ELSEVIER

Remote Sensing of Environment 82 (2002) 397–416

Remote Sensing of Environment

www.elsevier.com/locate/rse

Integration of lidar and Landsat ETM+ data for estimating and mapping forest canopy height

Andrew T. Hudak [a,*], Michael A. Lefsky [b,1], Warren B. Cohen [a,2], Mercedes Berterretche [b,3]

[a] *Pacific Northwest Research Station, USDA Forest Service, Corvallis, OR, USA*
[b] *Department of Forest Science, Oregon State University, Corvallis, OR, USA*

Received 30 July 2001; received in revised form 25 April 2002; accepted 28 April 2002

Abstract

Light detection and ranging (lidar) data provide accurate measurements of forest canopy structure in the vertical plane; however, current lidar sensors have limited coverage in the horizontal plane. Landsat data provide extensive coverage of generalized forest structural classes in the horizontal plane but are relatively insensitive to variation in forest canopy height. It would, therefore, be desirable to integrate lidar and Landsat data to improve the measurement, mapping, and monitoring of forest structural attributes. We tested five aspatial and spatial methods for predicting canopy height, using an airborne lidar system (Aeroscan) and Landsat Enhanced Thematic Mapper (ETM+) data: regression, kriging, cokriging, and kriging and cokriging of regression residuals. Our 200-km^2 study area in western Oregon encompassed Oregon State University's McDonald–Dunn Research Forest, which is broadly representative of the age and structural classes common in the region. We sampled a spatially continuous lidar coverage in eight systematic patterns to determine which lidar sampling strategy would optimize lidar–Landsat integration in western Oregon forests: transects sampled at 2000, 1000, 500, and 250 m frequencies, and points sampled at these same spatial frequencies. The aspatial regression model results, regardless of sampling strategy, preserved actual vegetation pattern, but underestimated taller canopies and overestimated shorter canopies. The spatial models, kriging and cokriging, produced less biased results than regression but poorly reproduced vegetation pattern, especially at the sparser (2000 and 1000 m) sampling frequencies. The spatial model predictions were more accurate than the regression model predictions at locations <200 m from sample locations. Cokriging, using the ETM+ panchromatic band as the secondary variable, proved slightly more accurate than kriging. The integrated models that kriged or cokriged regression residuals were preferable to either the aspatial or spatial models alone because they preserved the vegetation pattern like regression yet improved estimation accuracies above those predicted from the regression models alone. The 250-m point sampling strategy proved most optimal because it oversampled the landscape relative to the geostatistical range of actual spatial variation, as indicated by the sample semivariograms, while making the sample data volume more manageable. We concluded that an integrated modeling strategy is most suitable for estimating and mapping canopy height at locations unsampled by lidar, and that a 250-m discrete point sampling strategy most efficiently samples an intensively managed forested landscape in western Oregon.
© 2002 Published by Elsevier Science Inc.

1. Introduction

Currently, light detection and ranging (lidar) data provide detailed information on forest canopy structure in the vertical plane, but over a limited spatial extent (Lefsky, Cohen, Parker, & Harding, 2002). Landsat data provide useful structural information in the horizontal plane (Cohen & Spies, 1992) but are relatively insensitive to canopy height. Lidar–Landsat Enhanced Thematic Mapper (ETM+) integration is, therefore, a logical goal to pursue. No remote sensing instrument is suited for all applications, and there have been several calls for improving the applicability of remotely sensed data through multisensor integration. Most multisensor integration studies have involved Landsat imagery (e.g., Asner, Wessman, & Privette, 1997; Oleson et al., 1995) but none has integrated Landsat imagery with lidar data.

Lidar–Landsat ETM+ integration has immediate relevance due to the anticipated launches of the Ice, Cloud, and Land Elevation Satellite (ICESat) and Vegetation Canopy

* Corresponding author. Tel.: +1-208-883-2327; fax: +1-208-883-2318.
E-mail addresses: ahudak@fs.fed.us (A.T. Hudak), lefsky@fsl.orst.edu (M.A. Lefsky), wcohen@fs.fed.us (W.B. Cohen), meche@ica.com.uy (M. Berterretche).
[1] Tel.: +1-541-758-7765; fax: +1-541-758-7760.
[2] Tel.: +1-541-750-7322; fax: +1-541-758-7760.
[3] Tel.: +1-598-2-208-3577; fax: +1-598-2-208-5941.

0034-4257/02/$ - see front matter © 2002 Published by Elsevier Science Inc.
PII: S0034-4257(02)00056-1

Lidar (VCL) satellite missions. Global sampling of the Earth's forests, as VCL should provide, will be a huge boon for forest resource assessments. For example, the VCL mission has potential to greatly narrow the uncertainty surrounding estimates of global carbon pools (Drake et al., 2002). Discontinuous lidar data will need to be integrated with continuous optical imagery to produce comprehensive maps that have practical value to forest ecologists and forest resource managers (Lefsky, Cohen, Hudak, Acker, & Ohmann, 1999). Given the continued demand for Landsat imagery and the growing supply of ETM+ data, Landsat 7 is a logical choice for integration with lidar sample data.

In this study, our first objective was to estimate canopy height at locations unsampled by lidar, based on the statistical and geostatistical relationships between the lidar and Landsat ETM+ data at the lidar sample locations. We used basic data from lidar (maximum canopy height) and Landsat ETM+ (raw band values) and tested widely used, straightforward empirical estimation methods: ordinary least squares (OLS) regression, ordinary kriging (OK), and ordinary cokriging (OCK).

Prior research has shown that landscape pattern varies principally as a function of the areal size of individual stands in the heavily managed forests of western Oregon, or at a typical scale of 250–500 m (Milne & Cohen, 1999). Our second objective was to determine which spatial sampling pattern (contiguous transect or discrete point) and frequency (2000, 1000, 500, or 250 m) would optimize the integration of lidar and Landsat ETM+ data for accurately estimating and mapping canopy height in intensively managed, coniferous forest landscapes.

2. Background

2.1. Lidar

Lidar is an active remote sensing technology like radar but operating in the visible or near-infrared region of the electromagnetic spectrum. Lidar, at its most basic level, is a laser altimeter that determines the distance from the instrument to the physical surface by measuring the time elapsed between a laser pulse emission and its reflected return signal. This time interval multiplied by the speed of light measures twice the distance to the target; dividing this measurement by two can thus provide a measure of target elevation (Bachman, 1979). Processing of the return signal may identify multiple pulses and returns. As a result, trees, buildings, and other objects are apparent in the lidar signal, permitting accurate calculation of their heights (Nelson, Krabill, & Maclean, 1984). Studies using coincident field data have indicated that lidar data can provide nonasymptotic estimates of structural attributes such as basal area, biomass, stand volume (Lefsky, Cohen, Acker, et al., 1999; Lefsky, Harding, Cohen, Parker, & Shugart, 1999; Nelson, Oderwald, & Gregoire, 1997; Nilsson, 1996; Means et al., 1999, 2000), and leaf area index (LAI) (Lefsky, Cohen, Acker, et al., 1999), even in high-biomass forests. Lidar allows extraordinary structural differentiation between primary and secondary forest that is currently unrivaled by any other remote sensing technology (Drake et al., 2002; Lefsky, Cohen, Acker, et al., 1999; Weishampel, Blair, Knox, Dubayah, & Clark, 2000).

Lidar instruments can be divided into two general categories: discrete return and waveform sampling (Lefsky et al., 2002). They are distinguished in part by the size of the laser illumination area, or footprint, which typically is smaller with discrete return systems (0.25–1 m) than with waveform sampling systems (10–100 m). Waveform sampling systems compensate for their coarser horizontal resolution with finer vertical resolution, providing submeter vertical profiles, while discrete return systems record only one to five returns per laser footprint. Discrete return systems are more suited for supplying the demand for accurate, high-resolution topographic maps and digital terrain models, and are therefore becoming widely available in the commercial sector (Lefsky et al., 2002). Waveform sampling systems have been useful in forest canopy research applications, such as biomass assessments in temperate (Lefsky, Cohen, Acker, et al., 1999; Lefsky, Harding, et al., 1999) and tropical forests (Drake et al., 2002).

VCL is a spaceborne, waveform sampling lidar system that will inventory forests globally between $\pm 68^\circ$ latitude for an estimated 2 years. VCL footprints will be approximately 25 m in diameter and arrayed in each of three parallel transects separated by 4-km intervals. The 8-km ground swath will be randomly placed on the Earth's surface, with the ascending and descending orbital paths intersecting at 65° angles. The goals of the VCL mission are to characterize the three-dimensional structure of the Earth and its land cover, improve climate modeling and prediction, and to provide a global reference data set of topographic spot heights and transects (Dubayah et al., 1997; http://essp.gsfc.nasa.gov/vcl).

ICESat is a spaceborne, waveform sampling lidar system that will measure and monitor ice sheet and land topography as well as cloud, atmospheric, and vegetation properties. Like VCL, it will acquire data in the near-infrared region at 1064 nm, but also in the visible green region at 532 nm. ICESat will have 70-m spot footprints spaced at 175-m intervals (http://icesat.gsfc.nasa.gov/intro.html).

2.2. Landsat ETM+

Landsat imagery is the most common satellite data source used in terrestrial ecology. This is, in large part, due to its widespread availability and unrivaled length of record (since 1972), but also because the grain, extent, and multispectral features make Landsat suitable for a variety of environmental applications at landscape to regional scales.

Landsat spectral data are typically related to biophysical attributes via spectral vegetation indices (SVIs). Ecologically relevant structural attributes such as LAI have been estimated from SVIs of croplands (e.g., Asrar, Fuchs, Kanemasu, & Hatfield, 1984; Wiegand, Richardson, & Kanemasu, 1979), grasslands (e.g., Friedl, Michaelsen, Davis, Walker, & Schimel, 1994), shrublands (e.g., Law & Waring, 1994), and forests (e.g., Chen & Cihlar, 1996; Cohen, Spies, & Fiorella, 1995; Fassnacht, Gower, MacKenzie, Nordheim, & Lillesand, 1997; Turner, Cohen, Kennedy, Fassnacht, & Briggs, 1999). Sensitivity of SVIs to variation in LAI or biomass generally declines, however, as foliar densities increase within and among ecosystems (e.g., Turner et al., 1999). The greater structural complexity of forests requires, not surprisingly, more complex image processing techniques. For instance, Cohen and Spies (1992) used all six Landsat reflectance bands, rather than just the red and near-infrared bands as with most SVIs. Other notable, yet more complicated, approaches to enhancing the extraction of canopy structure information from Landsat imagery include using multitemporal TM data to capture variable illumination conditions (Lefsky, Cohen, & Spies, 2001) and spectral mixture analysis to quantify canopy shadows (e.g., Adams et al., 1995; Peddle, Hall, & LeDrew, 1999). The new ETM+ instrument on board Landsat 7 features enhanced radiometric resolution over its TM predecessor, which should aid all of the empirical methods just described. Yet there are fundamental limitations to the utility of passive optical sensors for characterizing vertical forest canopy structure, which will probably make them perpetually inferior to lidar for this task (Lefsky et al., 2001).

2.3. Estimation methods

All of the estimation methods we employed are empirical and were chosen for their broad use and general applicability: OLS regression, OK, and OCK. The literature documents many variations on these aspatial (e.g., Cohen, Maierpserger, Gower, & Turner, in press; Curran & Hay, 1986) and spatial (e.g., Cohen, Spies, & Bradshaw, 1990; Journel & Rossi, 1989; Knotters, Brus, & Voshaar, 1995; Pan, Gaard, Moss, & Heiner, 1993; Stein & Corsten, 1991) estimation methods. We deemed it less useful to conduct an exhaustive study of these than to concentrate on the three methods just named because they broadly represent basic empirical estimation techniques.

2.3.1. Ordinary Least Squares (OLS) regression

Regression is an aspatial method because it assumes spatial independence in the data. Users of regression models for mapping vegetation attributes must be mindful of this aspect, since geographic data are often autocorrelated. The OLS method determines the best fit for a straight line that minimizes the sum of squared deviations of the observed data values (dependent variable only) away from that line. The OLS multiple regression model takes the general form:

$$Z = \alpha + \beta_i(X_i) + \varepsilon \tag{1}$$

where, in this study, Z is the dependent variable (height), X_i is the i explanatory variable (Landsat Bands 1–7 and UTMX and UTMY locations), β_i is the linear slope coefficient corresponding to X_i, α is the intercept, and ε is the residual error. A thorough discussion of OLS and other multivariate regression methods is available in Kleinbaum, Kupper, Muller, and Nizam (1998).

