for segmentation using eCognition (Definiens Imaging GmbH, 2000). eCognition uses both image spatial and spectral information to create objects from the image data. The proprietary segmentation algorithm used in eCognition is purported to be based upon local homogeneity and subsequent grouping of pixels into objects of similar characteristics. The characteristics included in the segmentation are spatial, spectral, textural, and areal (i.e., size and shape). The heuristic optimization procedure maximizes homogeneity of image objects, over the entire scene, based upon the standard deviation of image object characteristics, including shape. Parameters may be set to weight the impact of a given characteristic upon the segmentation, for instance, if the scale parameter is larger, the image objects will be larger. In our case, we used the following settings for scale (10), color (0.7), shape (0.3), smoothness (0.5), and compactness (0.5). These objects were converted to polygons to allow polygon decomposition (Wulder and Franklin, 2001) approaches to be applied. Each segment was assigned a unique identification number (PID) that allowed for the relating of the segments to the lidar, image data, and forest inventory data. The tagging of each segment with a PID enabled the development of relationships between the within segment spectral values and the lidar data (**Figure 2**). Regression models were developed to empirically relate the lidar-estimated heights to the polygon attributes (as a benchmark), within polygon DNs, and within segment DNs. The Landsat-5 TM optical channels of 1 to 5, and 7 were used as the independent variables in the spectral-based empirical models. A single pixel buffer was also removed from the edge of the polygons and segments to further aid in the development of representative spectral signatures. Additionally, a minimum of five 40 m × 40 m boxes were required to be covered by the SLICER data for inclusion in the model building. The characteristics of the forest inventory polygons (used in the regression extension of height estimates) and the image segments varied (**Table 2**). The median area of the segments was larger than that of the forested polygons of the forest inventory.

Results and discussion

Height estimates from forest inventory categorical variables

The forest inventory GIS data provide complete coverage of the study area. As a result, the forest inventory data may provide data that will allow the extension of the lidar sample.

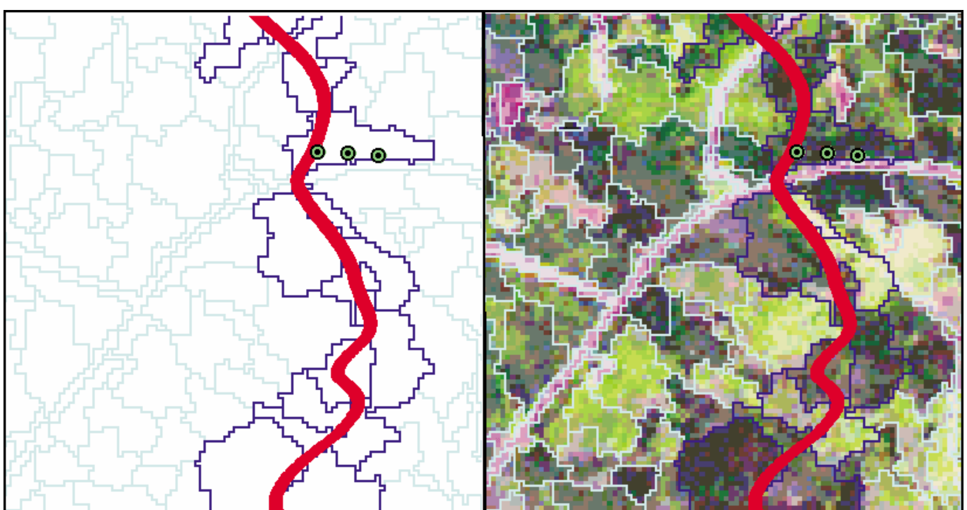

Figure 2. Illustration of segments with lidar data (dark blue), segments with no lidar data (light blue), a SLICER flight line (red), and some field plots (green/black circles). The two panels show the same segment information, but the right panel has the Landsat-5 TM (red, green, infrared) data as a backdrop to illustrate the segmentation results.

Table 2. Summary statistics of forest inventory polygons (containing forest attributes) and image segments.

