

Integrating airborne LiDAR and space-borne radar via multivariate kriging to estimate above-ground biomass.

Olivier W. Tsui^{a*}, Nicholas C. Coops^a, Michael A. Wulder^b, Peter L. Marshall^c

^a Olivier W. Tsui, Integrated Remote Sensing Studio, Forest Resources Management Department, University of British Columbia, 2424 Main Mall, Vancouver, British Columbia, V6T 1Z4, Canada. Phone: +1 604.822.6592. Email: Olivier.tsui@alumni.ubc.ca

^a Nicholas C. Coops, Integrated Remote Sensing Studio, Forest Resources Management Department, University of British Columbia, 2424 Main Mall, Vancouver, British Columbia, V6T 1Z4, Canada. Phone: +1 604.822.6592. Email: nicholas.coops@ubc.ca

^b Michael A. Wulder, Pacific Forestry Centre, Canadian Forest Service, Natural Resources Canada, 506 West Burnside Road, Victoria, BC, V8Z 1M5, Canada. Email: Mike.Wulder@nrcan-rncan.gc.ca

^c Peter L. Marshall, Faculty of Forestry, Forest Resources Management Department, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada. Email: peter.marshall@ubc.ca

*Corresponding author:

Olivier W. Tsui (email: olivier.tsui@alumni.ubc.ca)

Integrated Remote Sensing Studio

2424 Main Mall, Vancouver, BC V6T 1Z4

Tel: +1 604.822.6592

Fax: +1 604.822.9106

Keywords: Kriging, Co-kriging, Regression Kriging, Biomass, LiDAR, Radar, Carbon, REDD+

Pre-print of published version.

Reference:

Tsui, O.W., Coops, N.C., Wulder, M.A., and Marshall, P.L. (2013). Integrating airborne LiDAR and space-borne radar via multivariate kriging to estimate above-ground biomass. Remote Sensing of Environment. Vol. 139, pp. 340-352.

DOI.

<http://dx.doi.org/10.1016/j.rse.2013.08.012>

Disclaimer:

The PDF document is a copy of the final version of this manuscript that was subsequently accepted by the journal for publication. The paper has been through peer review, but it has not been subject to any additional copy-editing or journal specific formatting (so will look different from the final version of record, which may be accessed following the DOI above depending on your access situation).

Abstract

Understanding and investigating synergies between LiDAR (light detection and ranging) and SAR (synthetic aperture radar) provides new and innovative opportunities to characterize above-ground biomass. We demonstrate a spatial modeling framework that integrates above-ground biomass transects, derived from plot-based field data and small-footprint discrete return LiDAR, with complete wall-to-wall spaceborne L-band and C-band SAR to predict biomass over a larger area. Transect intervals of 2000 m, 1000 m, and 500 m were tested. Co-kriging, regression kriging, and regression co-kriging were used to extend the LiDAR-derived biomass transects. LiDAR-derived above-ground biomass and L-band backscatter (HV polarization) was moderately correlated, with a maximum semivariance distance between the LiDAR-derived biomass and SAR data of 374 m. Regression kriging at a sample interval of 500 m showed the smallest root mean squared error (RMSE) and mean absolute error (MAE) at 203.9 Mg ha⁻¹ and 131.6 Mg ha⁻¹, respectively. The mean error (ME) showed an average bias of -14.0 Mg ha⁻¹. Predictions using regression co-kriging at a sample interval of 2000 m resulted in the highest RMSE and MAE values at 238.2 Mg ha⁻¹ and 164.6 Mg ha⁻¹, respectively. ME also was highest, averaging -37.4 Mg ha⁻¹. Regardless of the spatial modeling technique employed, lower errors in predicted above-ground biomass were associated with smaller transect intervals. Moderate correlations between the LiDAR-derived above-ground biomass and the radar data impacted the predictive accuracy of the spatial models; however, overall variation in above-ground biomass in the study area was well represented. This study demonstrated that a sampling framework integrating LiDAR data with space-borne radar data using a spatial modeling approach can provide spatially-explicit above-ground biomass estimates for large areas. Such a sampling framework can be used in combination with ground plot and land cover data to assess carbon stocks under conditions where more common optical remote sensing approaches are difficult to implement.

1. Introduction

The impacts of climate change on many natural systems have been documented and reported by the Intergovernmental Panel on Climate Change (IPCC). It is recognized that changes in atmospheric concentration of greenhouse gasses (GHGs) such as CO₂ and modified land cover are key drivers of these changes (IPCC, 2007). Numerous studies have investigated the potential impacts of climate change on physical and biological processes (e.g., Bellard et al., 2012; Coops and Waring, 2011; Parmesan and Yohe, 2003; Zhu et al., 2012). Impacts on forest vegetation from climate change will increasingly affect

forest ecosystems processes (Metsaranta, Dymond, Kurz, & Spittlehouse, 2011). Gradual increases in temperature, changes in rainfall patterns or modification of atmospheric conditions such as cloud cover, will also likely impact vegetation growth, regeneration and natural rates of mortality (Chapin et al., 2010). Sustainable land management practices and policies are critical for mitigating climate change through the reduction of atmospheric CO₂.

Forests are important at multiple scales, from providing habitat for animals and non-timber products at local scales (Ahrends et al., 2010) to influencing climate systems and the carbon cycle globally (Lewis et al., 2009). Atmospheric carbon is removed from the atmosphere and sequestered in a number of pools, including soils, with sequestration in the form of forest vegetation forming a large and dynamic pool. Although terrestrial soils contain more carbon than the atmosphere and vegetation biomass together (Eswaran, Van Den Berg, & Reich, 1993), soil carbon is less dynamic and more difficult to monitor. Forests on the other hand can be monitored, and their dynamics understood. Consequently, forests are an important component in mitigating the effects of climate change. However, when forests are cleared and converted to other land types or degraded, much of their stored carbon is released into the atmosphere as CO₂. As of 2007, deforestation, including decay and peat fires and drained peat soils, is estimated to account for approximately 18 % of the global carbon emissions and is the second largest source of anthropogenic GHG emissions worldwide and the largest source for most tropical countries (IPCC, 2007). More damaging is that deforestation and degradation of tropical forests also removes globally important carbon sinks that currently sequester large quantities of CO₂ from the atmosphere, and which are critical to future climate stabilization (Stephens et al., 2007).

The principle of placing monetary value on forests based on the amount of carbon they store and their condition has produced national and regional strategies such as Reducing Emissions from Deforestation and Forest Degradation (REDD+) (Kimberly & Curran, 2009; Tacconi, Mahanty, & Suich, 2010; van de Sand, 2012). Through the retention of carbon and the avoidance of emissions from deforestation, REDD+ potentially may provide financial incentives to conserve forests and slow the drivers of land use change (Corbera, Soberanis, & Brown, 2009; Miles & Kapos, 2008). Spatially explicit maps of above-ground biomass are thus essential for quantifying the amount of carbon sequestered in forests, and changes in forest carbon stocks and forested areas.

The overall goal of this study is to test and demonstrate data integration methods for producing spatially explicit biomass products suitable for application over a range of environments, including remote or less data rich regions. We propose that samples of airborne LiDAR data, calibrated with field data, can be used in conjunction with space-borne radar data to produce viable wall-to-wall maps of above-ground

biomass. In support of this goal, we first provide background and context for this information need and describe the potential of various remotely sensed data sources. We then suggest appropriate data sets and methods, followed by an example. Finally we discuss the implementation opportunities and considerations associated with this approach.

1.1 Remote sensing and forest biomass

Obtaining comprehensive, timely and reliable forest inventory data is usually costly and labour intensive, especially for large areas (Kangas, Grove, & Scott, 2006). The advent of remote sensing technology has provided a practical and economical means to measure and monitor vegetation cover and structure, especially over large areas (Xie, Sha, & Yu, 2008). Optical data are commonly used for land cover mapping, capture of change, empirical estimates of structural attributes, and to provide strata for improving attribute estimation. Estimating biomass with optical data is also well established; see Lutz et al., 2008 for a review. However, the presence of clouds, shadows, and haze can impact the quality and completeness of data from optical sensors, especially in tropical areas (Roy, Ju, Mbow, Frost, & Loveland, 2010).

Synthetic Aperture Radar (SAR) (Santoro et al., 2011; Thiel, Thiel, & Schmulilius, 2009) and Interferometric SAR (InSAR) (Simard et al., 2006) have also been used for biomass assessment, particularly for areas with persistent cloud cover, because of their cloud penetration capacity, large spatial coverage, and sensitivity to forest biomass. As with most space-borne remote sensing technologies, obtaining accurate biomass estimates using these technologies is challenging as most direct observational methods suffer from saturation constraints (i.e., there is a clear upper limit to the level of forest biomass that can be estimated) or are less accurate at high biomass levels (Ardö, 1992; Duncanson, Niemann, & Wulder, 2010; Imhoff, 1995). In addition, challenges with using SAR can include poor accuracy and temporally unstable relationships due to variability in weather conditions such as moisture conditions, frost, and wind (Kasischke, Tanase, Bourgeau-Chavez, & Borr, 2011; Ranson & Sun, 1997). The possible use of InSAR for assessing biomass is also of interest due to the demonstrated provision of height measurements with no apparent saturation limit. For example, Solberg et al. (2010) showed the effective use of single-pass X-band InSAR data for measuring forest biomass in the boreal region of southern Norway and reported a linear relationship between biomass and InSAR heights with no apparent saturation effect. However, obtaining accurate height measurements requires reducing or compensating for temporal decorrelation, which necessitates the use of multiple baselines to improve interferometric processing or the use of single-pass interferometry. These requirements lower operational uptake given

that there are no space-borne L- or P-band SAR satellites currently operational, and no SAR sensors with single-pass configuration, except TanDEM-X launched in 2010.