2.3.2. Ordinary Kriging (OK)

Spatial models are appropriate if there is spatial dependence in the data, as in this study. Kriging interpolates the sample data to estimate values at unsampled locations, based solely on a linear model of regionalization. The linear model of regionalization essentially is a weighting function required to krig and can be graphically represented by a semivariogram. The semivariogram plots semivariance γ as a function of the distance between samples, known as the lag distance h, according to:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha + h)]^2 \tag{2}$$

where $\gamma(h)$ is semivariance as a function of lag distance h, $N(h)$ is the number of pairs of data locations separated by h, and z is the data value at locations $\mathbf{u}_\alpha$ and $(\mathbf{u}_\alpha + h)$ (Goovaerts, 1997). Semivariance typically increases with spatial resolution, which reflects decreasing spatial autocorrelation with increasing lag distance. Semivariograms have three main parameters. The *nugget* is the semivariance at a lag distance of zero; it represents spatial variability independent of the sampling frequency. The *sill* is the semivariance at lags where there is no spatial autocorrelation. The *range* is the lag distance at which the sill is reached.

The sample semivariogram describes the spatial autocorrelation unique to that sample dataset. Model semivariograms are formulated by fitting a mathematical function to simulate the nugget, sill, range, and thus the shape of the sample semivariogram. The functions used in a model semivariogram can be linear, but are almost always nonlinear: spherical, exponential, Gaussian, and power functions are most common. In this study, the sample semivariograms were simulated best by exponential models, which take the form:

$$\gamma(h) = c\left[1 - \exp\left(\frac{-3h}{a}\right)\right] \tag{3}$$

where a is the practical range of the semivariogram, defined as the lag distance at which the model value is at 95% of the sill c, as the exponential model approaches the sill asymptotically (Deutsch & Journel, 1998). Several of these functions may be nested to assemble a model semivariogram that fits the sample semivariogram well. The theory behind the

linear model of regionalization and practical issues associated with fitting model semivariograms have been thoroughly presented elsewhere (e.g., Goovaerts, 1997; Isaaks & Srivastava, 1989).

Kriging can be considered a weighted averaging procedure that minimizes errors about the mean, with weights assigned as a function of distance to neighboring samples. The OK model estimates a value Z^* at each location $\mathbf{u}$ and takes the general form:

$$Z^*(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u}) Z(\mathbf{u}_\alpha) \tag{4}$$

where Z is the primary variable and λ_α and $\mathbf{u}_\alpha$ are the weights and locations, respectively, of n neighboring samples (Goovaerts, 1997). The user typically controls the size of n by specifying a search neighborhood. The OK estimator minimizes the error variance and is nonbiased in that the error mean equals zero. The OK estimator allows for a locally varying mean by forcing the kriging weights to sum to one:

$$\sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u}) = 1. \tag{5}$$

2.3.3. OCK

Cokriging is a multivariate extension of kriging and relies on a linear model of coregionalization that exploits not only the autocorrelation in the primary variable, but also the cross-correlation between the primary variable and a secondary variable, which can be graphically represented by the cross-semivariogram, defined as:

$$\gamma_{ij}(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z_i(\mathbf{u}_\alpha) - z_i(\mathbf{u}_\alpha + h)] \times [z_j(\mathbf{u}_\alpha) - z_j(\mathbf{u}_\alpha + h)] \tag{6}$$

where $\gamma_{ij}(h)$ is the cross-semivariance between variables i and j, $N(h)$ is the number of pairs of data locations separated by lag distance h, z_i is the data value of variable i at locations $\mathbf{u}_\alpha$ and $(\mathbf{u}_\alpha + h)$, and z_j is the data value of variable j at the same locations (Goovaerts, 1997).

With the primary variable Z and a single secondary variable Y, the OCK estimator of Z^* at location $\mathbf{u}$ takes the form:

$$Z^*(\mathbf{u}) = \sum_{\alpha_1=1}^{n_1(\mathbf{u})} \lambda_{\alpha_1}(\mathbf{u}) Z(\mathbf{u}_{\alpha_1}) + \sum_{\alpha_2=1}^{n_2(\mathbf{u})} \lambda_{\alpha_2}(\mathbf{u}) Y(\mathbf{u}_{\alpha_2}) \tag{7}$$

where λ_{α_1} and $\mathbf{u}_{\alpha_1}$ are the weights and locations, respectively, of the n_1 primary data, and λ_{α_2} and $\mathbf{u}_{\alpha_2}$ are the weights and locations, respectively, of the n_2 secondary data (Goovaerts, 1997). As with kriging, the user can control the size of n_1 and n_2 by specifying search neighborhoods for each variable.

A useful feature of cokriging is that the locations of the primary and secondary variables need not correspond. Inclusion of auxiliary information in cokriging often results in more accurate estimates than from kriging, but at a cost of greater model complexity and higher computational demands. An important condition for a valid cokriging model is satisfying the positive definiteness constraint on the linear model of coregionalization, which ensures that the coefficient matrix for the system of kriging equations is invertible; details are available in Isaaks and Srivastava (1989).

In this study, we used the traditional OCK estimator, which operates under two nonbias constraints:

$$\sum_{\alpha_1=1}^{n_1(\mathbf{u})} \lambda_{\alpha_1}(\mathbf{u}) = 1; \quad \sum_{\alpha_2=1}^{n_2(\mathbf{u})} \lambda_{\alpha_2}(\mathbf{u}) = 0. \tag{8}$$

We used traditional OCK rather than standardized OCK, which operates under a single nonbias constraint: both the primary and secondary data weights must sum to one. Our rationale was based on the recommendation of Deutsch and Journel (1998); traditional OCK is more appropriate when the primary data (25 m lidar data, in this study) are undersampled with respect to the secondary data (15 m ETM+ panchromatic data, in this study). Details about standardized OCK, and other kriging or cokriging algorithms, may be found in Goovaerts (1997) or Isaaks and Srivastava (1989).

3. Methods

3.1. Study area

The 200-km^2 study area features Oregon State University's McDonald–Dunn Research Forest (Fig. 1) in the eastern foothills of the Coast Range in western Oregon. The area has elevations ranging from 58 to 650 m. Most of the area is coniferous forest dominated by *Pseudotsuga menziesii* (Douglas-fir) and codominated by *Tsuga heterophylla* (western hemlock), but hardwood stands featuring *Acer macrophyllum* (bigleaf maple) and *Quercus garryana* (Oregon white oak) also are common. Stands span the full range of successional stages: young, intermediate, mature, and old growth, and three management regimes: even-aged, two-storied, and uneven-aged (http://www.cof.orst.edu/resfor/mcdonald/purpose.sht).

3.2. Image processing

Small footprint lidar data were acquired from an airborne platform (Aeroscan, Spencer B. Gross, Portland, OR) in January 2000. The Aeroscan instrument records five vertical returns within small footprints having an average diameter of 60 cm and geolocated in real time using an on-board, differential global positioning system to an accuracy of 75 (horizontal) and 30 cm (vertical). North–south paths were flown to provide continuous lidar coverage (i.e., a pseudo-image) of the entire area. Maximum canopy height was calculated for each footprint as the difference between the

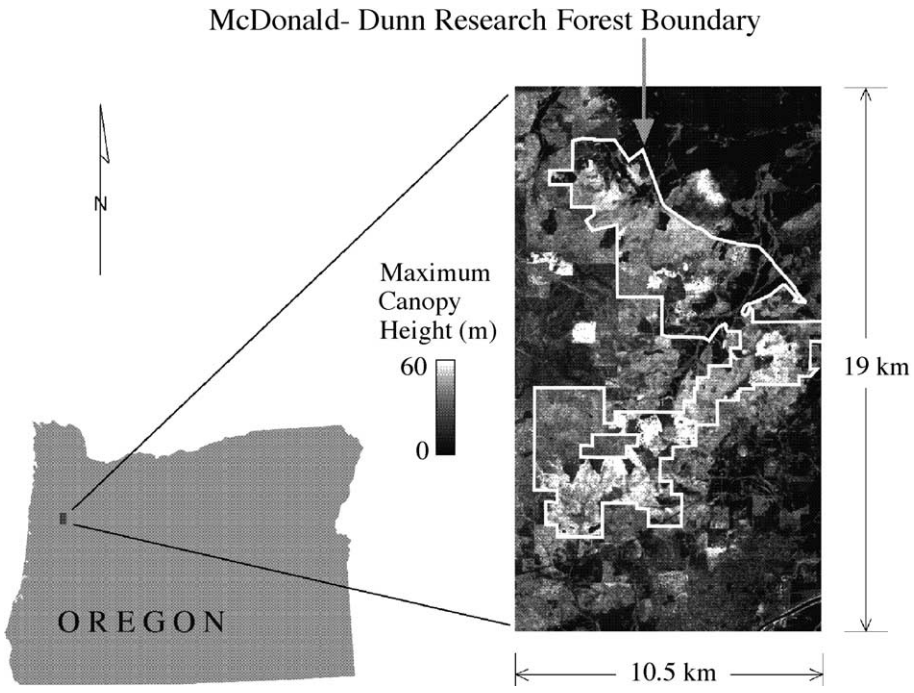

Fig. 1. Study area location and maximum canopy height image measured by lidar.

first (canopy top) and last (ground) returns. Maximum height values in each footprint were then spatially aggregated into 25-m bins to produce a maximum canopy height image of 25 m spatial resolution (Fig. 1). Every 25-m pixel was assigned a maximum canopy height value from a population of 10–764 lidar footprints, with a median of 26 footprints per pixel.

A Landsat ETM+ image (Path/Row = 46/29) acquired on September 7, 1999 was coregistered to a 1988 base image using 90 tie points selected through an automated spatial covariance procedure (Kennedy & Cohen, in press). Georegistration was performed in Imagine (ERDAS, Atlanta, GA) using a first-order polynomial function with nearest neighbor resampling (root mean square error = 14.3 m).

3.3. Sampling strategies

The lidar height image allowed us to sample a variety of spatial patterns and frequencies. Using ERDAS Imagine, we sampled the study area, in both transect and point patterns, at spatial frequencies of 2000, 1000, 500, and 250 m, for a total of eight sample height datasets (Fig. 2).

3.4. Estimation methods

The histogram of the maximum canopy height data exhibited a strong positive skew. We therefore normalized each of the eight height datasets with a square root transformation (SQRTHT) prior to applying any of the estimation methods; afterwards, all estimated SQRTHT values were backtransformed (squared) before comparing to measured height values.

3.4.1. Aspatial

The SQRTHT sample data were regressed on the raw ETM+ Bands 1–7, as well as the Universal Transverse Mercator (UTM) X and Y locations, using stepwise multiple linear regression in Interactive Data Language (IDL; Research Systems, Boulder, CO). Variables were assigned only if they added significantly to the model ($\alpha = 0.05$).

3.4.2. Spatial

The SQRTHT sample data were normal score transformed prior to modeling. This nonlinear, ranked transformation normalizes the data to produce a standard Gaussian cumulative distribution function with mean equal to zero and variance equal to one (Deutsch & Journel, 1998). After modeling, the estimates were backtransformed to the original SQRTHT data distribution; the estimates at the sample locations proved to be an exact reproduction of the original SQRTHT sample data.