Characteristic	Forest cover polygons	eCognition segments
Total no.	44 144	50 718
Mean area (ha)	8.18	14.23
Median area (ha)	4.14	11.07
Min. area (ha)	1.00	1.00
Max. area (ha)	1 103.13	194.22
SD area (ha)	17.16	11.95

Note: Forest inventory polygon statistics represent the forested land-base. SD, standard deviation.

Each polygonal spatial unit has attributes that relate the forest characteristics found within. The forest characteristics are related using categorical variables. In an attempt to use regression methods to extend the relationship found at the overlap between lidar and inventory data, we built empirical models with the independent variables of height, species, crown closure, and operability (access), which were indicated as appropriate variables through a stepwise multiple regression procedure. The resulting relationships were weak, with a poor relationship found between the lidar and polygon categorical data, with a r^2 of 0.23 and a SE of 4.15 m (n = 1143 polygons).

Height estimates from DNs within forest inventory polygons

The Landsat-5 TM imagery also provides a complete coverage of the study area. Using a polygon decomposition approach, each polygon is populated with the co-georegistered Landsat-5 DN. The presence of DN in all polygons allows the relationships created where the overlap of lidar occurs to be extended with a regression approach. Empirical regression relationships are developed between the lidar estimates of polygon height and the co-located DN. The lidar-estimated heights are the dependant variable, and the independent variables are the DNs found within the forest inventory polygons. The resulting relationship is weak and comparable to the relationship of lidar-estimated heights to the forest inventory categorical variables, with a r^2 of 0.26 and a SE of 4.06 m (n = 1143 polygons).

Height estimates from segments within polygons

When combining forest inventory and satellite imagery, the spectral contents found within the co-registered polygons are often poorly related to the polygon attributes. To enable the development of a stronger relationship between the polygon units and the spectral contents, we segmented the imagery using spectral and spatial characteristics. The polygon decomposition on the segmented units resulted in image-based contextual units (rather than photo-interpreted units). Empirical relationships were built between the lidar-estimated heights for each segment and the associated DNs. The resulting relationship between segment-based DNs and lidar-estimated segment heights was good, with a r^2 of 0.61 and a SE of 3.15 m (n = 1837 segments).

Comparison of methods

Figure 3 illustrates the distribution of error by methodological approach. The difference between polygons, segments, and polygons with prorated segment values and the field validation plots is computed to illustrate the accuracy of the differing approaches. The comparison of spatial units to point values is problematic and will result in a reduction of accuracy. There is an averaging of the conditions found within a spatial unit that results in the range of values representing areas not being as variable as it is for points. The largest difference between the plots and estimated heights is for the forest inventory polygons. The segments are well related to the plots. The segment-based height estimates, while in general near zero, also have some instances where the difference from the plot-based estimates is high. Once the segments are generalized