Airborne laser scanning or light detection and ranging (LiDAR) has received a great deal of scientific and operational attention for forest characterization in recent years (Hyypä et al., 2012). Although highly accurate, LiDAR data for large area monitoring is challenging because of operational considerations that limit widespread use, such as high data costs, aircraft scheduling and logistics, and large data volumes (Wulder et al., 2008). Wulder et al. (2008) discussed several factors that can affect cost. For instance, improvements in pulse rates enable higher flying heights which mean fewer lines are required to cover an area with the desired hit density; however, fuel costs, especially for remote locations, can be a key cost driver. Given the inverse relationship between spatial coverage and spatial resolution (Franklin, Lavigne, Wulder, & Stenhouse, 2002), cost may be the primary obstacle to using LiDAR for large-area forest characterizations (Wulder & Seemann, 2003). Even with anticipated reductions in LiDAR data costs in the near future (Li, Andersen, & McGaughey, 2008), it is still unlikely that LiDAR data would be available to provide wall-to-wall forest characterization measurements for large or remote locations (Wulder et al., 2012). A more likely scenario would be the integration of LiDAR with other data sources, where samples of LiDAR data provide highly accurate estimates of forest stand characteristics used to calibrate or validate broader observations obtained from other data sources (Wulder et al., 2012). However, there are examples where the expense in LiDAR collection is justified by the information needs present, such as the requirement for elevation data (e.g. Woods et al., 2011).

The ability to obtain frequent observations, demonstrated relationships with biomass, and the all-weather data collection capacity encourage further research with radar data. SAR and InSAR provide complimentary sets of information to LiDAR data for estimating forest biomass (Hyde et al., 2006; Wulder et al., 2012). In this integrative framework, forest biomass estimates derived from LiDAR data are used to calibrate (and subsequently validate) the wide area observations made by radar. For example Mitchard et al. (2012) combined direct observations of L-band radar, space-borne LiDAR, and ground data to map above-ground biomass for Lopé National Park (LNP) in Gabon, an area of 5,000 km². PALSAR backscatter (HH and HV polarization), elevation data, and a radar-derived forest degradation index (RFDI) were used to produce an unsupervised vegetation classification consisting of 40 classes for the entire national park. Each vegetation class was then assigned an average biomass value estimated by ground and LiDAR data. The LiDAR data consisted of sample profiles, covering only 17.85 km² but which intersected all 40 vegetation classes. They estimated the carbon stock of the LNP to be 173 Mg C ha⁻¹, which was consistent with the field data average of 181 Mg C ha⁻¹. The outcome was a 100 m spatial resolution biomass map for the LNP with an estimated uncertainty of ±25.0%.

The complimentary information content of LiDAR and radar promotes the investigation of additional analysis options, especially those incorporating spatial inter-relationships determined by geostatistics. Hudak et al., (2002), in an important paper, considered the integration of LiDAR with Landsat Enhanced Thematic Mapper + (ETM+) imagery to address some of the considerations still relevant today, such as exploiting the high accuracy of LiDAR acquired in a sampling configuration with wall – to – wall nature of most satellite imagery. Hudak et al. (2002) compared five spatial and a-spatial methods to integrate the two datasets and provided significant insights into the most appropriate methods that should be applied with this type of data. Results from Hudak et al., (2002) suggested that the spatial models (kriging and co-kriging), produced less biased results than regression and that co-kriging methods 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.

We believe Hudak et al., (2002) provides important insights for our work, and provides a proven, tried and tested, framework with which to assess how above-ground biomass predictions can be estimated using a LiDAR sampling framework, multivariate kriging, and wall-to-wall radar data to extend the biomass predictions to un-sampled areas. LiDAR is sensitive to tree crown characteristics and height of trees while radar is sensitive to the size and arrangement of structural elements of groups of trees (Sexton, Bax, Siqueira, Swenson, & Hensley, 2009; Tsui, Coops, Wulder, Marshall, & McCardle, 2012). Given these relationships between the LiDAR-derived forest biomass and radar measurements, accuracy in predicted biomass over large areas is expected to improve compared to predictions using SAR data alone.

1.2 Geostatistics

Geostatistics provide techniques to estimate variables that vary in space (Curran & Atkinson, 1998). Semivariograms are typically used to estimate the degree of dissimilarity (or variance) between multiple pairs of measurements and provides information on the scale and pattern of the spatial variance (Curran, 1988, Journel & Huijbregts, 1978). In order to describe the semivariogram and apply it in further analysis, it is necessary to fit a mathematical model from which insights can be gained into the overall variance in the scene (Webster, 1985). Generally, the variance that is spatially independent can be estimated by the nugget, the sill can provide an indication of the maximum semivariance observed where there is no spatial autocorrelation, and the range represents the lag value (95% of the maximum variance or above) where semivariance reaches a maximum

Ordinary kriging, the most common form of kriging (Krige, 1966), is a spatial modeling technique that provides optimal and unbiased estimates of unknown values from sample data (Curran & Atkinson, 1998).

The technique is only appropriate when there is spatial dependence in the data; it provides estimates by assigning weights to each sample data point that is in close proximity to the area of interest. Key to this process is that the weights are determined from the spatial dependence represented by the semivariogram (Curran & Atkinson, 1998); sample data closer to the estimate are given more weight because they are more likely to be similar to the unknown value.

Co-kriging extends ordinary kriging to account for more variables and is typically more appropriate when the primary variable to be estimated (in this study, forest biomass) is under-sampled with respect to the secondary variable (in this study, the radar observations) (Curran & Atkinson, 1998). Similar to kriging, estimates are calculated using the autocorrelation of the primary variable; however, co-kriging also exploits the crosscorrelation of the primary and secondary variable. If the two variables are cross-correlated (i.e., the spatial variability of the primary variable is also correlated with the spatial variability of the secondary variable), this information can be used to make predictions of the primary variable (Bivand, Pebesma, & Gómez-Rubio, 2008). Isotopic co-kriging requires that the data for the primary and auxiliary variables be measured at all sampling locations. Heterotopic co-kriging requires that only some of the sample points contain measurements of both primary and secondary variables (Wackernagel 2003). In cases involving remote sensing data, collocated co-kriging is a particular heterotopic situation and is applied when auxiliary variables are measured at all locations but the primary variable is available at only a few locations.

Regression kriging is a hybrid approach that combines either a simple or multiple linear regression model with kriging of the regression residuals (Goovaerts, 1997). The value of a target variable at some location can be modeled as a sum of deterministic (obtained from the regression) and stochastic (obtained from kriging of the residuals) components.

2. Methodology

2.1 Study Area

The study site is an intensively managed forest area dominated by Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) and western red cedar (*Thuja plicata* Donn ex D. Don) located on Vancouver Island, British Columbia, Canada. The area covers a 5 km by 5 km area around Oyster River (UTM Zone 10, NAD83: Upper left 329450E, 5531300N; Lower right 337550E, 5523500N), with a mean elevation of 240 m, and a range of 120 m to 460 m above sea level (Fig. 1). The site consists of predominately second-growth coniferous forest comprised of 70% Douglas-fir, 17% western red cedar, 3% western hemlock (*Tsuga heterophylla* (Raf.) Sarg.), and 10% red alder (*Alnus rubra* Bong.). It is highly productive

compared to most of Canada, with rotation cycles as short as 60 years (Morgenstern et al., 2004). The stand density ranges from 350 to 1200 stems ha⁻¹, with tree heights ranging between 10.0 and 35.0 m and diameter at breast height (dbh) ranging between 12.0 cm and 31.2 cm (Tsui et al., 2012). The present forest is the result of harvest of the original forest from 1920 to 1950. Much of the area did not regenerate naturally (Goodwin, 1937) and, as a result, some of the area was planted starting in the late 1940s, with second growth harvesting and subsequent planting beginning in 1989. The harvesting history has resulted in a patchwork of second growth stands at different successional stages. The site is located within the dry maritime Coastal Western Hemlock biogeoclimatic subzone (CWHxm), of the biogeoclimatic ecosystem classification (BEC) system of British Columbia. This subzone is characterized by cool summers and mild winters, with mean annual precipitation of 1,500 mm and a mean annual temperature of 9.1 °C (Meidinger & Pojar, 1991). A large portion of the forest was commercially harvested in the 2011 winter and replanted during the 2011 spring subsequent to the imagery used in this study..