OK and OCK operations were performed using algorithms in GSLIB (Statios, San Francisco, CA). We modeled the sample semivariograms by nesting nugget estimates with two exponential models. Only a model semivariogram for the primary variable was needed for OK. For cokriging, a model semivariogram was also required for the secondary variable, along with a cross-semivariogram. The ETM+ panchromatic band was the logical choice to serve as a secondary variable for cokriging, since this band has the highest resolution (15 m) among the ETM+ bands and, therefore, the highest spatial information content. The secondary data were also normal score transformed before modeling. We were careful to observe the positive definiteness constraint on the linear model of coregionalization

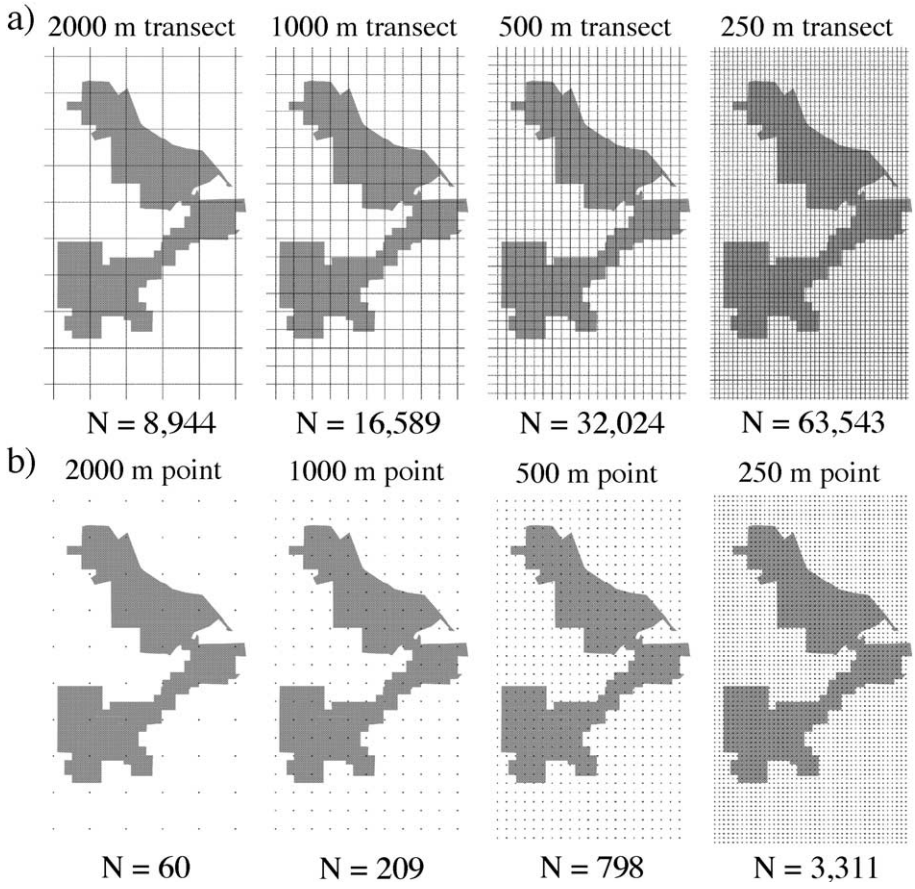

Fig. 2. (a) Transect and (b) Point sampling strategies tested and data volume of each sample dataset. The McDonald–Dunn Research Forest is shown in the background (see Fig. 1) for a spatial frame of reference.

while developing the three semivariogram models required for each cokriging operation (Goovaerts, 1997; Isaaks & Srivastava, 1989).

3.4.3. Integrated

Residuals from the OLS regression models were exported from IDL as ASCII files and imported into GSLIB for kriging/cokriging. The same rules and procedures were followed for modeling the residuals as for modeling the SQRTHT data.

3.5. Validation

The image of lidar-measured height values allowed exhaustive validation of the five estimation methods and eight sampling strategies tested. To ensure comparability, the same validation points were used to evaluate all estimation methods and sampling strategies. Two sets of validation points were systematically selected to compare measured and estimated height values using Pearson's correlation statistic. One set of validation points was designed to assess the height estimates for the study area as a whole, with no regard to distance from sample locations; the other set was designed to assess the height estimates as a function of distance from sample locations (Fig. 3).

Histograms, scatterplots, and graphs of measured vs. estimated height values were plotted, and correlation coefficients were calculated, in IDL. Estimated height and estimation error images were mapped in Arc/Info GRID (ESRI, Redlands, CA). Moran's coefficient (I) calculations for spatial autocorrelation in the model residuals were performed using S-PLUS (Insightful, Seattle, WA) functions developed by Dr. Robin Reich (http://www.cnr.colostate.edu/~robin/). The significance test to evaluate each I statistic assumed normality in 700 residual values sampled (Fig. 3) from the population of errors. The theory underlying Moran's I statistic can be pursued more thoroughly in Cliff and Ord (1981) and Moran (1948).

4. Results

4.1. Empirical models

Separate stepwise multiple regression models were developed for the eight sampling strategies tested. In

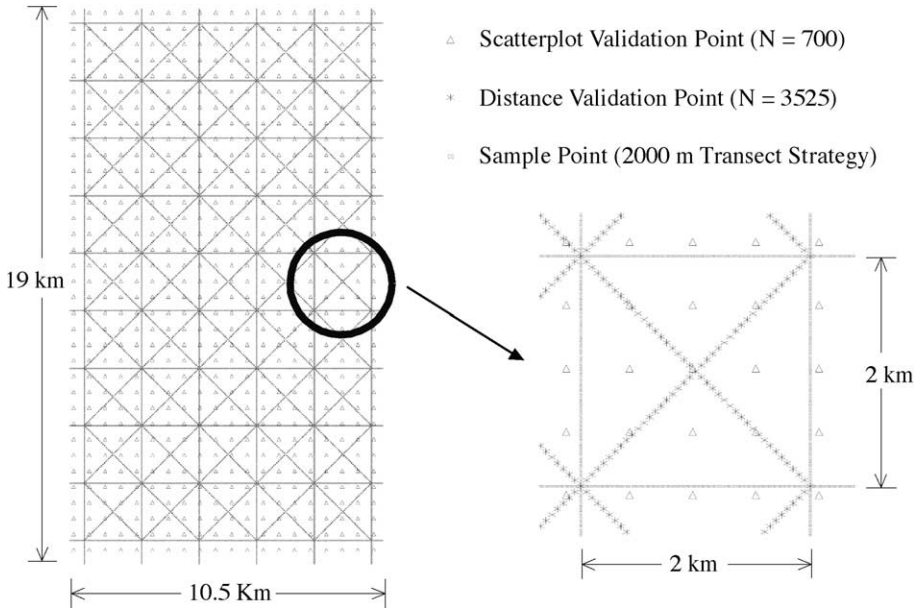

Fig. 3. Validation point locations. A grid of points representing the whole study area, but excluding any samples used for modeling, was systematically sampled to produce scatterplots of measured vs. estimated height (△). Separate validation points were systematically sampled to evaluate the height estimates as a function of distance from sample locations (*). Sample points for the 2000-m transect sampling strategy are shown in the background (see Fig. 2) for a spatial frame of reference (▩).

every case, ETM+ Band 7 was the first variable selected (Table 1). All nine independent variables contributed significantly, and were therefore included, in the four transect cases. The number of variables included in the point models decreased as sample data volume decreased, with only one variable selected in the lower extreme case (2000 m point strategy).

For the spatial and integrated models, unique semivariogram models of the height and height residual datasets were generated for all eight of the sampling strategies tested. The range and sill parameters, and the shape of the semivariograms, were very similar among the eight height datasets, and among the eight height residual datasets (Fig. 4). Nugget variance increased in the cases of the relatively sparse 1000- and 2000-m point samples. For cokriging, each of the eight sampling strategies also required unique model semivariograms of the secondary data semivariograms and the respective cross-semivariograms. As with the primary datasets, the range and sill parameters and semivariogram shapes were consistent amongst all eight sample datasets, and nugget variance was again greater in the 1000- and 2000-m point samples. There was less spatial autocorrelation to exploit in the residual data than in the SQRTHT data. Similarly, the spatial cross-correlation between the primary and secondary data was considerable with regard to the SQRTHT datasets, but relatively low with regard to the residual datasets. Very

Table 1
Multiple regression models for the (A) transect and (B) point sample datasets; explanatory variables are listed in the order of forward stepwise selection

(A) Transects	
2000 m	SQRTHT=144.18−0.0132546(B_7)−0.0768883(B_6)−0.0000242784(UTMY)−0.0119655(B_5)+0.00631775(B_4)+0.0633678(B_1)−0.0195735(B_3)−0.000023421(UTMX)−0.0343936(B_2)
1000 m	SQRTHT=169.426−0.0144084(B_7)−0.0733225(B_6)−0.0000299485(UTMY)−0.0126307(B_5)+0.00528321(B_4)+0.065575(B_1)−0.0528681(B_2)−0.0000165147(UTMX)−0.00862129(B_3)
500 m	SQRTHT=177.901−0.0149903(B_7)−0.0734557(B_6)−0.000031937(UTMY)−0.0132014(B_5)+0.0051818(B_4)−0.00837438(B_3)+0.0600543(B_1)−0.0506589(B_2)−0.0000129396(UTMX)
250 m	SQRTHT=162.93−0.0148533(B_7)−0.0700705(B_6)−0.0000286052(UTMY)−0.0136385(B_5)−0.00704695(B_3)+0.0608232(B_1)−0.0550868(B_2)+0.00571048(B_4)−0.000016865(UTMX)
(B) Points	
2000 m	SQRTHT=6.64595−0.0621982(B_7)
1000 m	SQRTHT=12.8103−0.0445905(B_7)−0.0517594(B_6)
500 m	SQRTHT=155.678−0.0213743(B_7)−0.0813073(B_6)−0.0000281888(UTMY)−0.0136584(B_5)+0.0048719(B_4)
250 m	SQRTHT=154.923−0.0206861(B_7)−0.0682174(B_6)−0.0000263197(UTMY)−0.0168956(B_5)+0.00383822(B_4)−0.000020857(UTMX)

SQRTHT=height (square root transformed); B_1–B_7=Landsat ETM+ Bands 1–7; UTMX, UTMY=X and Y UTM locations.

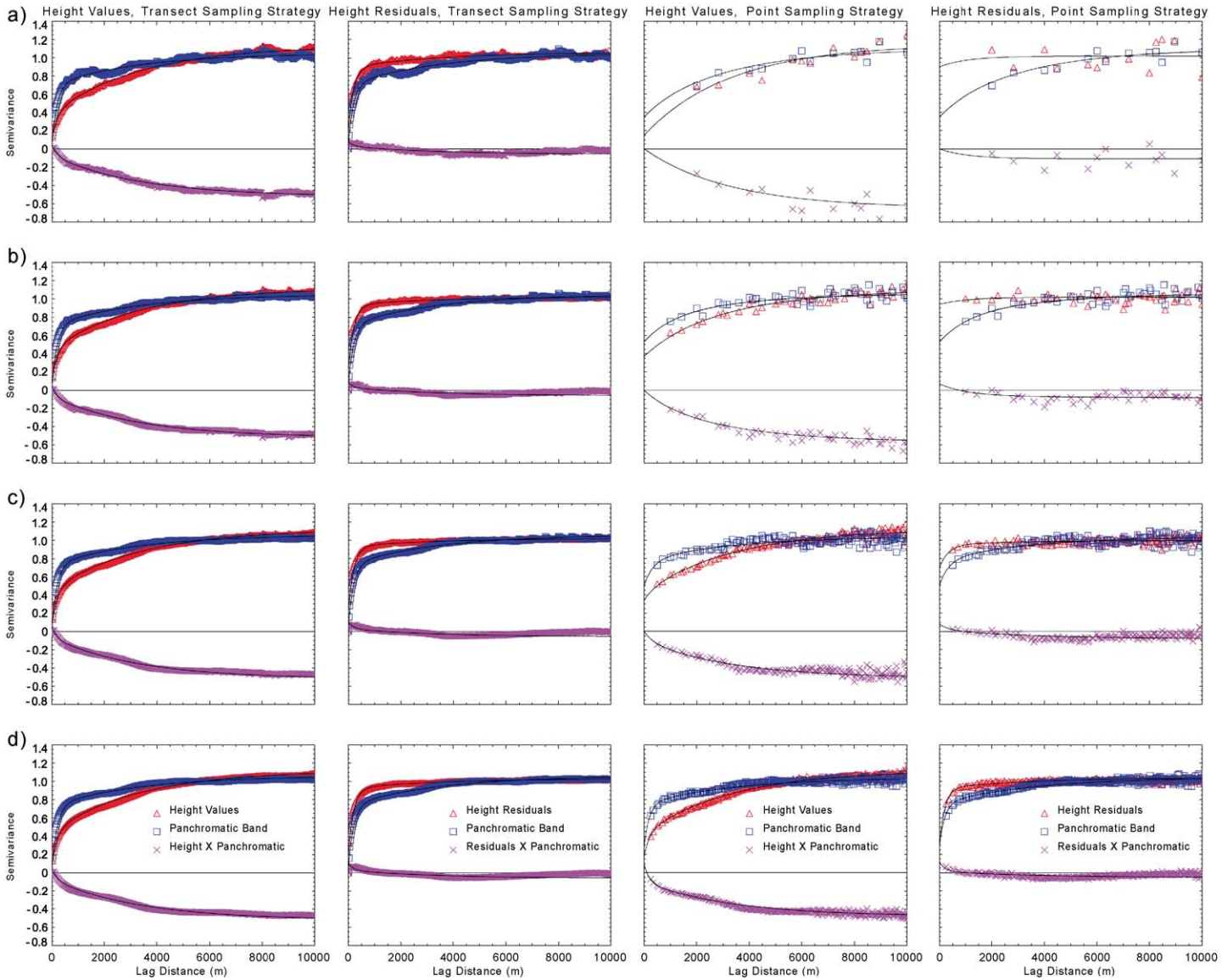

Fig. 4. Sample and model semivariograms of the primary and secondary datasets, from transect and point sampling intervals of: (a) 2000, (b) 1000, (c) 500, and (d) 250 m. The lines plotted over the sample semivariograms are the fitted model semivariograms. The cross-semivariograms are negative because the primary variables (height or height residuals) are negatively correlated to the secondary variable (ETM+ panchromatic band).

tight model fits were achieved for all primary, secondary, and cross-semivariograms (Fig. 4) by nesting a nugget value and two exponential models (Table 2).