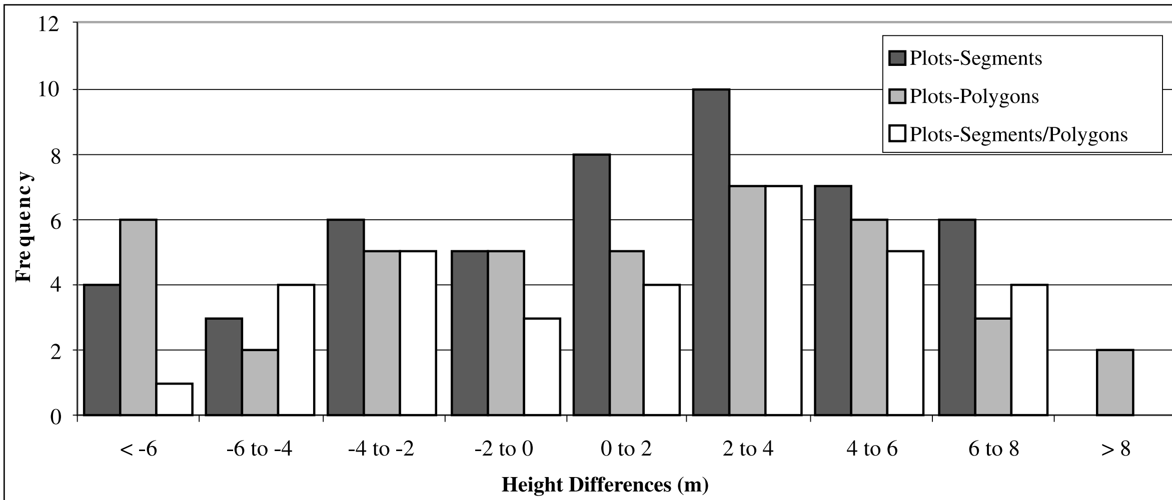

Figure 3. Distribution of error of estimates for segment- and polygon-based lidar height estimates. Height differences are from point heights subtracting the lidar-calibrated spectral-based height estimates.

to represent polygon areas, the instances with the greatest differences with the plot values are reduced. The segment-based height estimates appear to be moderated when summarized to represent the polygons, with some smaller differences introduced, but the large differences removed. **Tables 3** and **4** illustrate the differences for polygon- and segment-based estimates, respectively. The cumulative percentages indicate that 82% of the difference between segment-based heights and plot values is less than 6 m, while for the polygon-based height estimates, 73% of the predictions are less than 6 m. The number of validation points differ owing to the size, shape, and number of segments (n = 49) and polygons (n = 41), with points falling on polygon boundaries not included (**Table 1**).

As described in the field data section, a set of validation points were kept for the accuracy assessment of the final height maps. **Figure 4** illustrates the general trends for each height map. In general, low heights were overestimated and higher heights were underestimated. Against the validation data, coefficients of determination were lowest when estimating heights from the DNs extracted from within polygons (r^2 = 0.30, SE = 4.8 m, n = 41). Using DNs found within the segments resulted in an improved relationship to plot heights (r^2 = 0.45, SE = 4.2 m, n = 49). Prorating the segment-based height estimates within the polygon boundaries results further improved the relationship to the validation plot values (r^2 = 0.67, SE = 3.3 m, n = 41).

Table 3. Height difference between plot- and polygon-based height estimates.

Height difference (m)	Count	% of total	Cumulative (%)
<2	10	24	
2–4	12	29	53
4–6	8	20	73
6–8	8	20	93
8–10	3	7	100
>10	0	0	
Total	41	100.0	

Table 4. Height difference between validation plots and segment-based height estimates.

Height difference (m)	Count	% of total	Cumulative (%)
<2	13	27	
2–4	16	33	60
4–6	11	22	82
6–8	8	16	98
8–10	1	2	100
>10	0	0	
Total	49	100.0	

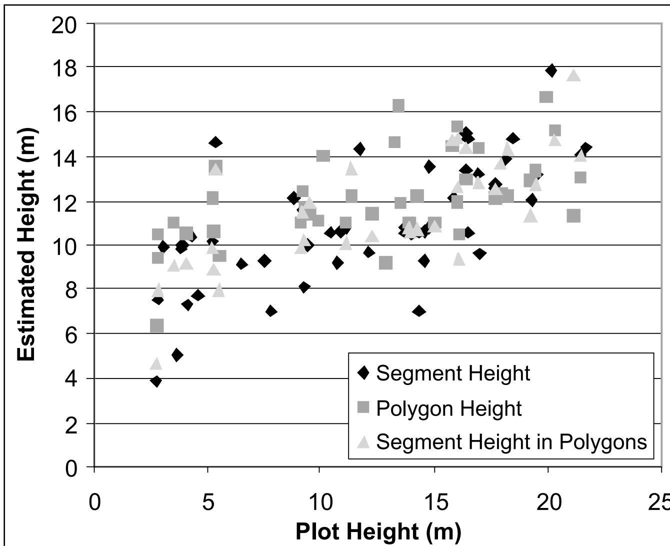

Figure 4. Comparison of plot-based height estimates with segment, polygon, and segments prorated within polygon estimates of height (polygon, n = 41; segment, n = 49). See text for additional description of the relationships.