2.2 Data: Biomass map

The spatially explicit predictions of above-ground biomass used as the reference data set was estimated by Tsui et al. (2012) from plot-based field data and small-footprint discrete-return LiDAR. LiDAR data were acquired in August 2008 at a mean flying altitude of 2,303m with a bald earth density of between 0.4 - 1.0 points m⁻² and a non-ground density of 0.7 points m⁻². Standard plot-based LiDAR metrics (i.e., mean first return height, standard deviation, coefficient of variation, percentiles of first return heights, percentages of first returns above 2m, and percentage of first returns above the first return mean height) were computed. Biomass, species and age class were determined using 18 fixed-area field plots measuring 30 m x 30 m, with all trees greater than 10 cm dbh within the plot measured for dbh, height, height to the base of the live crown, and species. Above-ground biomass for each tree was calculated using species-specific biomass equations and then summed to obtain plot-level biomass values. The final LiDAR-based biomass model had a root mean squared error (RMSE) of 56.43 Mg ha⁻¹, a relative RMSE of 18%, and an adjusted R² of 0.82.

Mapping of the empirical model produced a spatially explicit above-ground biomass image at a spatial resolution of 20 m, which matched the spatial resolution of the coarsest radar co-variable used in this study (Fig. 1). Final biomass values ranged from 0 to 1100 Mg ha⁻¹ with a mean biomass value of 304 Mg ha⁻¹.

2.3 Data: Radar

Co-incident to the LiDAR-derived biomass data, five radar images were acquired over the study site. Three Fine Beam Dual polarization (FBD) images acquired by the Phased Array type L-band Synthetic Aperture Radar (PALSAR) instrument on the Advanced Land Observing Satellite (ALOS), and two RADARSAT-2 Quad-pol Fine Beam images. All data sets were stored in single-look complex (SLC) format (Table 1).

PALSAR and RADARSAT-2 SLC data were multi-looked using factors of 2 and 8, and factors of 1 and 2 respectively, for range and azimuth directions and then calibrated to obtain SAR backscatter images. Radiometric, geometric, and terrain correction of the ALOS PALSAR and RADARSAT-2 data were performed using the Alaska Satellite Facility (ASF) MapReady software package. Following radiometric and geometric correction, conversion from sigma nought to gamma nought, which normalizes the radar cross-section, was completed for an improved representation of backscatter values for distributed targets such as forests. In addition to backscatter images, InSAR coherence magnitudes were also calculated using PALSAR and RADARSAT-2 complex image pairs (i.e., two radar images with similar orbits with amplitude and phase information). By measuring the difference in the phase of the microwave pulses after interacting with an object (e.g., a tree branch), the coherence of the phase can be calculated. As the phase difference or phase shift increases, mainly caused by random fluctuations (e.g., wind induced movements), the coherence decreases within a range from 1 to 0. Since vegetation causes signal coherence to decrease because the exact point of scattering and the travel path of the signals vary between the radar scenes, coherence is highest for open areas and decreases as vegetation increases (Rosen et al., 2000). Lastly, the RADARSAT-2 data were down-sampled to a spatial resolution of 20 m and reduced to a spatial subset equivalent to the extent of the estimated above-ground biomass data set. For a more complete description of the radar processing steps completed, see Tsui et al. (2012).

2.4 Above-ground biomass sampling

Sample forest biomass values were extracted from the processed biomass data set for the study area. Sampling of these data consisted of both a north-south and east-west continuous transect with each transect line consisting of a data point every 20 m. Transect lines were separated by three distances, 500 m, 1000m, or 2000m, with a total number of points of $n = 5,540$; $n = 3,020$; and $n = 1,512$ respectively (Fig. 2). Histograms of the three forest biomass data sets exhibited a strong positive skew. Therefore each biomass data set was normalized using a log transformation followed by a normal score transformation to obtain a sample distribution closely resembling a standard normal distribution (Olea, 1977) before

geostatistical analyses were performed. After modeling, all predicted above-ground biomass values were transformed back to the original units (i.e., Mg h^{-1}) before evaluating the various models.

2.5 Biomass modeling

Above-ground biomass, derived from LiDAR height measurements, was used as the primary variable, and backscatter intensities and coherence magnitudes from ALOS PALSAR and RADARSAT-2 data were used as the wall-to-wall co-variables. The PALSAR and RADARSAT-2 radar backscatter coefficients and coherence magnitudes were also transformed to normalize the data prior to use in the estimation processes. Sampling of the primary variable was performed to simulate airborne profiling LiDAR observations. Co-kriging, regression kriging, and regression co-kriging were performed using algorithms found in the GSTAT package designed for R (Pebesma, 2004).

For each biomass data set, created from the three sampling strategies, the most suitable mathematical model was determined by fitting the most common semivariogram models to each of the estimated semivariograms using weighted least squares and evaluating the quality of fit through the residual sum of squared errors. To support the process of predicting biomass values at non-sampled locations, wall-to-wall coverage of the secondary variable was included in the kriging process, as illustrated in Fig. 3 for 1000 m transects. For co-kriging, a model semivariogram was also calculated for the secondary variable (PALSAR HV polarization), as well as a cross-semivariogram describing the cross-correlation between the primary and secondary variables. Even though it is possible to model any number of co-variables, it has been shown that co-kriging results are virtually identical to kriging outputs when spatial cross-correlation between the primary and secondary variable is minimal or absent (Wackernagel, 2003). Care was taken when developing the three semi-variogram models required for each co-kriging operation to ensure that each model conformed to the linear model of co-regionalization (i.e., all models (direct and cross) have the same shape and range, but have different sills and nuggets to ensure that the covariance matrices were always positive) (Goovaerts, 1997).

A three stage process was followed to obtain predictions from regression kriging and regression co-kriging. For regression kriging, the deterministic part of the predictions was performed first by regressing above-ground biomass on the various radar data sets using ordinary least squares to get an estimate of forest biomass from the radar variables. Secondly, residuals from the ordinary least squares regression were interpolated across the study area using ordinary kriging. Lastly, the deterministic and stochastic components were combined together to obtain the final predicted value.

To determine which radar variables were significant for predicting biomass, a stepwise multiple regression analysis was employed using the `regsubsets` function from the `Leaps` package (Leaps 2009) with a significance level of $\alpha = 0.05$. To interpolate the regression residuals for the study area, a model semivariogram was determined following the same procedures outlined for co-kriging. The difference between regression kriging and regression co-kriging primarily lies in the interpolation of the regression residuals. Instead of interpolating the residuals using ordinary kriging, regression co-kriging uses a secondary variable for interpreting the residuals. We selected PALSAR HV backscatter as the secondary variable for regression co-kriging since this data set had the highest correlation with forest biomass.

2.6 Model Evaluation

Discrepancies between actual and predicted forest biomass were evaluated based on the validation sample outlined in Fig. 2(d). The validation data consisted of 580 forest biomass points extracted from the reference above-ground biomass data set. These points were outside the training data set used in the kriging process. Following Alsamamra et al. (2009), Meng et al. (2009), and Murphy and Katz (1985) three different and common indices were used to assess the predictions: root mean squared error (RMSE), mean absolute error (MAE), and mean error (ME). The RMSE and MAE measure the average precision of the prediction and provide an indication of how close the predictions are to the observed values. The ME provides an indication of bias in the predictions and it should be close to zero for unbiased methods. Lastly, t-tests were used to compare whether the residuals differed statistically from one another for each of the distances between the transects and modeling techniques.

3. Results

Semivariograms were generated to assess the spatial properties of the sampled above-ground biomass, biomass residuals, and the radar co-variable data sets. For above-ground biomass, all sample intervals showed similar spatial dependence with similar shape and nugget, partial sill, and range parameters. Similar results were found for the biomass residuals and the cross-semivariogram data set. For the above-ground biomass and biomass residual data, nugget variance increased as sample interval increased from 500 m to 2000m. The distance where maximum semivariance was observed ranged from 328 m to 374 m and 313 m to 324 m for the above-ground biomass and residuals, respectively. The cross-semivariograms between the above-ground biomass and PALSAR HV polarization showed higher nugget variance and a range distance from 324 m to 374 m. Less spatial autocorrelation was observed between biomass and PALSAR HV polarization, while spatial autocorrelation was evident for the biomass and residual datasets. The Whittle-Matern model had the best fit for all cases and was selected as the theoretical mathematical

model for the use in spatial predictions. Example model semivariograms for each variable for the 1000 m sampling interval are provided in Fig. 4 and Table 2.

Point pairs for each variable were graphed to assess if any global trends existed. No obvious trend was found among the individual variables. Anisotropy was also checked. Similar spatial dependence and semi-variance for all sample intervals at directions of 0, 45, 90, 135, 180, 225, 270, and 315 degrees were found.

3.1 Biomass Estimates

Above-ground biomass was underestimated in all cases. Regression kriging at a sample interval of 500 m showed the smallest RMSE and MAE at 203.9 Mg ha⁻¹ and 131.6 Mg ha⁻¹, respectively. The ME showed an average bias of -14.0 Mg ha⁻¹. Moderate correlation ($r = 0.68$) was observed between the predicted and reference above-ground biomass. Predictions using regression co-kriging at a sample interval of 2000 m resulted in the highest RMSE and MAE values at 238.2 Mg ha⁻¹ and 164.6 Mg ha⁻¹, respectively. ME also was highest, averaging -37.4 Mg ha⁻¹. Correlation between the observed and reference above-ground biomass was also lower ($r = 0.52$) (Table 3). Regression kriging generally showed the lowest RMSE, MAE, and ME for all distances between the transects. As expected, biases in above-ground biomass predictions decreased as distance between the transects decreased.