4.2. Estimation accuracy

4.2.1. Global

Histograms of the full populations of estimated height values were used to evaluate global accuracy (Fig. 5). Deviations in the estimated height histograms away from the measured height histogram were a good indicator of estimation biases at various heights. These biases were most pronounced in all of the regression results, and in the kriging/cokriging results based on sparse point samples (1000 or 2000 m). Biases in the estimates from the integrated methods were relatively minor, and decreased as sampling frequency increased. Correlations between measured and estimated heights were always better using the integrated models than using either the regression or spatial models alone. Cokriging produced slightly higher correlations than kriging. Correlations also were higher with the transect samples than with the point samples at each spatial sampling frequency.

Scatterplots of measured vs. estimated height values were also generated to compare the five models and eight sampling strategies tested (Fig. 6). Deviations in the slope of the fitted trendlines away from the 1:1 line helped show that the regression models suffered the most from underestimating the taller heights while overestimating the shorter heights. These deviations corresponded closely with the

Table 2
Model semivariograms for the (A) transect and (B) point sample datasets

(A) Transects		
2000 m	SQRTHT	$\gamma(h) = 0.10 + 0.28*\exp_{(600\ m)} + 0.75*\exp_{(10{,}000\ m)}$
	Residuals	$\gamma(h) = 0.20 + 0.69*\exp_{(600\ m)} + 0.17*\exp_{(10{,}000\ m)}$
	Panchromatic	$\gamma(h) = 0.05 + 0.65*\exp_{(600\ m)} + 0.35*\exp_{(10{,}000\ m)}$
	SQRTHT × Panchromatic	$\gamma(h) = 0.04 - 0.11*\exp_{(600\ m)} - 0.45*\exp_{(10{,}000\ m)}$
	Residuals × Panchromatic	$\gamma(h) = 0.08 - 0.05*\exp_{(600\ m)} - 0.09*\exp_{(10{,}000\ m)}$
1000 m	SQRTHT	$\gamma(h) = 0.10 + 0.35*\exp_{(600\ m)} + 0.67*\exp_{(10{,}000\ m)}$
	Residuals	$\gamma(h) = 0.20 + 0.73*\exp_{(600\ m)} + 0.09*\exp_{(10{,}000\ m)}$
	Panchromatic	$\gamma(h) = 0.05 + 0.66*\exp_{(600\ m)} + 0.34*\exp_{(10{,}000\ m)}$
	SQRTHT × Panchromatic	$\gamma(h) = 0.04 - 0.11*\exp_{(600\ m)} - 0.45*\exp_{(10{,}000\ m)}$
	Residuals × Panchromatic	$\gamma(h) = 0.08 - 0.05*\exp_{(600\ m)} - 0.09*\exp_{(10{,}000\ m)}$
500 m	SQRTHT	$\gamma(h) = 0.08 + 0.37*\exp_{(600\ m)} + 0.67*\exp_{(10{,}000\ m)}$
	Residuals	$\gamma(h) = 0.20 + 0.73*\exp_{(600\ m)} + 0.09*\exp_{(10{,}000\ m)}$
	Panchromatic	$\gamma(h) = 0.05 + 0.65*\exp_{(600\ m)} + 0.35*\exp_{(10{,}000\ m)}$
	SQRTHT × Panchromatic	$\gamma(h) = 0.04 - 0.11*\exp_{(600\ m)} - 0.45*\exp_{(10{,}000\ m)}$
	Residuals × Panchromatic	$\gamma(h) = 0.10 - 0.07*\exp_{(600\ m)} - 0.09*\exp_{(10{,}000\ m)}$
250 m	SQRTHT	$\gamma(h) = 0.08 + 0.37*\exp_{(600\ m)} + 0.67*\exp_{(10{,}000\ m)}$
	Residuals	$\gamma(h) = 0.20 + 0.72*\exp_{(600\ m)} + 0.11*\exp_{(10{,}000\ m)}$
	Panchromatic	$\gamma(h) = 0.05 + 0.68*\exp_{(600\ m)} + 0.33*\exp_{(10{,}000\ m)}$
	SQRTHT × Panchromatic	$\gamma(h) = 0.04 - 0.11*\exp_{(600\ m)} - 0.45*\exp_{(10{,}000\ m)}$
	Residuals × Panchromatic	$\gamma(h) = 0.10 - 0.07*\exp_{(600\ m)} - 0.09*\exp_{(10{,}000\ m)}$
(B) Points		
2000 m	SQRTHT	$\gamma(h) = 0.15 + 0.01*\exp_{(3000\ m)} + 0.99*\exp_{(10{,}000\ m)}$
	Residuals	$\gamma(h) = 0.90 + 0.11*\exp_{(3000\ m)} + 0.01*\exp_{(10{,}000\ m)}$
	Panchromatic	$\gamma(h) = 0.35 + 0.10*\exp_{(3000\ m)} + 0.65*\exp_{(10{,}000\ m)}$
	SQRTHT × Panchromatic	$\gamma(h) = 0.00 - 0.01*\exp_{(3000\ m)} - 0.64*\exp_{(10{,}000\ m)}$
	Residuals × Panchromatic	$\gamma(h) = 0.00 - 0.10*\exp_{(3000\ m)} - 0.01*\exp_{(10{,}000\ m)}$
1000 m	SQRTHT	$\gamma(h) = 0.37 + 0.08*\exp_{(3000\ m)} + 0.65*\exp_{(10{,}000\ m)}$
	Residuals	$\gamma(h) = 0.94 + 0.07*\exp_{(3000\ m)} + 0.01*\exp_{(10{,}000\ m)}$
	Panchromatic	$\gamma(h) = 0.53 + 0.23*\exp_{(3000\ m)} + 0.30*\exp_{(10{,}000\ m)}$
	SQRTHT × Panchromatic	$\gamma(h) = 0.00 - 0.13*\exp_{(3000\ m)} - 0.44*\exp_{(10{,}000\ m)}$
	Residuals × Panchromatic	$\gamma(h) = 0.07 - 0.12*\exp_{(3000\ m)} - 0.03*\exp_{(10{,}000\ m)}$
500 m	SQRTHT	$\gamma(h) = 0.34 + 0.08*\exp_{(1000\ m)} + 0.70*\exp_{(10{,}000\ m)}$
	Residuals	$\gamma(h) = 0.70 + 0.25*\exp_{(1000\ m)} + 0.05*\exp_{(10{,}000\ m)}$
	Panchromatic	$\gamma(h) = 0.50 + 0.27*\exp_{(1000\ m)} + 0.27*\exp_{(10{,}000\ m)}$
	SQRTHT × Panchromatic	$\gamma(h) = 0.00 - 0.10*\exp_{(1000\ m)} - 0.41*\exp_{(10{,}000\ m)}$
	Residuals × Panchromatic	$\gamma(h) = 0.08 - 0.06*\exp_{(1000\ m)} - 0.10*\exp_{(10{,}000\ m)}$
250 m	SQRTHT	$\gamma(h) = 0.18 + 0.25*\exp_{(600\ m)} + 0.69*\exp_{(10{,}000\ m)}$
	Residuals	$\gamma(h) = 0.30 + 0.62*\exp_{(600\ m)} + 0.12*\exp_{(10{,}000\ m)}$
	Panchromatic	$\gamma(h) = 0.30 + 0.45*\exp_{(600\ m)} + 0.29*\exp_{(10{,}000\ m)}$
	SQRTHT × Panchromatic	$\gamma(h) = 0.09 - 0.19*\exp_{(600\ m)} - 0.38*\exp_{(10{,}000\ m)}$
	Residuals × Panchromatic	$\gamma(h) = 0.12 - 0.10*\exp_{(600\ m)} - 0.07*\exp_{(10{,}000\ m)}$

deviations in the estimated height histograms from the measured height histogram (Fig. 5). Furthermore, correlations between measured and estimated height values in the scatterplots agreed well with the correlations calculated from the global height estimates (Fig. 5). It is thus safe to conclude that the 700 points in these scatterplots were highly representative of the full population of height estimates, and their errors.

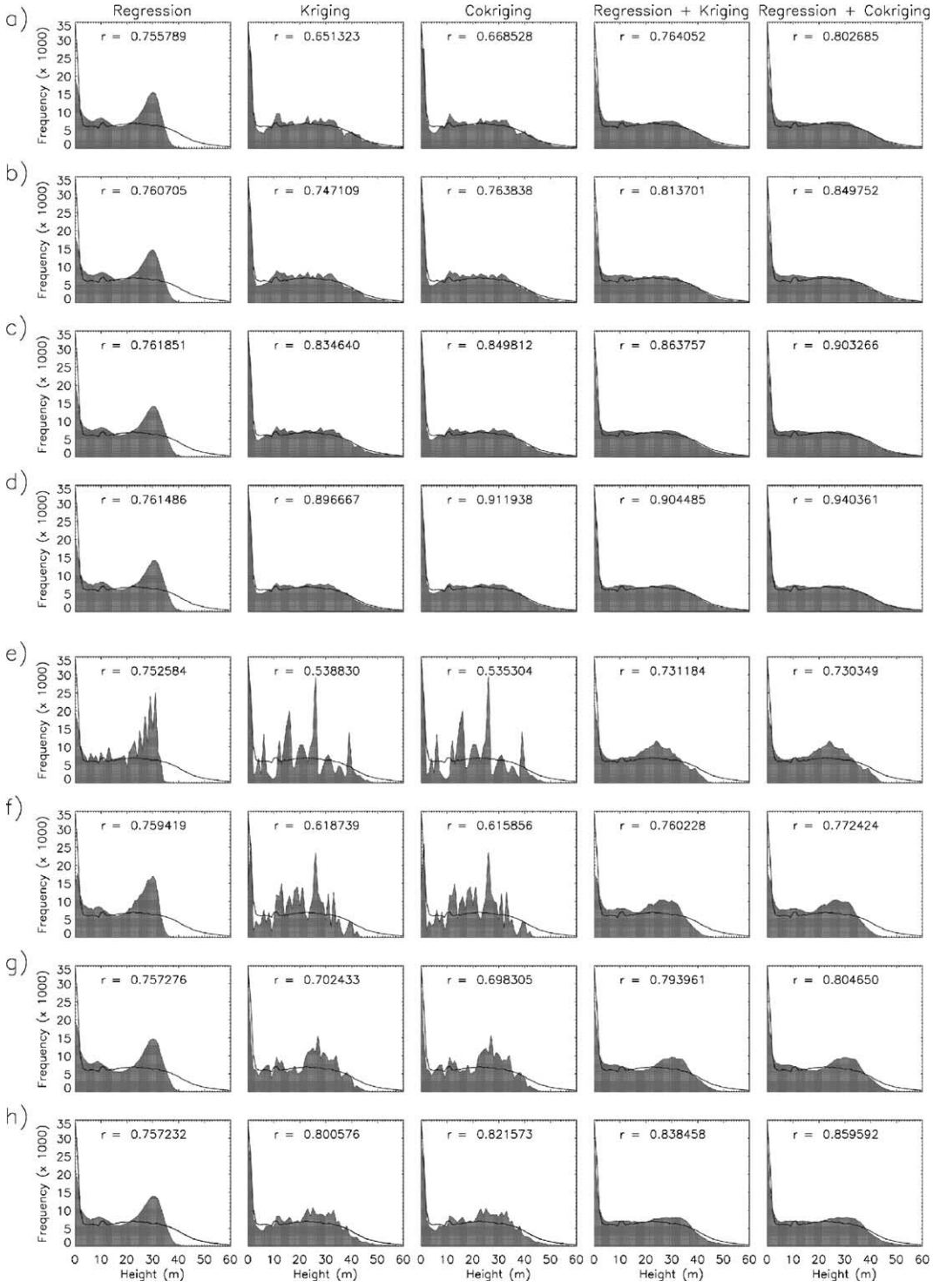

Fig. 5. Histograms of the entire population of estimated height values (shaded in gray, $N = 337{,}464$) from the five models tested, for the eight sampling strategies tested: (a) 2000-m transect, (b) 1000-m transect, (c) 500-m transect, (d) 250-m transect, (e) 2000-m point, (f) 1000-m point, (g) 500-m point, and (h) 250-m point. The outline of the measured height value histogram is plotted over each estimated height value histogram for comparison.