Conclusions

In this paper we presented the integration of remotely sensed imagery, forest inventory, and lidar data to estimate height values for a landscape from a sample of lidar data. Empirical relationships were developed between lidar heights and DNs of co-geolocated polygons and segments. The regression relationship between lidar height and segment DNs was the strongest of the approaches illustrated with a r^2 of 0.61. From this extension relationship, heights were estimated for the ~7000 km^2 inventory area based upon a 0.48% lidar sample covering 33.9 km^2. Using the set of validation points, we found a r^2 of 0.67 with a SE of 3.3 m for the polygon height estimates from the lidar height estimates from the DNs within the image segments. Further, over 80% of the differences between segment-based heights and plot values were less than 6 m. In Saskatchewan, heights are often grouped into 5-m classes (Gillis and Leckie, 1993). The procedure presented here provides useful update data opportunities in such a context. Following the methods presented in this paper, we were able to use a sample of lidar data, representing less than 0.48% of the inventory region, to calibrate empirical models to estimate an update of the forest inventory heights from 1994 to 1996 conditions.

Based upon this research, future work is encouraged for the use of small-footprint discrete-return lidar. The estimation and extension of an extended suite of vertically distributed attributes, such as biomass and volume, is also indicated. Future work may also be based upon improvements in image processing techniques applied, size and range of validation data used, and statistical techniques applied.

Acknowledgements

Xilin Fang of the Forestry Branch of Saskatchewan Parks and Renewable Resources is thanked for provision of the forest inventory data. Dr. David Harding and the researchers associated with the BOREAS project are also thanked for dataset development and documentation.

References

Blackburn, G. 2002. Remote sensing of forest pigments using airborne image spectrometer and LIDAR imagery. *Remote Sensing of Environment*, Vol. 82, pp. 311–321.

Blair, J., Coyle, D., Bufton, J., and Harding, D. 1994. Optimization of an airborne laser altimeter for remote sensing of vegetation and tree canopies. In *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'94)*, 8–12 August 1994, Pasadena, Calif. IEEE, Piscataway, N.J. Vol. II, pp. 939–941.

Definiens Imaging GmbH. 2000. *eCognition user's guide*. Version 2.1 [computer program]. Definiens Imaging GmbH, München, Germany. Available from <http://www.definiens-imaging.com/central/index.htm>. 237 p.

Gillis, M., and Leckie, D. 1993. *Forest inventory mapping procedures across Canada*. Petawawa National Forestry Institute (PNFI), Forestry Canada, Chalk River, Ont. Information Report PI-X-114, 79 p.

Gougeon, F., St-Onge, B., Wulder, M., and Leckie, D. 2001. Synergy of Airborne Laser Altimetry and Digital Videography for Individual Tree Crown Delineation. In *Proceedings of the 23rd Canadian Symposium on Remote Sensing – 10e Congrès de l'Association québécoise de télédétection, Remote Sensing in the Third Millennium: From Global to Local*, 21–24 August 2001, Québec City, Que. CASI, Ottawa, Ont. pp. 251–257.

Halliwell, D., and Apps, M. 1997a. *Boreal Ecosystem–Atmosphere Study (BOREAS) biometry and auxiliary sites: locations and descriptions*. Northern Forestry Centre, Canadian Forest Service, Natural Resources Canada, Edmonton, Alta. Catalogue No. F.42-266/1-1997E. 120 p.

Halliwell, D., and Apps, M. 1997b. *Boreal Ecosystem–Atmosphere Study (BOREAS) biometry and auxiliary sites: overstory and understory data*. Northern Forestry Centre, Canadian Forest Service, Natural Resources Canada, Edmonton, Alta. Catalogue No. F.42-266/2-1997E. 254 p.

Harding, D. 2000. *BOREAS Scanning Lidar Imager of Canopies by Echo Recovery: Level-3 Data* [online]. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tenn. Available from <http://www.daac.ornl.gov/>.

Hudak, A.T., Lefsky, M.A., Cohen, W., and Berterretche, M. 2002. Integration of lidar and Landsat ETM+ for estimating and mapping forest canopy height. *Remote Sensing of Environment*, Vol. 82, pp. 397–416.