Histograms of above-ground biomass (Fig. 5) were used to visualize the accuracy of the predicted values. Deviations of the predicted above-ground biomass histogram from the reference histogram provided further insight into where the majority of the prediction bias occurred. As indicated by ME values, all kriging methods and transect widths underestimated the biomass, with the majority of the deviations occurring at higher biomass levels. Mean predicted above-ground biomass for all methods ranged from 255 to 278 Mg ha⁻¹, a negative deviation of 49 Mg ha⁻¹ to 26 Mg ha⁻¹ from the mean above-ground biomass of the reference biomass data set. The histograms also showed that the predicted biomass was overestimated for low biomass areas.

Improvements can be seen with shorter transect intervals (e.g. larger samples) (Fig. 6). None of the geostatistical methods predicted above-ground biomass above approximately 1250 Mg ha⁻¹, with the exception of co-kriging which had an outlier above 1400 Mg ha⁻¹.

Although the range in prediction errors was large for each case, the majority of the deviations between predicted and reference above-ground biomass were between -200 Mg ha⁻¹ and 200 Mg ha⁻¹ (Fig. 7). The variance in residuals decreased as sampling frequency increased. There was not much change in variance

among the geostatistical methods within the same sampling frequency. Differences in mean prediction errors were not significant among the methods. Co-kriging showed no significant change in mean predicted errors between sampling distances. While no differences were observed between the 500 and 1000 m for regression kriging and regression co-kriging ($n = 580$; $\alpha = 0.05$; $p\text{-value} = 0.15$ and 0.50 respectively), significant differences were observed between the 1000 and 2000 m width ($n = 580$; $\alpha = 0.05$; $p\text{-value} = 1.2 \times 10^{-3}$ and 5.7×10^{-4} respectively), and the 500 and 2000 m width ($n = 580$; $\alpha = 0.05$; $p\text{-value} = 3.7 \times 10^{-6}$ and 7.2×10^{-6} respectively).

3.2 Biomass Mapping Models

The maps produced captured the overall variation in above-ground biomass in the study area (Fig. 1 vs. Fig. 8). As expected, increased sampling frequency provided better definition of the variation. Artifacts were observed in the co-kriging maps and were more pronounced at the wider sampling distances. The regression kriging and regression co-kriging maps were quite similar and the artifacts present in the co-kriging maps were virtually indistinguishable. Figure 9 provides an indication of where the error in the prediction is greatest, computed as the difference between the predicted and the reference biomass dataset. As expected the co-kriging difference maps (top line) show, in general, an over prediction of biomass amounts compared to the reference. Some spatial structure is observable in the difference maps however with areas of high biomass under predicted and areas of low biomass over predicted. The regression kriging and co-kriging difference maps (middle and bottom row) show less overall error and less spatial structure, consistent with improved overall prediction.

4. Discussion

We found lower errors in predicted biomass with less distance between adjacent sampling transects. There were some differences among the kriging methods employed and sampling strategies. Systematic sampling at 1000 m intervals together with regression kriging was a good compromise between ease of use, increase in accuracy, and cost of obtaining LiDAR data for this study area. The accuracy of all predictions was affected by the low correlation between the LiDAR-derived above-ground biomass estimates and the radar data, which was reflected in the cross-semivariogram by the high nugget and partial sill value. Regardless of the sampling distance between transects, range values at which no spatial dependence was detected were consistently less than 400 m. As a result, the semi-variogram had little influence on the estimation process beyond this distance. This resulted in a smoothing effect, particularly evident when co-kriging was employed. This smoothing effect was less evident in regression kriging and regression co-kriging due to the deterministic portion of these models.

The tendency for kriging to underestimate large values and overestimate small values is supported by previous studies (Hudak et al., 2002 and Meng et al., 2009). This tendency may help account for the inflated RMSE values, given that large errors are given disproportional weight because of the squaring of the differences. It is possible to correct for backtransformation bias which has been shown to reduce RMSE in some cases (Hudak et al 2006). In our analysis of the individual plot data which resulted in the biomass map used in this study, by Tusi et al (2012), this backtransformation bias was less than 9 Mg h⁻¹ indicating it is not contributing significantly to the errors in these models. Ordinary kriging and co-kriging are the best approaches in cases where spatial interpolation is required, since kriging coefficients rely on the spatial variation between sample points (Hudak et al., 2002). However, regression kriging is better suited for spatial extrapolation, since the main kriging coefficients only depend on the correlation between the dependent variable and independent variables (Meng et al., 2009). The techniques and sampling framework presented in this study are relevant to large area above-ground biomass assessments. Due to the asymptotic relationships between biomass and radar backscatter, and the high variance between the two, integrating direct radar backscatter and LiDAR data using a spatial modeling approach is likely more suited for large area mapping of moderate forest biomass levels where correlation would be expectedly higher. Integrating radar backscatter and LiDAR to estimate areas with high biomass could possibly benefit from aspatial methods as shown by Tsui et al. (2012). However, such an approach is most suitable for smaller areas where both LiDAR and radar data are available at all locations. In the past five years a number of additional modeling frameworks have become more popular such as K-NN imputation methods and random forest techniques. These techniques may offer some additional predictive power to this type of analysis; however neither explicitly attempt to model the observed spatial correlation in the data as geo-statistical based approaches do. As stated by Hudak et al. (2002) most remote sensing analysis is afforded the luxury of sampling the entire population, reducing the need for spatial interpolation strategies such as discussed in this paper. The widespread use of LiDAR in a sampling framework however demonstrates there is a continual examination of these geo-statistical methods is needed to reveal insights and to maximise the value of these datasets.

Given that tree heights are known to be highly correlated to biomass, the potential for integrating LiDAR with wall-to-wall canopy heights derived from Polarimetric InSAR (Pol-InSAR) or InSAR is high. By virtue of the properties of radar, InSAR heights usually correspond to the location of the scattering phase center, which typically underestimates actual canopy heights (Balzter, Luckman, Skinner, Rowland, & Dawson, 2007). Therefore, integrating InSAR canopy height measurements with highly accurate LiDAR observations is one way to obtain higher accuracy. The planned suite of future satellites (NASA's

Deformation, Ecosystem, Structure, and Dynamics of Ice – Radar (DESDynI-R) L-band mission, Japan’s ALOS-2 L-band Mission, and ESA’s BIOMASS P-band mission), in addition to advancements in SAR interferometric processing, such as multiple baseline InSAR (Neumann, Ferro-Famil, & Reigber, 2010), and SAR tomography (Reigber & Moreira, 2000), should allow for future operational integration of SAR and LiDAR data. Additionally, new means of optical image understanding and processing are providing novel opportunities for composting that may mitigate the negative impacts of cloud cover (Hansen & Loveland, 2012).

Regardless of which secondary variable is selected, highly accurate assessments of forest biomass for large areas will likely require integration with LiDAR observations, whether airborne or spaceborne. Large area coverage of LiDAR requires a sampling framework to capture the variation and structural characteristics of the forested area (Wulder et al., 2012). The cost/benefit of such a sampling framework can be illustrated by performing a cost comparison for this study area using the approximate costs in Wulder et al. (2008). For a total area of 25 km², complete airborne LiDAR coverage to obtain a posting size of 30 cm would cost approximately \$1000 CAD per km² (\$25,000 CAD total). Implementing a sampling framework incorporating continuous profiling transects with a swath size of 100 m at the same posting size for every 1000 meters would cost approximately \$6,000 CAD, a quarter of the cost of full LiDAR coverage. However, direct one-to-one comparison of these data is problematic as the data differ in quality and information content. If accurate tree or stand-level height information is required, LiDAR data may not be considered expensive. However, if lower costs are needed, while achieving reasonably accurate depiction of stand height or biomass, a sampling framework and modeling approach may be preferred.

Cost savings generated from a LiDAR sampling framework and a multi-sensor approach would conceivably benefit activities like assessing above-ground carbon stocks and carbon stock change, particularly for tropical areas where forest lands are regularly cloud covered and are inaccessible. Strategies for implementing a LiDAR sampling framework for such a scenario could involve several general steps:

1. Stratify the land primarily into forest land (FL) and non-forest land (NFL) using pre-existing land cover data if available (Frazer et al., 2011). If such land cover data are not available, separation of broad land cover types can be obtained through the use of SAR as demonstrated by Hoekman et al. (2010) for the island of Borneo.