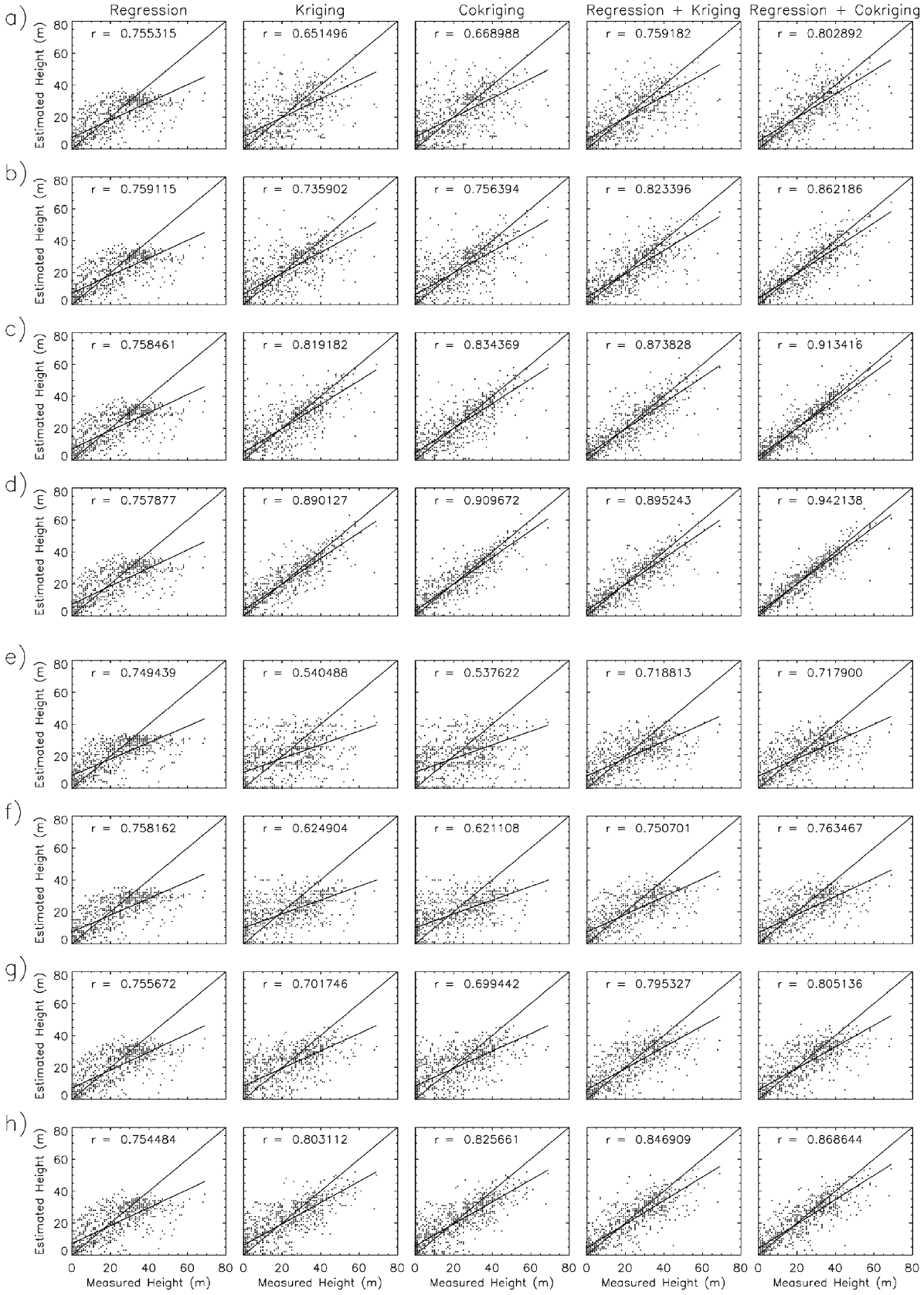

Fig. 6. Scatterplots of measured vs. estimated height values from the five models tested, for the eight sampling strategies tested: (a) 2000-m transect, (b) 1000-m transect, (c) 500-m transect, (d) 250-m transect, (e) 2000-m point, (f) 1000-m point, (g) 500-m point, and (h) 250-m point. Locations of the plotted values ($N = 700$) are shown in Fig. 3.

4.2.2. Local

Local estimation accuracy also was assessed according to Pearson's correlation statistic. Accuracy decreased as the distance from sample locations increased (Fig. 7). The spatial models were more accurate than the regression models below distances of approximately 200 m from the sample locations. The integrated models preserved the accuracy of the regression estimates beyond this distance to the nearest sample. A sampling interval of 250 m ensured that all estimates were < 180 m from the nearest sample (i.e.,

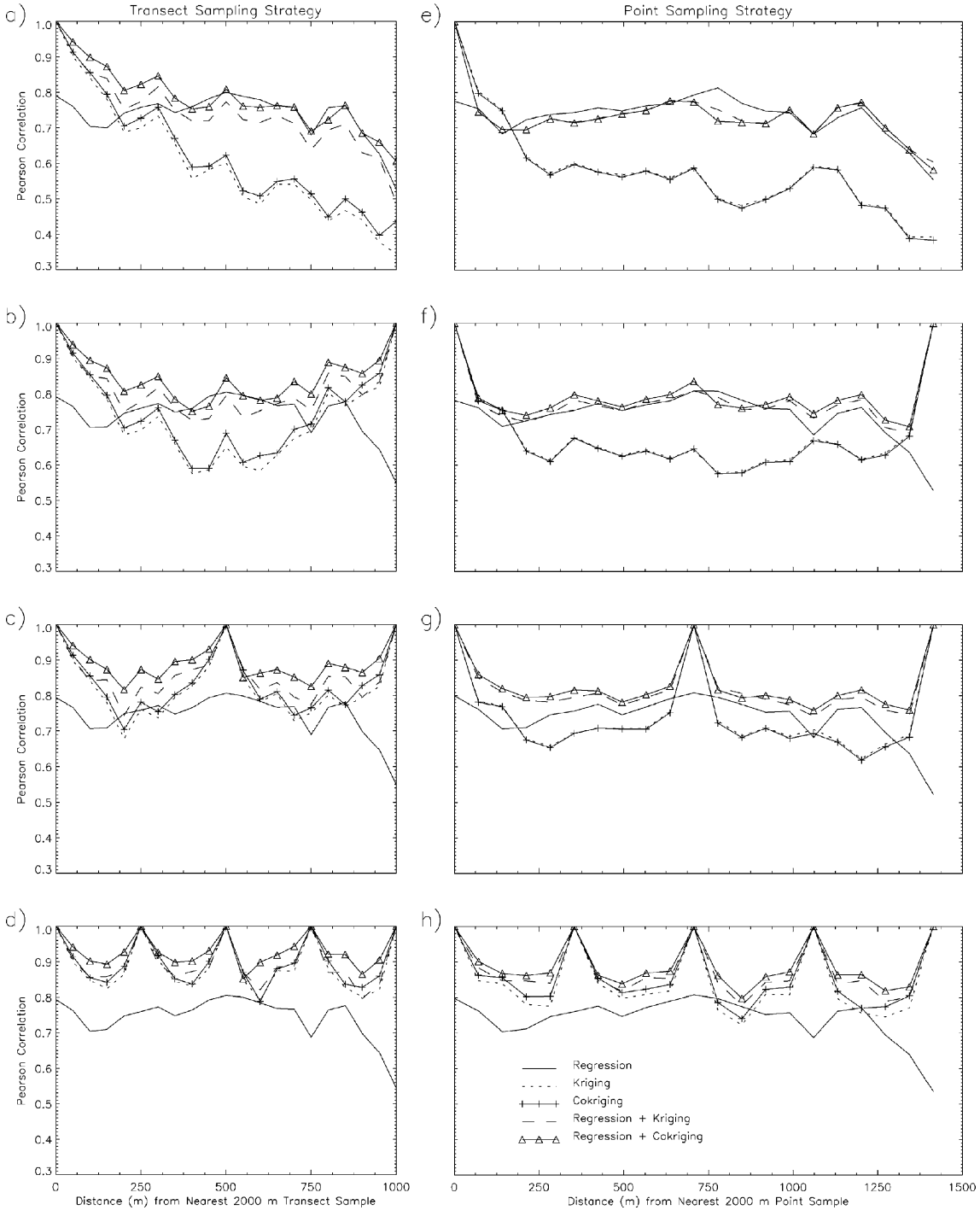

Fig. 7. Distance vs. Pearson's correlation coefficient for the eight sampling strategies tested: (a) 2000-m transect, (b) 1000-m transect, (c) 500-m transect, (d) 250-m transect, (e) 2000-m point, (f) 1000-m point, (g) 500-m point, and (h) 250-m point. The validation points are farther from the nearest point sample location than from the nearest transect sample location by a factor of $\sqrt{2}$. The graphed value at each distance is based on $n = 180$ points, except $n = 60$ at 0 m, and $n = 45$ at 1000 (plots a–d) or 1414 m (plots e–h). Perfect correlations result in the spatial and integrated models where the validation and sample data locations intersect. Locations of the plotted values ($N = 3525$) are shown in Fig. 3.

below the range of the semivariograms; see Fig. 4), which improved estimation accuracies of the spatial and integrated models above those of regression, at all locations.

4.3. Mapping

Regression-based maps (Fig. 8) were virtually indistinguishable regardless of the sampling strategy (Fig. 2) or number of variables included (Table 1). In dramatic contrast, the sampling strategy caused obvious artifacts in the kriging or cokriging maps that were most pronounced at the sparser sampling frequencies. These artifacts were, however, greatly attenuated in the maps produced from the integrated models. The kriging and cokriging maps were virtually indistinguishable when the same primary data were modeled.