Leckie, D., and Gillis, M. 1995. Forest inventory in Canada with an emphasis on map production. *Forestry Chronicle*, Vol. 71, No. 1, pp. 74–88.

Lefsky, M.A., and Cohen, W.B. 2003. Selection of remotely sensed data. In *Remote Sensing of Forest Environments: Concepts and Case Studies*. Edited by M. Wulder and S. Franklin. Kluwer Academic Publishers, Boston. pp. 13–46.

Lefsky, M.A., Cohen, W.B., and Spies, T.A. 2001. An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Canadian Journal of Forest Research*, Vol. 31, pp. 78–87.

Lefsky, M.A., Cohen, W.B., Parker, G.G., and Harding, D.J. 2002. Lidar remote sensing for ecosystem studies. *BioScience*, Vol. 52, No. 1, pp. 19–30.

Lim, K., Treitz, P., Wulder, M., St-Onge, B., and Flood, M. 2002. Lidar remote sensing of forest structure. *Progress in Physical Geography*, Vol. 27, No. 1. pp. 88–106.

Lowe, J., Power, K., and Marsan, M. 1996. *Canada's forest inventory 1991: Summary by terrestrial ecozones and ecoregions*. Pacific Forestry Centre, Canadian Forest Service, Natural Resources Canada, Victoria, B.C. Information Report BC-X-364E. 56 pp.

Magnussen, S., Boudewyn, P., Wulder, M., and Seemann, D. 2000. Predictions of forest inventory cover type proportions using Landsat TM. *Silva Fennica*, Vol. 34, No. 4, pp. 351–370.

Markham, B., and Barker, J. 1986. Landsat MSS and TM post calibration dynamic ranges, exoatmospheric reflectances and at satellite temperature. *EOSAT Landsat Technical Notes*, August, No. 1, pp. 3–7.

Næsset, E. 1996.Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 52, pp. 49–56.

Pal, N., and Pal, S. 1993. A review on image segmentation techniques. *Pattern Recognition*, Vol. 26, No. 9, pp. 1277–1294.

Rowe, J. 1977. *Forest regions of Canada*. Canadian Forest Service, Natural Resources Canada, Ottawa Ont. 172 p.

Sellers, P., Hall, F., Margolis, H., Kelly, B., Baldocchi, D., den Hartog, G., Cihlar, J., Ryan, M., Goodison, B., Crill, P., Ranson, K., Lettenmaier, D., and Wickland, D. 1995. The Boreal Ecosystem–Atmosphere Study (BOREAS): An Overview and Early Results from the 1994 Field Year. *Bulletin of the American Meteorological Society*, Vol. 76, No. 9, pp. 1549–1577.

Wulder, M. 1998a. Optical remote sensing techniques for the assessment of forest inventory and biophysical parameters. *Progress in Physical Geography*, Vol. 22, No. 4, pp. 449–476.

Wulder, M. 1998b. The prediction of leaf area index from forest polygons decomposed through the integration of remote sensing, GIS, UNIX, and C. *Computers & Geosciences*, Vol. 24, No.2, pp. 151–157.

Wulder, M., and Franklin, S. 2001. Polygon decomposition with remotely sensed data: Rationale, methods, and applications. *GEOMATICA*, Vol. 55, No. 1, pp. 11–21.

Wulder, M., Magnussen, S., Boudewyn, P., and Seemann, D. 1999. Spectral variability related to forest inventory polygons stored within a GIS. In *Proceedings of the IUFRO Conference on Remote Sensing and Forest Monitoring*, 1–3 June 1999, Rogow, Poland. European Communities, Italy. pp.141–153.

Wulder, M., Seemann, D., and Bouchard, A. 2001. Within polygon grid based sampling for height estimation with LIDAR data. In *Proceedings of the 23rd Canadian Symposium on Remote Sensing – 10e Congrès de l'Association québécoise de télédétection, Remote Sensing in the Third Millennium: From Global to Local*, 21–24 August 2001, Québec City, Que. CASI, Ottawa, Ont. pp. 265–269.