2. Acquire space-borne radar at the spatial resolution required to meet information need and use land cover data to mask-out (i.e., remove) areas identified as NFL (similar to approaches described in Tsui et al., (2012)
3. Establish a series of sample ground plots for all areas stratified as FL and ensure the complete range of biomass is captured to characterize the population. Then, calculate above-ground biomass through the use of appropriate allometric equations (Hawbaker et al. 2009). The number of ground plots will depend on the size of the study area, but also must consider the spatial distribution of individual forest types and the within type variability (Wulder et al., 2012). A complete discussion is not provided here, please refer to Brown (1999) and White et al., (2013) for details.
4. Collect airborne LiDAR data to intersect each of the sample ground plots. Regress calculated above-ground biomass data with the LiDAR measurements to obtain sample above-ground biomass transects. A concise discussion on the use of profiling LiDAR as a sampling tool is provided in Nelson et al. (2003). Because of the nature of how airborne LiDAR data are collected (i.e., always linear) Wulder et al. (2012) suggests possible sampling designs that can be used.
5. Estimate and model semi-variograms; from the LiDAR transects and the SAR data and predict above-ground biomass for un-sampled locations using regression kriging.
6. Complete quantitative validation to compare predicted values to actual values. Reference data may consist of new ground plot data or an independent subset of LiDAR transects not used in the modeling (White et al., 2013).

5. Conclusion

In this research we focused on the predictive accuracy of three kriging techniques to estimate above-ground biomass at a spatial resolution of 20 m. We demonstrated how samples of forest biomass, derived from airborne LiDAR and ground plot data, can be combined with wall-to-wall spaceborne radar observations to achieve spatially continuous estimates. In this integrative framework, spatial modeling methods provide an effective means to overcome challenges in acquiring large area biomass assessments. Although improvements can be made through additional research, LiDAR and space-borne radar data will likely prove useful where comprehensive national forest inventories do not exist or are too expensive to implement, and where frequent cloud cover make other methods of quantifying forest biomass difficult. The wall-to-wall mapping opportunity enabled through integrating with radar data provides additional value to existing data and practices.

It can be seen from the example provided in this study that accurate large-area biomass maps are possible, provided appropriate samples of LiDAR-based biomass area are available and these samples represent the population statistically and geographically. Next steps would be to compare this method to the more common “combine and assign” or “stratify and multiply” approaches (Goetz et al., 2009) and examine the errors associated with each of these approaches.

Acknowledgements

We thank Prof. Andy Black and staff for allowing us access to the FLUXNET-Canada site and Timberwest and Island Timberlands for providing access to their forest inventories and access to their private lands. We also thank Colin Ferster, Jean-Simon Michaud, and Martin van Leeuwen (UBC) for field assistance. Components of this study were funded by NSERC Engage and an NSERC Discovery grant to Coops. All RADARSAT-2 data were provided through the Canadian Space Agency’s RADARSAT-2 Science and Operational Applications Research Education Initiative (SOAR-E). We are grateful to the two reviewers for comments and advice on the manuscript.

References

- Ahrends, A., Burgess, N. D., Milledge, S. A. H., Bulling, M. T., Fisher, B., Smart, J. C. R., Lewis, S. L. (2010). Predictable waves of sequential forest degradation and biodiversity loss spreading from an African city. *Proceedings of the National Academy of Sciences*, 107(33), 14556–14561.
- Alsamamra, H., Ruiz-Arias, J. A., Pozo-Vázquez, D., & Tovar-Pescador, J. (2009). A comparative study of ordinary and residual kriging techniques for mapping global solar radiation over southern Spain. *Agricultural and Forest Meteorology*, 149(8), 1343–1357.
- Ardö, J. (1992). Volume quantification of coniferous forest compartments using spectral radiance recorded by Landsat Thematic Mapper. *International Journal of Remote Sensing*, 13(9), 1779–1786
- Balzter, H., Luckman, A., Skinner, L., Rowland, C., & Dawson, T. (2007). Observations of forest stand top height and mean height from interferometric SAR and LiDAR over a conifer plantation at Thetford Forest, UK. *International Journal of Remote Sensing*, 28(6), 1173–1197.
- Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. *Ecology Letters*, 15(4), 365–377.
- Bivand, R. S., Pebesma, E. J., & Gómez-Rubio, V. (2008). *Applied spatial data analysis with R*. New York: Springer.
- Brown, S. (1999). *Guidelines for inventorying and monitoring carbon offsets in forest-based projects*. Arlington, VA: Winrock International. Retrieved from http://www.winrock.org/ecosystems/files/Guidelines_for_Inventorying_and_Monitoring.pdf (Accessed 20 December, 2012)

525 Chapin, F. S., McGuire, A. D., Ruess, R. W., Hollingsworth, T. N., Mack, M. C., Johnstone, J. F., Taylor,
526 D. L. (2010). Resilience of Alaska's boreal forest to climatic change. *Canadian Journal of Forest*
527 *Research*, 40(7), 1360–1370.

528 Coops, N. C., & Waring, R. H. (2011). Estimating the vulnerability of fifteen tree species under changing
529 climate in Northwest North America. *Ecological Modelling*, 222(13), 2119–2129.

530 Corbera, E., Soberanis, C. G., & Brown, K. (2009). Institutional dimensions of Payments for Ecosystem
531 Services: An analysis of Mexico's carbon forestry programme. *Ecological Economics*, 68(3), 743–
532 761.

533 Curran, P. J. (1988). The semivariogram in remote sensing: An introduction. *Remote Sensing of*
534 *Environment*, 24(3), 493–507.

535 Curran, P. J., & Atkinson, P. M. (1998). Geostatistics and remote sensing. *Progress in Physical*
536 *Geography*, 22(1), 61–78.

537 Dobson, M. C., Ulaby, F. T., LeToan, T., Beaudoin, A., Kasischke, E. S., & Christensen, N. (1992).
538 Dependence of radar backscatter on coniferous forest biomass. *IEEE Transactions on Geoscience and*
539 *Remote Sensing*, 30(2), 412–415.

540 Duncanson, L., Niemann, K., & Wulder, M. (2010). Integration of GLAS and Landsat TM data for
541 aboveground biomass estimation. *Canadian Journal of Remote Sensing*, 36(2), 129–141.

542 Eswaran, H., Van Den Berg, E., & Reich, P. (1993). Organic carbon in soils of the world. *Soil Science*
543 *Society of America Journal*, 57(1), 192–194.

544 Franklin, S. E., Lavigne, M. B., Wulder, M. A., & Stenhouse, G. B. (2002). Change detection and
545 landscape structure mapping using remote sensing. *The Forestry Chronicle*, 78(5), 618–625.

546 Frazer, G.W., Magnussen, S., Wulder, M.A., Niemann, K.O. (2011). Simulated impact of sample plot size
547 and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand
548 biomass. *Remote Sensing of Environment*, 115: 636–649.

549 Goetz, S. J., Baccini, A., Laporte, N. T., Johns, T., Walker, W., Kelldorfer, J., Sun, M. (2009). Mapping
550 and monitoring carbon stocks with satellite observations: a comparison of methods. *Carbon Balance*
551 *and Management*, 4, 2.

552 Goodwin, G. (1937). Regeneration study on the logged-off lands of the Comox Logging and Railway
553 Company Oyster River forest survey. Vol. R 72, Surv. File 0124780. B. C. Forest Service, Canada.

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

555 Hansen, M. C., & Loveland, T. R. (2012). A review of large area monitoring of land cover change using
556 Landsat data. *Remote Sensing of Environment*, 122, 66–74.

557 Hawbaker, T.J., Keuler, N.S., Lesak, A.A., Gobakken, T., Contrucci, K., Radeloff, V.C. (2009). Improved
558 estimates of forest vegetation structure and biomass with Lidar-optimized sampling design. *Journal of*
559 *Geophysical Research*, 114: doi:10.1029/2008JG000870.

560 Hoekman, D. H., Vissers, M. A. M., & Welaard, N. (2010). PALSAR wide-area mapping of Borneo:
 561 Methodology and map validation. *IEEE Journal of Selected Topics in Applied Earth Observations and*
 562 *Remote Sensing*, 3(4), 605–617.

563 Hudak, A. T., Lefsky, M. A., Cohen, W. B., & Berterretche, M. (2002). Integration of lidar and Landsat
 564 ETM+ data for estimating and mapping forest canopy height. *Remote Sensing of Environment*, 82(2-
 565 3), 397–416.

566 Hyde, P., Dubayah, R., Walker, W., Blair, J. B., Hofton, M., & Hunsaker, C. (2006). Mapping forest
 567 structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird)
 568 synergy. *Remote Sensing of Environment*, 102(1-2), 63–73.

569 Hyypä, J., Yu, X., Hyypä, H., Vastaranta, M., Holopainen, M., Kukko, A., Alho, P. (2012). Advances in
 570 forest inventory using airborne laser scanning. *Remote Sensing*, 4(5), 1190–1207.

571 Imhoff, M. L. (1995). Radar backscatter and biomass saturation: ramifications for global biomass
 572 inventory. *IEEE Transactions on Geoscience and Remote Sensing*, 33(2), 511–518.

573 IPCC. (2007). *Climate change 2007: synthesis report (AR7)* (p. 104). Geneva: Intergovernmental Panel
 574 on Climate Change.