Maps of estimation errors (Fig. 9) were produced by subtracting the actual height map (Fig. 1) from the estimated height maps (Fig. 8). Overall, every model underestimated canopy height, although the estimation bias was an order of magnitude greater for the regression models than for any of the spatial or integrated models (Table 3). The standard deviation of the estimation errors for the spatial and inte-

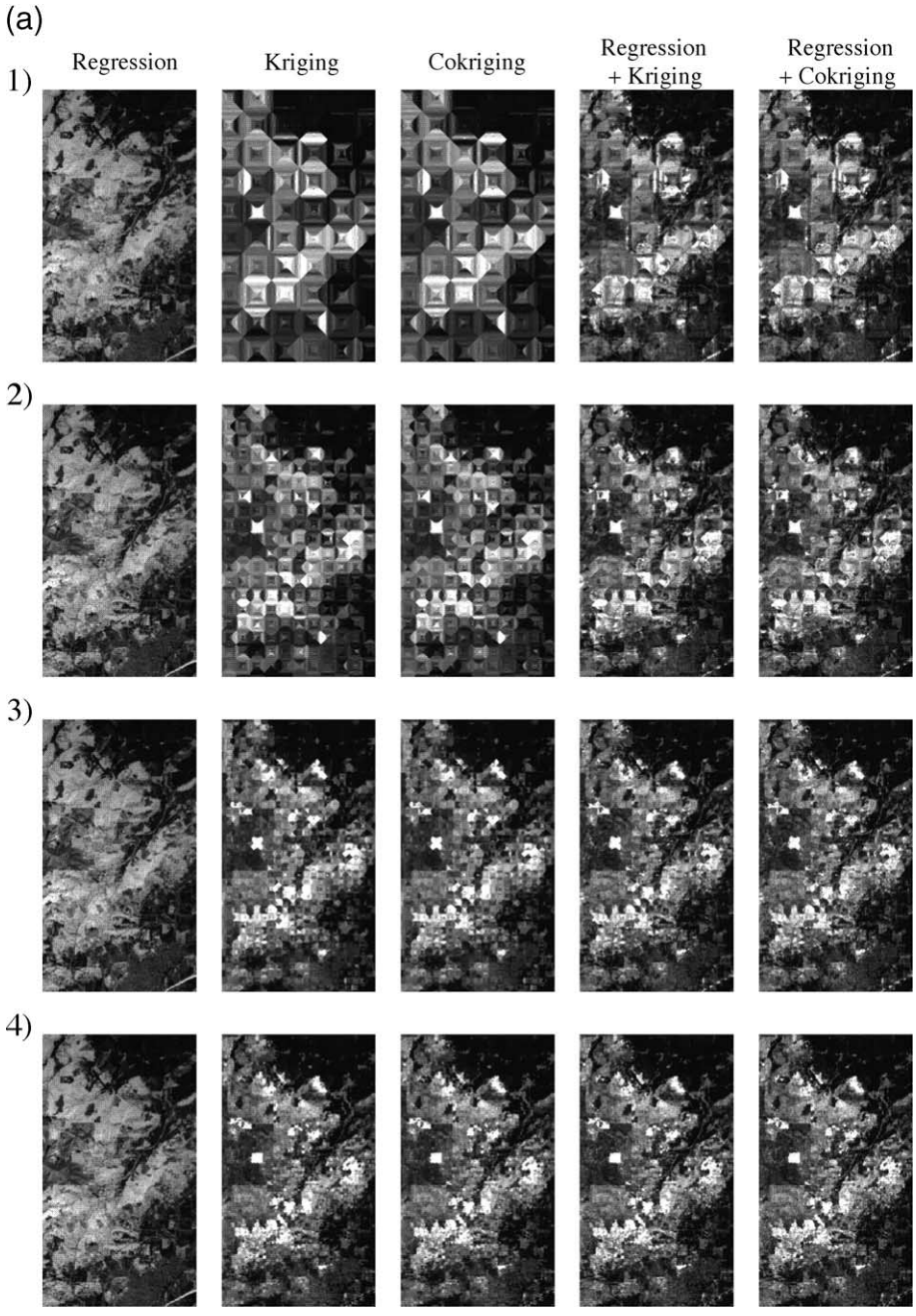

Fig. 8. (a) Estimated height maps from the five models tested, for the four transect sampling strategies: (1) 2000, (2) 1000, (3) 500, (4) 250 m, and (b) for the four point sampling strategies: (1) 2000, (2) 1000, (3) 500, and (4) 250 m. Brightness values are scaled to height values as in Fig. 1.

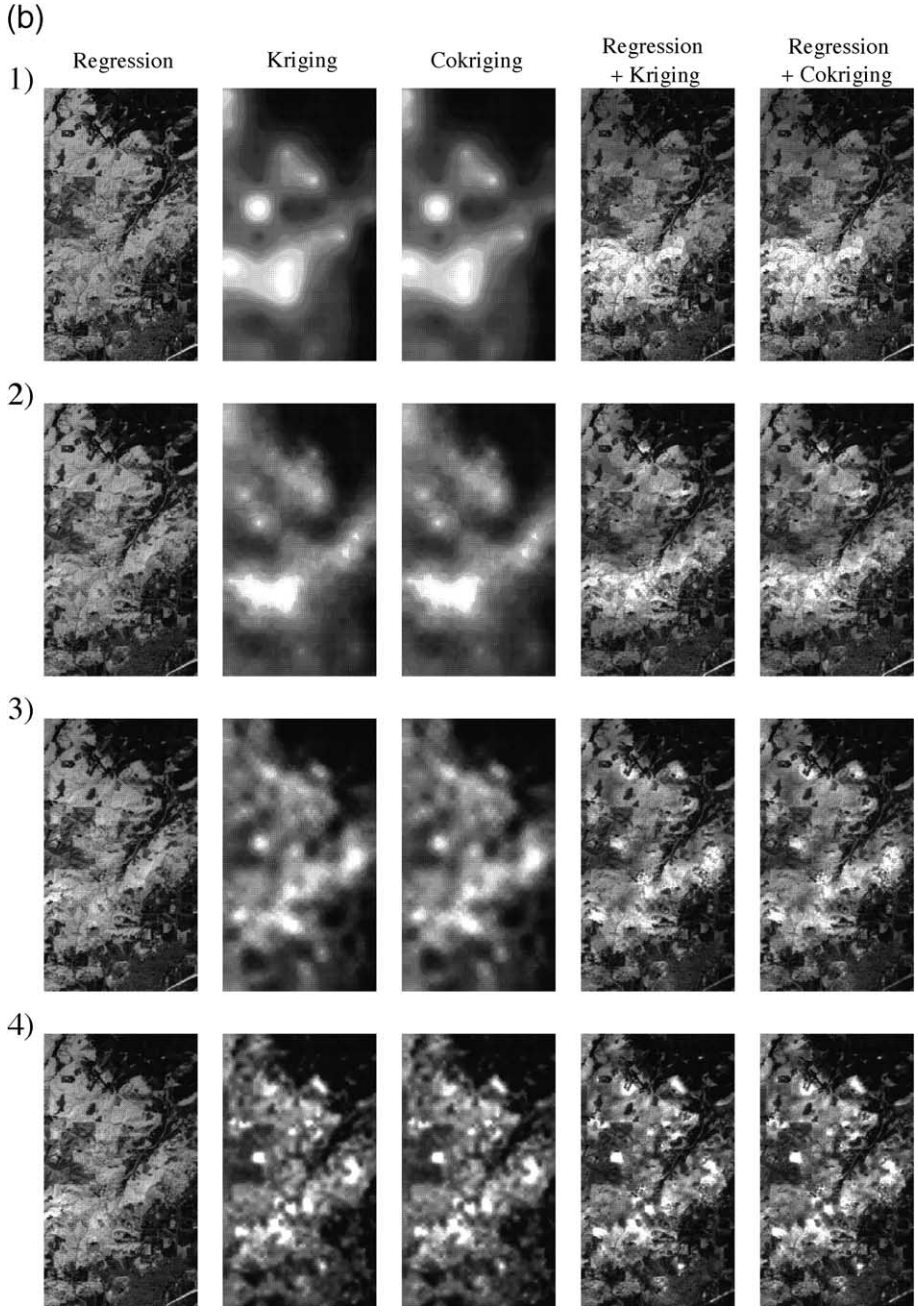

Fig. 8 (*continued*).

grated models decreased as the spatial sampling frequency increased.

Spatial patterns in the error maps for the spatial and integrated models became less apparent as sampling density increased, while sampling density had no effect on error patterns for the aspatial regression models (Fig. 9). Moran's *I* statistic was useful for quantifying the significance of the spatial autocorrelation remaining in the height estimation errors for all models. All regression models, and all models derived from the two sparser point sample datasets (2000 and 1000 m), failed to remove the spatial dependence from the residuals (Table 3). The spatial models applied to the 2000-m transect sample dataset also left significant spatial autocorrelation in the residual variance, although the integrated models did not. All other models successfully accounted for spatial autocorrelation in the sample data.

5. Discussion

5.1. Aspatial Models

The high similarity among all regression estimates of height (Figs. 5–8) indicates the insensitivity of the regres-

sion models to sample size, sampling pattern, sampling frequency, or number of ETM+ bands selected (Fig. 2, Table 1). Regression suffered the worst from a consistent estimation bias (Table 3), overestimating shorter stands while underestimating taller stands (Figs. 5, 6, and 9). This effect is discussed in detail (as variance ratio) by Cohen et al. (in press). On the other hand, regression did preserve the spatial pattern of stands across the study landscape (Figs. 1 and 8).

We included the UTMX and UTMY location variables in the regression models as an easy way to account for a potential geographic trend across our study area, following the approach of Metzger (1997). Yet most of the height data variance explainable with regression were explained by ETM+ Band 7 alone (Table 1, Figs. 5 and 6). The location variables (particularly UTMY) were selected by some of the stepwise regression models but only for those sampling strategies with a high data volume (Fig. 2). In these cases,

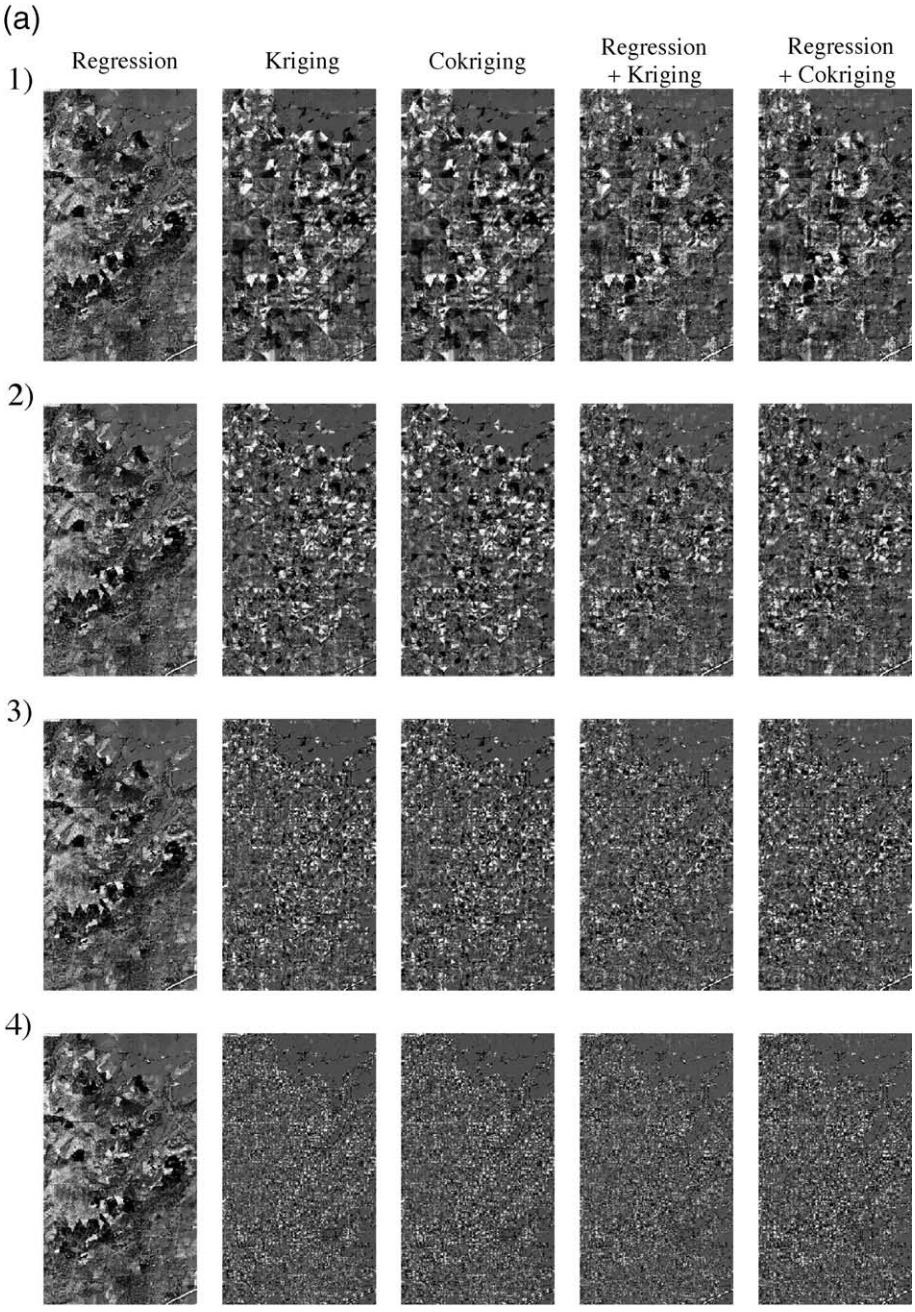

Fig. 9. (a) Estimated height error maps from the five models tested, for the four transect sampling strategies: (1) 2000, (2) 1000, (3) 500, and (4) 250 m, and (b) for the four point sampling strategies: (1) 2000, (2) 1000, (3) 500, and (4) 250 m. Bright areas are overestimates while dark areas are underestimates; see Table 3 for error magnitudes.

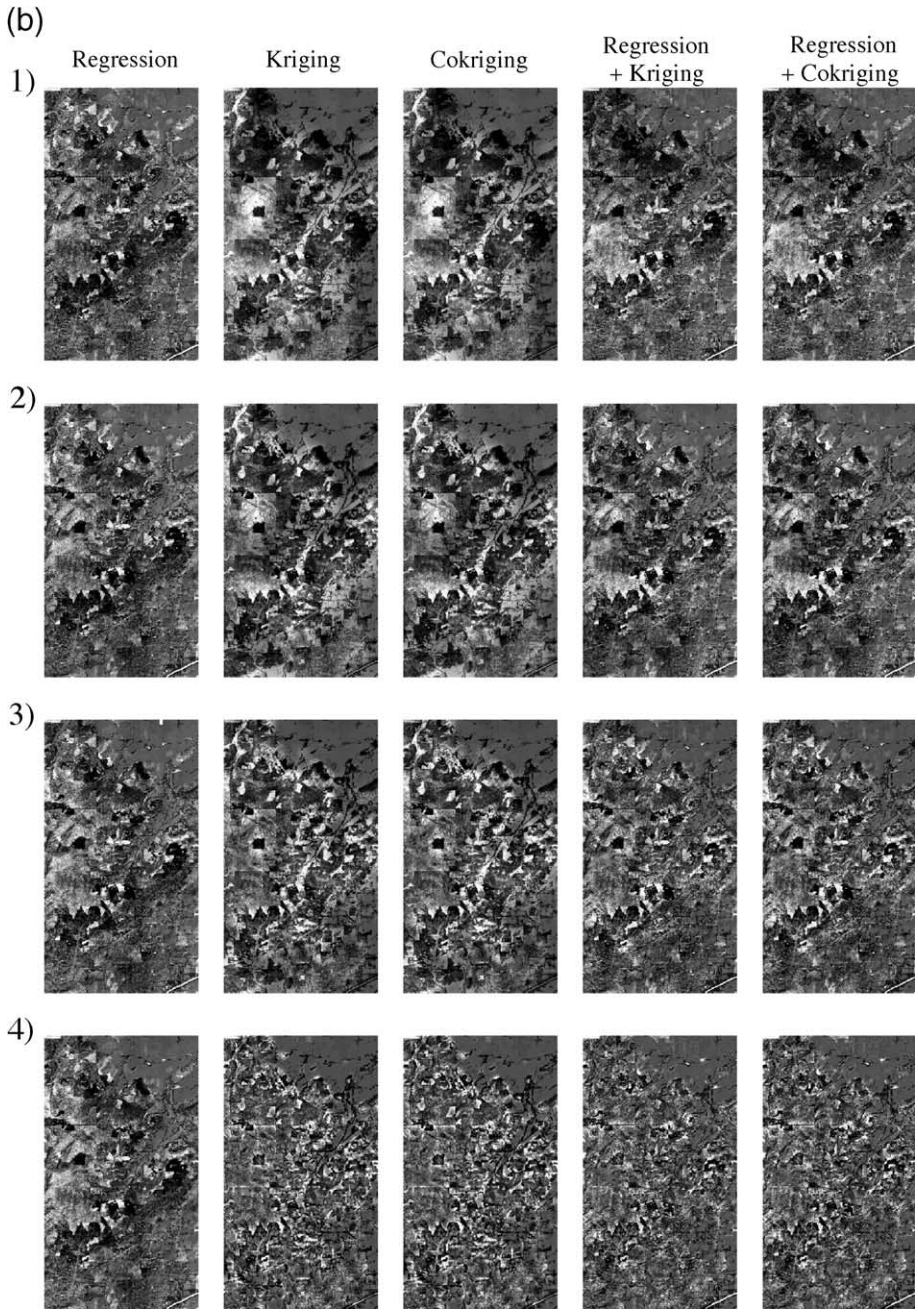

Fig. 9 (*continued*).

the addition of the location variables and other ETM+ bands as explanatory variables carried statistical significance but probably lacked biological significance over the small spatial extent studied.

5.2. Spatial Models

In stark contrast to regression, height estimates from the spatial methods were only slightly biased (Figs. 5 and 6, Table 3), but were highly sensitive to sampling pattern and frequency (Fig. 2), which produced spatial discontinuities in the resulting maps (Fig. 8). These discontinuities were visually distracting when the modeled variable (canopy height in this case) was undersampled relative to the spatial frequency at which it actually varies; the semivariograms indicate that the range of spatial autocorrelation in canopy height is no more than 500 m in this landscape (Fig. 4). Beyond 500 m from the nearest sample, the semivariograms carried little or no weight in the estimation; this produced the smoothing effect visible especially in the 2000- and 1000-m kriged/cokriged maps (Fig. 8). At sampling intervals of 500 or 250 m, all estimates were at, or below, the

Table 3
(A) Mean and (B) S.D. of the residuals (meters) at the gridded validation points ($N = 700$), and (C) P value of Moran's I test for spatial dependence; P values in bold indicate that significant ($\alpha = 0.05$) spatial autocorrelation remains in the residual variance

	Regression	Kriging	Cokriging	Regression+Kriging	Regression+Cokriging
(A) Mean					
2000-m transect	− 1.7757	− 0.3143	− 0.2257	− 0.5286	− 0.2900
1000-m transect	− 1.7686	− 0.3043	− 0.2143	− 0.7371	− 0.4400
500-m transect	− 1.5271	− 0.1386	− 0.0257	− 0.4386	− 0.1043
250-m transect	− 1.4386	− 0.2157	− 0.1457	− 0.4443	− 0.1500
2000-m point	− 1.8014	− 1.4257	− 1.4286	− 1.5486	− 1.5414
1000-m point	− 2.1271	− 1.2929	− 1.2986	− 1.6843	− 1.6900
500-m point	− 1.6271	− 0.6029	− 0.6229	− 0.6229	− 0.6200
250-m point	− 1.6386	− 0.5900	− 0.5657	− 0.7014	− 0.6286
(B) S.D.					
2000-m transect	10.3101	12.5512	12.2680	10.4557	9.5455
1000-m transect	10.2471	10.8950	10.5106	8.9759	8.0058
500-m transect	10.2468	9.1390	8.7843	7.6854	6.4406
250-m transect	10.2561	7.1690	6.5299	7.0124	5.2700
2000-m point	10.4529	13.8177	13.8534	10.9392	10.9550
1000-m point	10.2999	12.3177	12.3704	10.3862	10.1577
500-m point	10.2957	11.2657	11.3053	9.5506	9.3366
250-m point	10.3173	9.3750	8.8693	8.3667	7.7890
(C) P value of Moran's I					
2000-m transect	**0.0000**	**0.0047**	**0.0040**	0.0618	0.1168
1000-m transect	**0.0000**	0.1303	0.1140	0.9353	0.7139
500-m transect	**0.0000**	0.3477	0.4307	0.2702	0.2752
250-m transect	**0.0000**	0.6157	0.6081	0.4824	0.2438
2000-m point	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
1000-m point	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
500-m point	**0.0000**	0.1519	0.1794	0.1474	0.1571
250-m point	**0.0000**	0.7851	0.7926	0.6611	0.8508

range of spatial autocorrelation for this landscape (Figs. 4 and 7), so little smoothing occurred.

Stein and Corsten (1991) found that kriging and cokriging estimates differ only slightly from each other, and that the advantage of cokriging is greater when a highly correlated secondary variable is sampled intensively. We also found cokriging only slightly more advantageous than kriging at all sampling frequencies, perhaps because canopy height and the ETM+ panchromatic band were only weakly correlated ($r = -0.43$).

5.3. Integrated method

Most of the biases in the regression estimates were eliminated in the integrated models, where the regression residuals were subsequently kriged and added back to the regression surface (Figs. 5 and 6, Table 3). We found the advantage of cokriging over kriging to be greater with the height residuals than with the height values (Figs. 5–7, Table 3). Perhaps because the regression models explain such a large proportion of the total variation in canopy height ($r^2 = 0.58$), the height residuals may correspond more closely than the height values to the fine scale structural features in the panchromatic image.

The integrated methods proved superior because they preserved the spatial pattern in canopy height, like the regression models (Fig. 8), while also improving global and local estimation accuracy, like the spatial models (Figs. 5–7). They have no apparent disadvantage relative to aspatial or spatial methods alone (Table 3).

The estimation methods applied to lidar canopy height data in this analysis are applicable to field data, as has already been demonstrated by Atkinson, Webster, and Curran (1992, 1994). The samples need not be situated along a systematic grid; the methods are as applicable to random or subjective sampling strategies, as long as the samples represent the population in both statistical and geographical space, and, for spatial methods, are dense enough to capture the range of the semivariograms.

5.4. Alternative modeling techniques

As an aspatial estimation method, inverse regression models (Curran & Hay, 1986) should be considered when the explanatory variables are dependent on the variable of interest. Surface radiance is influenced by canopy height, however, Landsat imagery is much more sensitive to the spectral properties of the surface materials than to their height. Another criticism of regression models is that they account for errors in only one set of variables (e.g., Landsat bands) and assume a lack of measurement error in the variable of interest (e.g., lidar height). All remotely sensed

data including lidar are subject to several sources of error: irradiance variation, sensor calibration, sensor radiometric resolution, sensor drift, signal digitization, atmospheric attenuation, and atmospheric path radiance. An alternative approach that accounts for errors in both the independent and dependent variables is reduced major axis (RMA) regression (Cohen et al., in press; Curran & Hay, 1986). Regardless of the regression method selected, we argue against using regression models alone to estimate canopy height. Our regression equations were useful for explaining a large proportion of the total variance in canopy height due to high covariance with measured radiance, but not due to any functional relationship. As stated in our objectives, we considered it most useful to present the most commonly used techniques for this paper, and OLS regression is clearly the standard empirical modeling tool.

With regard to spatial estimation methods, OK or OCK is advisable only in interpolation situations such as in this study; in extrapolation situations, it may be better to use universal kriging (Journel & Rossi, 1989; Stein & Corsten, 1991) or OK with an external drift (Berterretche, 2001). In cases where anisotropy exists in the landscape, anisotropic kriging models having a directional component can be employed. Goovaerts (1997) thoroughly presents the many kriging/cokriging procedures available.

For mapping, conditional simulation can be a good alternative to the estimation methods presented here (Dungan, 1998, 1999). Conditional simulation "conditions" stochastic predictions of the modeled variable within the spatial range of the sample data, as defined by the same semivariogram model used for kriging. Although locally inaccurate, conditional simulation preserves the global accuracy and spatial pattern of the data modeled. These qualities can be important for some applications, such as modeling variables as input for ecological process models. We ran conditional stochastic simulations of canopy height, and height residuals, from our eight sample datasets. In every case, local accuracy was markedly lower than for any of the estimation methods we tested. Since local accuracy was important for our objectives, while multiple realizations were not, we pursued simulation methods no further for this paper. The decision of which estimation or simulation methods to use for modeling height or any other structural variable ultimately depends on user objectives.

5.5. Sampling strategy

Traditionally, most remote sensors have afforded analysts with a certain luxury by sampling the entire population within the extent of coverage. This has precluded any need to apply spatial interpolation strategies such as kriging, yet imagery is full of underexploited spatial information. A number of studies have demonstrated the value of geostatistical analysis tools such as semivariograms (e.g., Cohen et al., 1990; Curran, 1988; Glass, Carr, Yang, & Myers, 1988; Hudak & Wessman, 1998; Woodcock, Strahler, & Jupp, 1988). As remote sensing technology has advanced towards increasing spectral, spatial, and temporal resolution, data processing and storage technologies have kept pace, enabling the continued availability of comprehensive data even as those data volumes have exponentially increased. While these trends may very well continue, it is instructive and useful to consider the applicability of future remote sampling instruments for estimating and mapping forest attributes.