575 Journel, A. G., & Huijbregts, C. J. (1978). *Mining geostatistics*. London: Academic Press.

576 Kangas, A., Grove, J. H., & Scott, C. T. (2006). Forest inventory: methodology and applications. In A.
 577 Kangas & M. Maltamo (Eds.). Netherlands: Springer.

578 Kasischke, E. S., Tanase, M. A., Bourgeau-Chavez, L. L., & Borr, M. (2011). Soil moisture limitations on
 579 monitoring boreal forest regrowth using spaceborne L-band SAR data. *Remote Sensing of*
 580 *Environment*, 115, 227–232.

581 Kimberly, C. M., & Curran, L. M. (2009). REDD pilot project scenarios: are costs and benefits altered by
 582 spatial scale? *Environmental Research Letters*, 4(3), 031003.

583 Krige, D. G. (1966). Two-dimensional weighted moving average trend surfaces for ore-evaluation.
 584 *Journal of the South African Institute of Mining and Metallurgy*, 66, 13–38.

585 Le Toan, T., Beaudoin, A., Riou, J., & Guyon, D. (1992). Relating forest biomass to SAR data. *IEEE*
 586 *Transactions on Geoscience and Remote Sensing*, 30(2), 403–411.

587 Lewis, S. L., Lopez-Gonzalez, G., Sonké, B., Affum-Baffoe, K., Baker, T. R., Ojo, L. O., Wöll, H. (2009).
 588 Increasing carbon storage in intact African tropical forests. *Nature*, 457(7232), 1003–1006.

589 Li, Y., Andersen, H.-E., & McGaughey, R. (2008). A comparison of statistical methods for estimating
 590 forest biomass from light detection and ranging data. *Western Journal of Applied Forestry*, 23, 223–
 591 231.

592 Lutz, D. A., Washington-Allen, R. A., & Shugart, H. H. (2008). Remote sensing of boreal forest
 593 biophysical and inventory parameters: a review. *Canadian Journal of Remote Sensing*, 34(S2), 286–
 594 313.

595 Meidinger, D. V., & Pojar, J. (1991). *Ecosystems of British Columbia*. Victoria: Research Branch,
596 Ministry of Forests.

597 Meng, Q., Cieszewski, C., & Madden, M. (2009). Large area forest inventory using Landsat ETM+: A
598 geostatistical approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1), 27–36.

599 Metsaranta, J. M., Dymond, C. C., Kurz, W. A., & Spittlehouse, D. L. (2011). Uncertainty of 21st century
600 growing stocks and GHG balance of forests in British Columbia, Canada resulting from potential
601 climate change impacts on ecosystem processes. *Forest Ecology and Management*, 262(5), 827–837.

602 Miles, L., & Kapos, V. (2008). Reducing Greenhouse Gas Emissions from Deforestation and Forest
603 Degradation: Global Land-Use Implications. *Science*, 320(5882), 1454–1455.

604 Mitchard, E. T. A., Saatchi, S. S., White, L. J. T., Abernethy, K. A., Jeffery, K. J., Lewis, S. L. Meir, P.
605 (2012). Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park,
606 Gabon: overcoming problems of high biomass and persistent cloud. *Biogeosciences*, 9(1), 179–191.

607 Morgenstern, K., Black, A. T., Humphreys, E. R., Griffis, T. J., Drewitt, G. B., Cai, T., Livingston, N. J.
608 (2004). Sensitivity and uncertainty of the carbon balance of a Pacific Northwest Douglas-fir forest
609 during an El Niño/La Niña cycle. *Agricultural and Forest Meteorology*, 123(3–4), 201–219.

610 Murphy, A. H., & Katz, R. W. (1985). *Probability, statistics, and decision making in the atmospheric*
611 *sciences*. Boulder, Colorado: Westview Press.

612 Nelson, R., Valenti, M. A., Short, A., & Keller, C. (2003). A multiple resource inventory of Delaware
613 using airborne laser data. *BioScience*, 53(10), 981–992.

614 Neumann, M., Ferro-Famil, L., & Reigber, A. (2010). Estimation of forest structure, ground, and canopy
615 layer characteristics from multibaseline polarimetric interferometric SAR data. *IEEE Transactions on*
616 *Geoscience and Remote Sensing*, 48(3), 1086–1104.

617 Olea, R. A. (1977). *Measuring spatial dependence with semivariograms*. Lawrence, Kansas: Kansas
618 Geological Survey.

619 Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across
620 natural systems. *Nature*, 421(6918), 37–42.

621 Pebesma, E. J. (2004). Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*,
622 30(7), 683–691.

623 Ranson, K. J., & Sun, G. (1997). An evaluation of AIRSAR and SIR-C/X-SAR images for mapping
624 northern forest attributes in Maine, USA. *Remote Sensing of Environment*, 59(2), 203–222.

625 Reigber, A., & Moreira, A. (2000). First demonstration of airborne SAR tomography using multibaseline
626 L-band data. *IEEE Transactions on Geoscience and Remote Sensing*, 38(5), 2142–2152.

627 Rignot, E., Way, J., Williams, C., & Viereck, L. (1994). Radar estimates of aboveground biomass in
628 boreal forests of interior Alaska. *IEEE Transactions on Geoscience and Remote Sensing*, 32(5), 1117–
629 1124.

- 630 Rosen, P. A., Hensley, S., Joughin, I. R., Li, F. K., Madsen, S. N., Rodriguez, E., & Goldstein, R. M.
631 (2000). Synthetic aperture radar interferometry. *Proceedings of the IEEE*, 88(3), 333–382.
- 632 Roy, D. P., Ju, J., Mbow, C., Frost, P., & Loveland, T. (2010). Accessing free Landsat data via the Internet:
633 Africa's challenge. *Remote Sensing Letters*, 1(2), 111–117.
- 634 Santoro, M., Beer, C., Cartus, O., Schmullius, C., Shvidenko, A., McCallum, I., Wiesmann, A. (2011).
635 Retrieval of growing stock volume in boreal forest using hyper-temporal series of Envisat ASAR
636 ScanSAR backscatter measurements. *Remote Sensing of Environment*, 115(2), 490–507.
- 637 Sexton, J. O., Bax, T., Siqueira, P., Swenson, J. J., & Hensley, S. (2009). A comparison of lidar, radar, and
638 field measurements of canopy height in pine and hardwood forests of southeastern North America.
639 *Forest Ecology and Management*, 257(3), 1136–1147.
- 640 Simard, M., Zhang, K., Rivera-Monroy, V. H., Ross, M. S., Ruiz, P. L., Castañeda-Moya, E., Rodriguez,
641 E. (2006). Mapping height and biomass of mangrove forests in Everglades National Park with SRTM
642 elevation data. *Photogrammetric Engineering Remote Sensing*, 72(3), 299–311.
- 643 Solberg, S., Astrup, R., Gobakken, T., N sset, E., & Weydahl, D. J. (2010). Estimating spruce and pine
644 biomass with interferometric X-band SAR. *Remote Sensing of Environment*, 114, 2353 – 2360.
- 645 Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Denning, A. S. (2007).
646 Weak northern and strong tropical Land carbon uptake from vertical profiles of atmospheric CO₂.
647 *Science*, 316(5832), 1732–1735.
- 648 Tacconi, L., Mahanty, S., & Suich, H. (2010). *Payments for environmental services, forest conservation
649 and climate change: Livelihoods in the REDD?* Cheltenham, UK: Edward Elgar Publishing.
- 650 Thiel, C. J., Thiel, C., & Schmullius, C. C. (2009). Operational large-area forest monitoring in Siberia
651 using ALOS PALSAR summer intensities and winter coherence. *IEEE Transactions on Geoscience
652 and Remote Sensing*, 47(12), 3993–4000.
- 653 Tsui, O. W., Coops, N. C., Wulder, M. A., Marshall, P. L., & McCardle, A. (2012). Using multi-frequency
654 radar and discrete-return LiDAR measurements to estimate above-ground biomass and biomass
655 components in a coastal temperate forest. *ISPRS Journal of Photogrammetry and Remote Sensing*,
656 69(0), 121–133.
- 657 Van de Sand, I. (2012). *Payments for ecosystem services in the context of adaptation to climate change.*
658 *Ecology and Society*, 17(1), 11.
- 659 Wackernagel, H. (2003). *Multivariate geostatistics*. New York: Springer.
- 660 Webster, R. (1985). *Quantitative Spatial Analysis of Soil in the Field*. Springer-Verlag, 3, 1–70.
- 661 White, J.C., M.A. Wulder, A. Varhola, M. Vastaranta, N.C. Coops, B.D. Cook. (2013). A best practices
662 guide for generating forest inventory attributes from LiDAR. Version 1. Internal Report. Pacific
663 Forestry Centre, Canadian Forest Service, Natural Resources Canada. Victoria, BC. Canada. 67p.
- 664 Woods, M., Pitt, D., Penner, M., Lim, K., Nesbitt, D., Etheridge, D., & Treitz, P. (2011). Operational
665 implementation of a LiDAR inventory in Boreal Ontario. *The Forestry Chronicle*, 87(4), 512–528.

- 666 Wulder, M A, Bater, C. W., Coops, N. C., Hilker, T., & White, J. C. (2008). The role of LiDAR in
667 sustainable forest management. *The Forestry Chronicle*, 84(6), 807–826.
- 668 Wulder, M A, & Seemann, D. (2003). Forest inventory height update through the integration of lidar data
669 with segmented Landsat imagery. *Canadian Journal of Remote Sensing*, 29(5), 536–543.
- 670 Wulder, M A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., Gobakken, T. (2012).
671 Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment*,
672 121(0), 196–209.
- 673 Xie, Y., Sha, Z., & Yu, M. (2008). Remote sensing imagery in vegetation mapping: a review. *Journal of*
674 *Plant Ecology*, 1(1), 9–23.
- 675 Zhu, K., Woodall, C. W., & Clark, J. S. (2012). Failure to migrate: lack of tree range expansion in
676 response to climate change. *Global Change Biology*, 18(3), 1042–1052.

677 **Tables and Figures**

678	Fig. 1.	Location of study site and above-ground biomass values estimated by discrete-return
679		LiDAR.
680	Fig. 2.	Sampling strategies tested and data volumes for each sample forest biomass data
681		set: (a) 2000 m, (b) 1000 m, (c) 500m, and (d) validation points. Shaded grey area
682		represents the extent of the reference LiDAR derived above-ground biomass data set.
683	Fig. 3.	Image lattices showing characteristics of the experimental design for multivariate
684		kriging. Above-ground biomass transects simulate airborne profiling LiDAR flight lines
685		at 1000 m intervals.
686	Fig. 4.	Experimental (black points) and model (black line) semivariograms for (a). above-
687		ground biomass; (b) PALSAR HV polarization, (c) cross-semivariogram, and (d) OLS
688		residuals for the 1000m sampling interval.
689	Fig. 5.	Histograms of estimated above-ground biomass values (shaded in black) for all
690		sampling strategies tested: (1) co-kriging (a,b,c); (2) regression kriging (a,b,c); and (3)
691		regression co-kriging (a,b,c). Histogram of reference biomass values provided as
692		reference (shaded in grey, N= 80,025).
693	Fig. 6.	Scatterplots of estimated vs. observed above-ground biomass values for all sampling
694		strategies tested: (1) co-kriging (a,b,c); (2) regression kriging (a,b,c); and (3)
695		regression co-kriging (a,b,c). Pearson's correlation coefficient provided for each
696		sampling strategy. Scatterplots represent accuracy of estimated values based on
697		validation points (N = 580).
698	Fig. 7.	Violin plot showing the interquartile range (mid-spread) of residuals in predicted
699		biomass for all sampling strategies tested. Mean error indicated by red cross. OCK -
700		Ordinary Co-kriging; RK – Regression kriging; and RCK – Regression co-kriging.
701	Fig. 8.	Estimated above-ground biomass maps for all sampling strategies tested using: (1)
702		co-kriging (a,b,c); (2) regression kriging (a,b,c); and (3) regression co-kriging (a,b,c).
703	Fig. 9.	Difference maps of reference minus predicted above-ground biomass maps for all
704		sampling strategies tested using: (1) co-kriging (a,b,c); (2) regression kriging (a,b,c);
705		and (3) regression co-kriging (a,b,c).
706		

707	Table 1.	Radar data sets in terms of products, acquisition dates and image configurations.
708		Complex pairs used to calculate InSAR coherence are denoted by * and †.
709	Table 2	Calculated model semivariograms and cross-semivariogram used in spatial
710		predictions for each variable for the 1000m sampling interval.
711	Table 3	Evaluation of global accuracy for co-kriging, regression kriging, and regression co-
712		kriging based on the validation dataset. Root Mean Squared Error (RMSE), Mean
713		Absolute Error (MAE), Mean Error or Bias (ME), Pearson's Correlation Coefficient and
714		Coefficient of Determination are used to measure the differences between predicted
715		and observed values.
716		

Table 1. Radar data sets in terms of products, acquisition dates and image configurations. Complex pairs used to calculate InSAR coherence are denoted by * and †.

ID	SAR Sensor	Product	Acquisition date	Incidence angle (deg)	Flight Direction	Polarisations	Ground Resolution (m)
1	PALSAR	FBD	30-Aug-2008	34.3	Ascending	HH+HV	~20
2*		FBD	02-Sep-2009	34.3	Ascending	HH+HV	~20
3*		FBD	18-Jul-2009	34.3	Ascending	HH+HV	~20
4†	RADARSAT-2	Fine-Quad	07-Aug-2010	39.2	Ascending	HH+HV+VH+VV	~8
5†		Fine-Quad	31-Aug-2010	39.2	Ascending	HH+HV+VH+VV	~8

Table 2. Calculated model semivariograms and cross-semivariogram used in spatial predictions for each variable for the 1000m sampling interval.

Semi-Variogram Model (1000 m)	Model	Nugget	Partial Sill	Range (m)
Above-ground Biomass	Whittle-Matern	0.116	0.877	387.493
Radar co-variable	Whittle-Matern	0.156	0.807	519.417
Cross-variogram (biomass and radar)	Whittle-Matern	0.397	0.441	328.268
OLS residuals	Whittle-Matern	0.013	0.801	313.421

Table 3. Evaluation of global accuracy for co-kriging, regression kriging, and regression co-kriging based on the validation dataset. Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Relative Error (MRE), Mean Error or Bias (ME), Pearson's Correlation Coefficient and Coefficient of Determination are used to measure the differences between predicted and observed values.

Estimated Above-ground Biomass						
Prediction Method and Sampling Strategy	RMSE - (Mg ha ⁻¹)	MAE - (Mg ha ⁻¹)	MRE - (%)	ME - (Mg ha ⁻¹)	Multiple R-Squared	Pearson's Correlation (r)
Ordinary Co-kriging						
2000 m transect	234.750	162.370	28.0	-22.730	0.285	0.533
1000 m transect	219.413	142.802	24.6	-34.164	0.383	0.619
500 m transect	205.00	126.758	21.6	-24.362	0.458	0.676
Regression kriging						
2000 m transect	237.759	161.951	28.7	-35.347	0.255	0.505
1000 m transect	218.729	147.638	25.5	-19.523	0.380	0.616
500 m transect	203.900	131.643	22.7	-14.007	0.460	0.678
Regression Co-kriging						
2000 m transect	238.228	164.621	28.4	-37.349	0.276	0.525
1000 m transect	221.940	152.050	26.3	-18.110	0.364	0.604
500 m transect	205.613	134.259	23.2	-15.690	0.449	0.670