We found that sampling frequency is critical for mapping of managed, high-biomass coniferous forests of the Pacific Northwest, where forest structure predominantly varies at the scale of individual stands with spatial frequencies of <500 m (Milne & Cohen, 1999). Transect sampling consistently produced more accurate results than point sampling but at the cost of data volumes that were orders of magnitude higher. The 250-m point sampling strategy was the most efficient of those tested in our study area, but we caution that the optimal sampling frequency probably varies tremendously for different forests around the globe.

The objectives of the VCL and ICESat missions, as described in the Introduction, are necessarily much broader than those of this study. We sought not to simulate the VCL or ICESat sampling designs but to more generally explore the effects of sampling pattern (transect vs. point) and frequency (2000, 1000, 500, and 250 m) on estimating and mapping a basic forest structural attribute (height). Our analyses suggest that in dense, managed forests, the wide spacing of VCL transects would likely be problematic for mapping structural attributes using geostatistical integration models at landscape scales. However, at regional–global scales, the contiguous nature of the VCL transects should provide a valuable, unbiased inventory of the scale of variation in forest canopy structure. We also anticipate some disadvantages and advantages of ICESat data for vegetation mapping. One difficulty of ICESat data may be the 70-m footprint diameter, which is likely too large for accurately measuring forest canopy height in areas with steep slopes. On the other hand, the 175-m spacing of the samples may be advantageous for applying geostatistical integration models for mapping forest canopy height in flat, forested regions as in Canada, Siberia, and the Amazon and Congo Basins.

6. Conclusion

Integration of lidar and Landsat ETM+ data using straightforward empirical modeling procedures can be used to improve the utility of both datasets for forestry applications. In this study, an integrated technique of ordinary cokriging of the height residuals from an OLS regression model proved the best method for estimating and mapping forest canopy height, and an equitable distribution of lidar sampling points proved critical for efficient lidar–Landsat

ETM+ integration. We encourage testing of integration models in a variety of ecosystems once lidar sample data become readily available.

Acknowledgements

This work was funded by the NASA Terrestrial Ecology Program (NRA-97-MTPE-08), through the Enhanced Thematic Mapper Plus Lidar for Forested Ecosystems (ETM + LIFE) Project. Aeroscan data were provided by Mike Renslow of Spencer B. Gross, Portland, OR. We also thank Ralph Dubayah for his critical review and Jeff Evans for graphics assistance.

References

Adams, J. B., Sabol, D. E., Kapos, V., Filho, R. A., Roberts, D. A., Smith, M. O., & Gillespie, A. R. (1995). Classification of multispectral images based on fractions of endmembers: application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment*, *52*, 137–154.

Asner, G. P., Wessman, C. A., & Privette, J. L. (1997). Unmixing the directional reflectances of AVHRR sub-pixel landcovers. *IEEE Transactions on Geoscience and Remote Sensing*, *35*, 868–878.

Asrar, G., Fuchs, M., Kanemasu, E. T., & Hatfield, J. L. (1984). Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat. *Agronomy Journal*, *76*, 300–306.

Atkinson, P. M., Webster, R., & Curran, P. J. (1992). Cokriging with ground-based radiometry. *Remote Sensing of Environment*, *41*, 45–60.

Atkinson, P. M., Webster, R., & Curran, P. J. (1994). Cokriging with airborne MSS imagery. *Remote Sensing of Environment*, *50*, 335–345.

Bachman, C. G. (1979). *Laser radar systems and techniques* (193 pp.). Norwood, MA: Artech House.

Berterretche, M. (2001). *Comparison of regression and geostatistical methods to develop LAI surfaces for NPP modeling.* Master's thesis (169 pp.). Corvallis, OR: Oregon State University.

Chen, J. M., & Cihlar, J. (1996). Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sensing of Environment*, *55*, 153–162.

Cliff, A. D., & Ord, J. K. (1981). *Spatial processes: models and applications* (266 pp.). London: Pion.

Cohen, W. B., Maierpserger, T. K., Gower, S. T., & Turner, D. P. (in press). An improved strategy for regression of biophysical variables and Landsat ETM+ data. *Remote Sensing of Environment*.

Cohen, W. B., & Spies, T. A. (1992). Estimating structural attributes of Douglas-fir/western hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment*, *41*, 1–17.

Cohen, W. B., Spies, T. A., & Bradshaw, G. A. (1990). Semivariograms of digital imagery for analysis of conifer canopy structure. *Remote Sensing of Environment*, *34*, 167–178.

Cohen, W. B., Spies, T. A., & Fiorella, M. (1995). Estimating the age and structure of forests in a multi-ownership landscape of western Oregon, USA. *International Journal of Remote Sensing*, *16*, 721–746.

Curran, P. J. (1988). The semivariogram in remote sensing: an introduction. *Remote Sensing of Environment*, *24*, 493–507.

Curran, P. J., & Hay, A. M. (1986). The importance of measurement error for certain procedure in remote sensing of optical wavelengths. *Photogrammetric Engineering and Remote Sensing*, *52*, 229–241.

Deutsch, C. V., & Journel, A. G. (1998). *GSLIB geostatistical software library and user's guide* (369 pp.). New York: Oxford University Press.

Drake, J. B., Dubayah, R. O., Clark, D. B., Knox, R. G., Blair, J. B., Hofton, M. A., Chazdon, R. L., Weishampel, J. F., & Prince, S. D. (2002). Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment*, *79*, 305–319.

Dubayah, R., Blair, J. B., Bufton, J. L., Clark, D. B., JaJa, J., Knox, R., Luthcke, S. B., Prince, S., & Weishampel, J. (1997). The Vegetation Canopy Lidar mission. *Land satellite information in the next decade: II. Sources and applications* (pp. 100–112). Washington, DC: ASPRS.

Dungan, J. L. (1998). Spatial prediction of vegetation quantities using ground and image data. *International Journal of Remote Sensing*, *19*, 267–285.

Dungan, J. L. (1999). Conditional simulation: an alternative to estimation for achieving mapping objectives. In A. Stein, F. van der Meer, & B. Gorte (Eds.), *Spatial statistics for remote sensing* (pp. 135–152). Dordrecht: Kluwer Academic Publishing.

Fassnacht, K. S., Gower, S. T., MacKenzie, M. D., Nordheim, E. V., & Lillesand, T. M. (1997). Estimating the LAI of North Central Wisconsin forests using the Landsat Thematic Mapper. *Remote Sensing of Environment*, *61*, 229–245.

Friedl, M. A., Michaelsen, J., Davis, F. W., Walker, H., & Schimel, D. S. (1994). Estimating grassland biomass and leaf area index using ground and satellite data. *International Journal of Remote Sensing*, *15*, 1401–1420.

Glass, C. E., Carr, J. R., Yang, H.-M., & Myers, D. E. (1988). Application of spatial statistics to analyzing multiple remote sensing data sets. In A. I. Johnson, & C. B. Pettersson (Eds.), *Geotechnical applications of remote sensing and remote data transmission, ASTM special technical publication*, *vol.* 967 (pp. 138–150). Philadelphia: American Society for Testing and Materials.

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. New York: Oxford University Press (483 pp.) .

Hudak, A. T., & Wessman, C. A. (1998). Textural analysis of historical aerial photography to characterize woody plant encroachment in South African savanna. *Remote Sensing of Environment*, *66*, 317–330.

Isaaks, E. H., & Srivastava, R. M. (1989). *Applied geostatistics* (561 pp.). New York: Oxford University Press.

Journel, A. G., & Rossi, M. E. (1989). When do we need a trend model in kriging? *Mathematical Geology*, *21*, 715–739.

Kennedy, R. E., & Cohen, W. B. (in press). Automated designation of tiepoints for image-to-image registration. *International Journal of Remote Sensing*.

Kleinbaum, D. G., Kupper, L. L., Muller, K. E., & Nizam, A. (1998). *Applied regression analysis and other multivariable methods* (3rd ed.) (736 pp.). Pacific Grove, CA: Duxbury Press.

Knotters, M., Brus, D. J., & Voshaar, J. H. O. (1995). A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. *Geoderma*, *67*, 227–246.

Law, B. E., & Waring, R. H. (1994). Remote sensing of leaf area index and radiation intercepted by understory vegetation. *Ecological Applications*, *4*, 272–279.

Lefsky, M. A., Cohen, W. B., Acker, S. A., Parker, G. G., Spies, T. A., & Harding, D. (1999). Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment*, *70*, 339–361.

Lefsky, M. A., Cohen, W. B., Hudak, A. T., Acker, S. A., & Ohmann, J. L. (1999). Integration of lidar, Landsat ETM+ and forest inventory data for regional forest mapping. *International Archives of Photogrammetry and Remote Sensing*, *32*, 119–126 (Part 3W14).

Lefsky, M. A., Cohen, W. B., Parker, G. G., & Harding, D. J. (2002). Lidar remote sensing for ecosystem studies. *Bioscience*, *52*, 19–30.

Lefsky, M. A., Cohen, W. B., & 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*, *31*, 78–87.

Lefsky, M. A., Harding, D., Cohen, W. B., Parker, G., & Shugart, H. H. (1999). Surface lidar remote sensing of basal area and biomass in de-

ciduous forests of Eastern Maryland, USA. *Remote Sensing of Environment*, *67*, 83–98.

Means, J. E., Acker, S. A., Fitt, B. J., Renslow, M., Emerson, L., & Hendrix, C. J. (2000). Predicting forest stand characteristics with airborne scanning lidar. *Photogrammetric Engineering and Remote Sensing*, *66*, 1367–1371.

Means, J. E., Acker, S. A., Harding, D. J., Blair, J. B., Lefsky, M. A., Cohen, W. B., Harmon, M. E., & McKee, W. A. (1999). Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the Western Cascades of Oregon. *Remote Sensing of Environment*, *67*, 298–308.

Metzger, K. L. (1997). *Modeling forest stand structure to a ten meter resolution using Landsat TM data.* Master's thesis (123 pp.). Fort Collins, CO: Colorado State University.

Milne, B. T., & Cohen, W. B. (1999). Multiscale assessment of binary and continuous landcover variables for MODIS validation, mapping, and modeling applications. *Remote Sensing of Environment*, *70*, 82–98.

Moran, P. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society*, *10B*, 243–251.

Nelson, R., Krabill, W., & Maclean, G. (1984). Determining forest canopy characteristics using airborne laser data. *Remote Sensing of Environment*, *15*, 201–212.

Nelson, R., Oderwald, R., & Gregoire, T. G. (1997). Separating the ground and airborne laser sampling phases to estimate tropical forest basal area, volume, and biomass. *Remote Sensing of Environment*, *60*, 311–326.

Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment*, *56*, 1–7.

Oleson, K. W., Sarlin, S., Garrison, J., Smith, S., Privette, J. L., & Emery, W. J. (1995). Unmixing multiple land-cover type reflectances from coarse spatial resolution satellite data. *Remote Sensing of Environment*, *54, 98–112.*

Pan, G., Gaard, D., Moss, K., & Heiner, T. (1993). A comparison between cokriging and ordinary kriging: case study with a polymetallic deposit. *Mathematical Geology*, *25*, 377–398.

Peddle, D. R., Hall, F. G., & LeDrew, E. F. (1999). Spectral mixture analysis and geometric–optical reflectance modeling of boreal forest biophysical structure. *Remote Sensing of Environment*, *67*, 288–297.

Stein, A., & Corsten, L. C. A. (1991). Universal kriging and cokriging as a regression procedure. *Biometrics*, *47*, 575–587.

Turner, D. P., Cohen, W. B., Kennedy, R. E., Fassnacht, K. S., & Briggs, J. M. (1999). Relationships between LAI and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing of Environment*, *70*, 52–68.

Weishampel, J. F., Blair, J. B., Knox, R. G., Dubayah, R., & Clark, D. B. (2000). Volumetric lidar return patterns from an old-growth tropical rainforest canopy. *International Journal of Remote Sensing*, *21*, 409–415.

Wiegand, C. L., Richardson, A. J., & Kanemasu, E. T. (1979). Leaf area index estimates for wheat from LANDSAT and their implications for evapotranspiration and crop modeling. *Agronomy Journal*, *71*, 336–342.

Woodcock, C. E., Strahler, A. H., & Jupp, D. L. B. (1988). The use of variograms in remote sensing: II. Real digital images. *Remote Sensing of Environment*, *25*, 349–379.