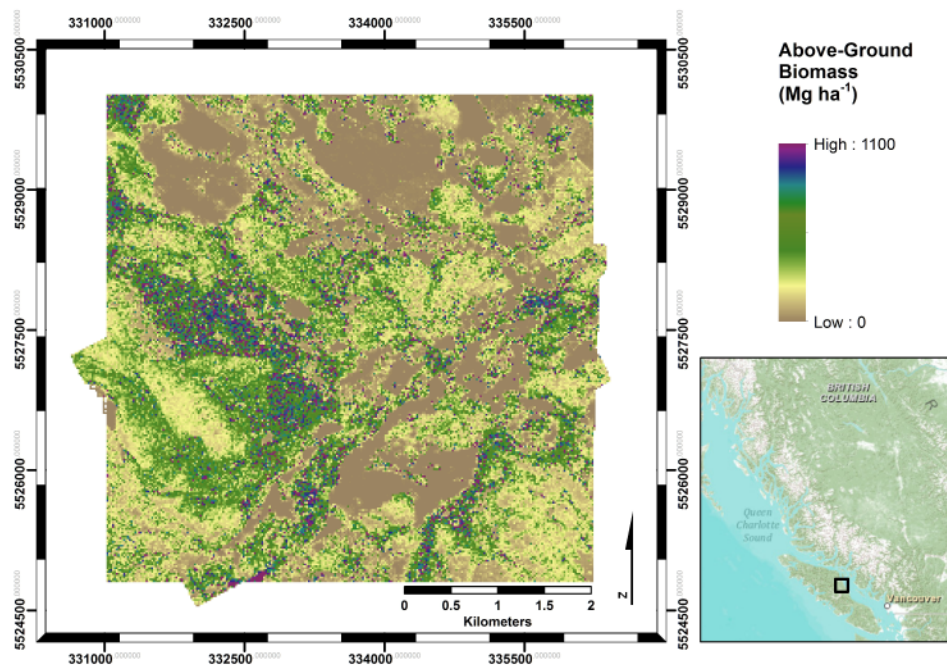


Figure 1

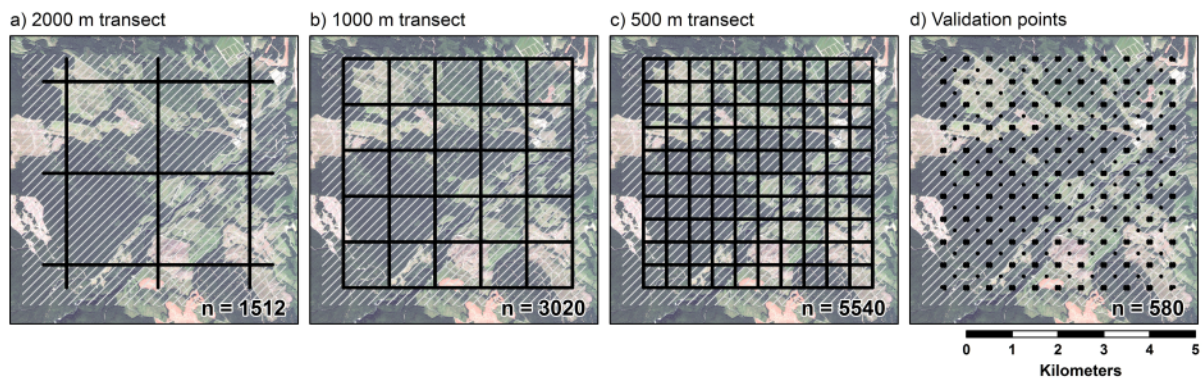


Figure 2

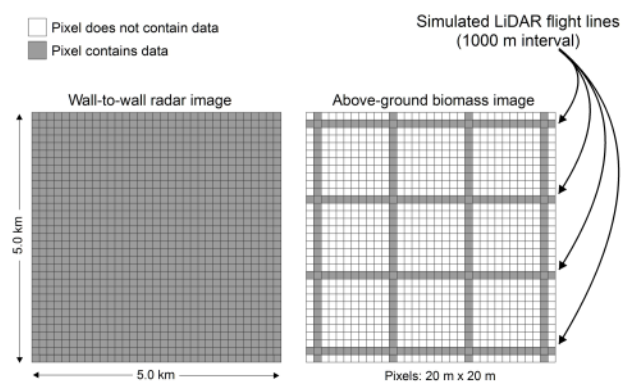


Figure 3

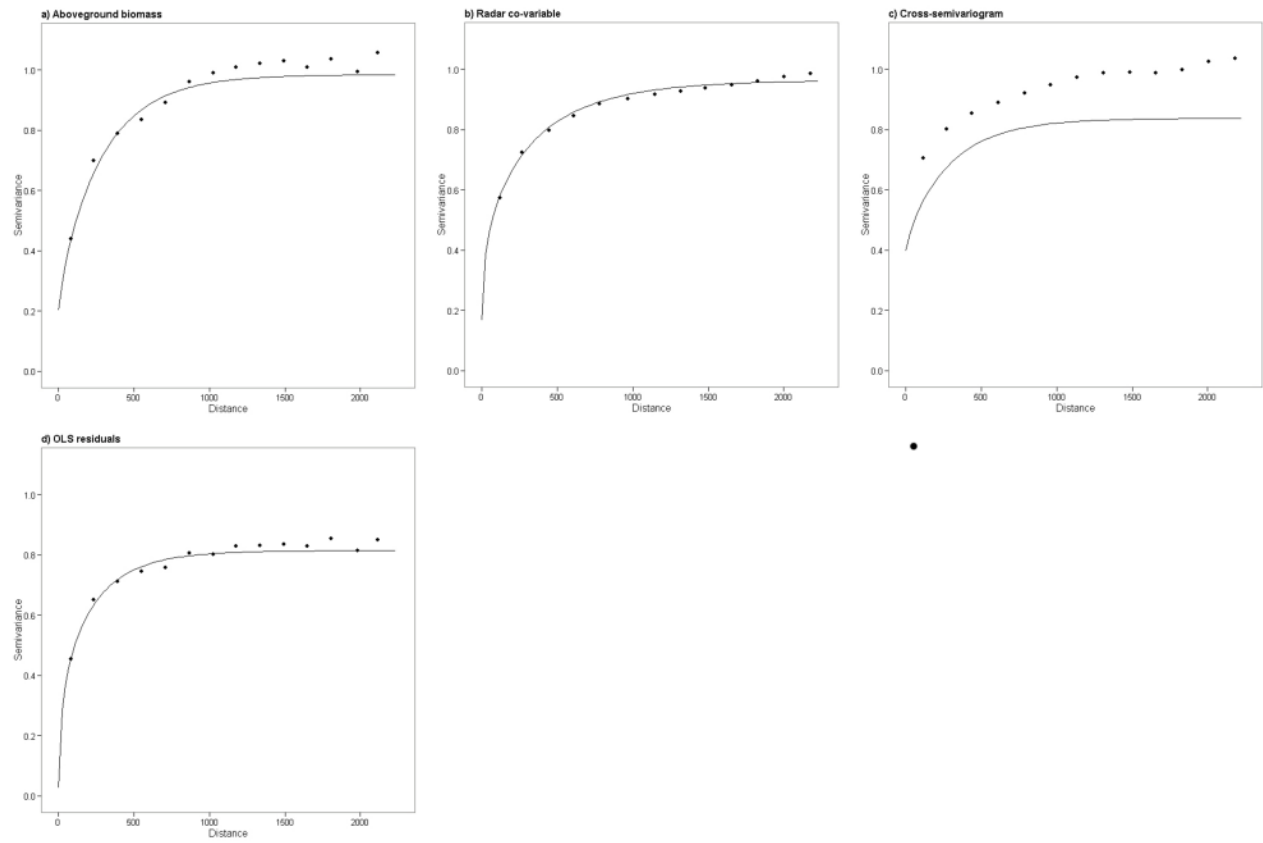


Figure 4

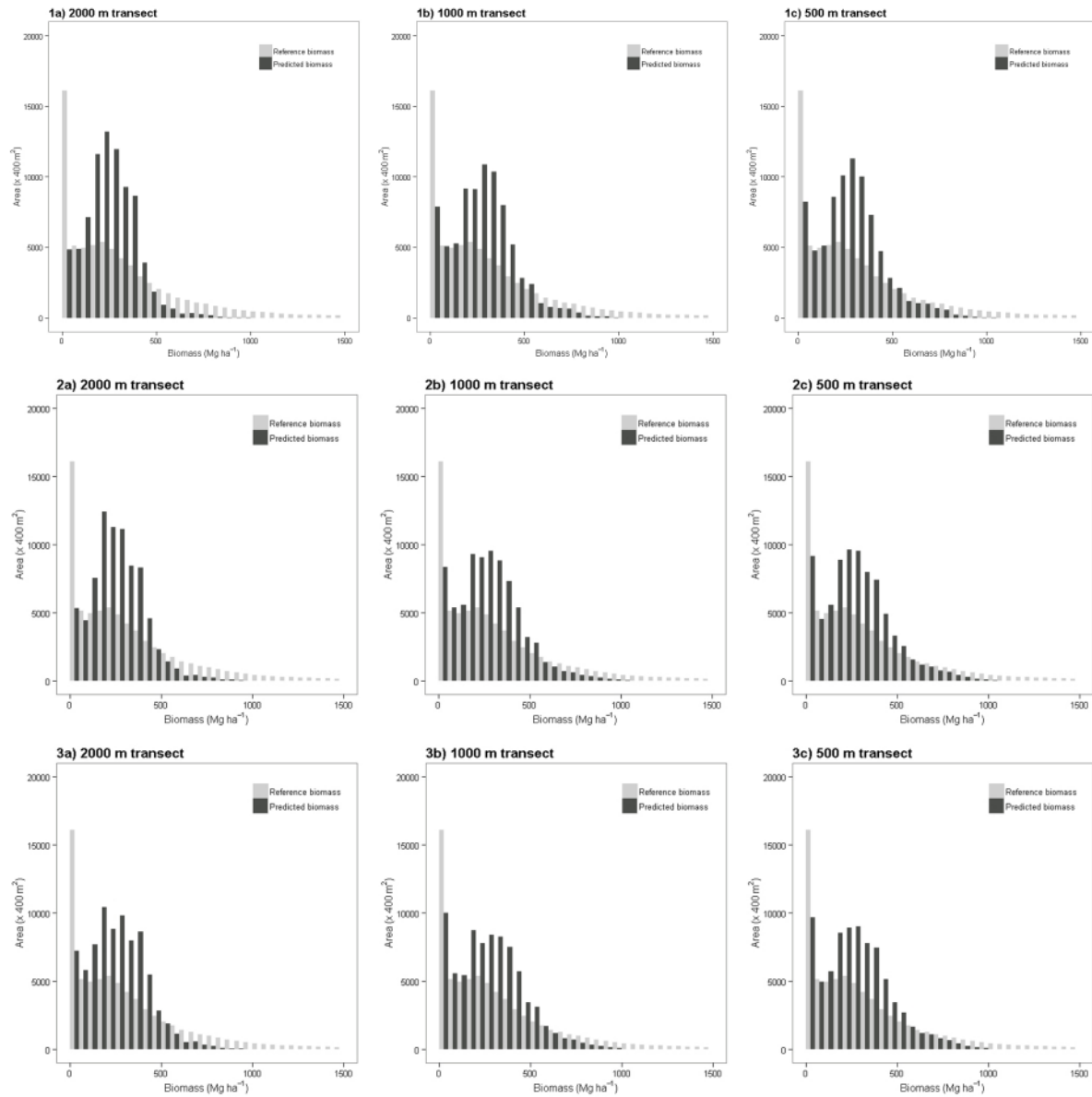


Figure 5

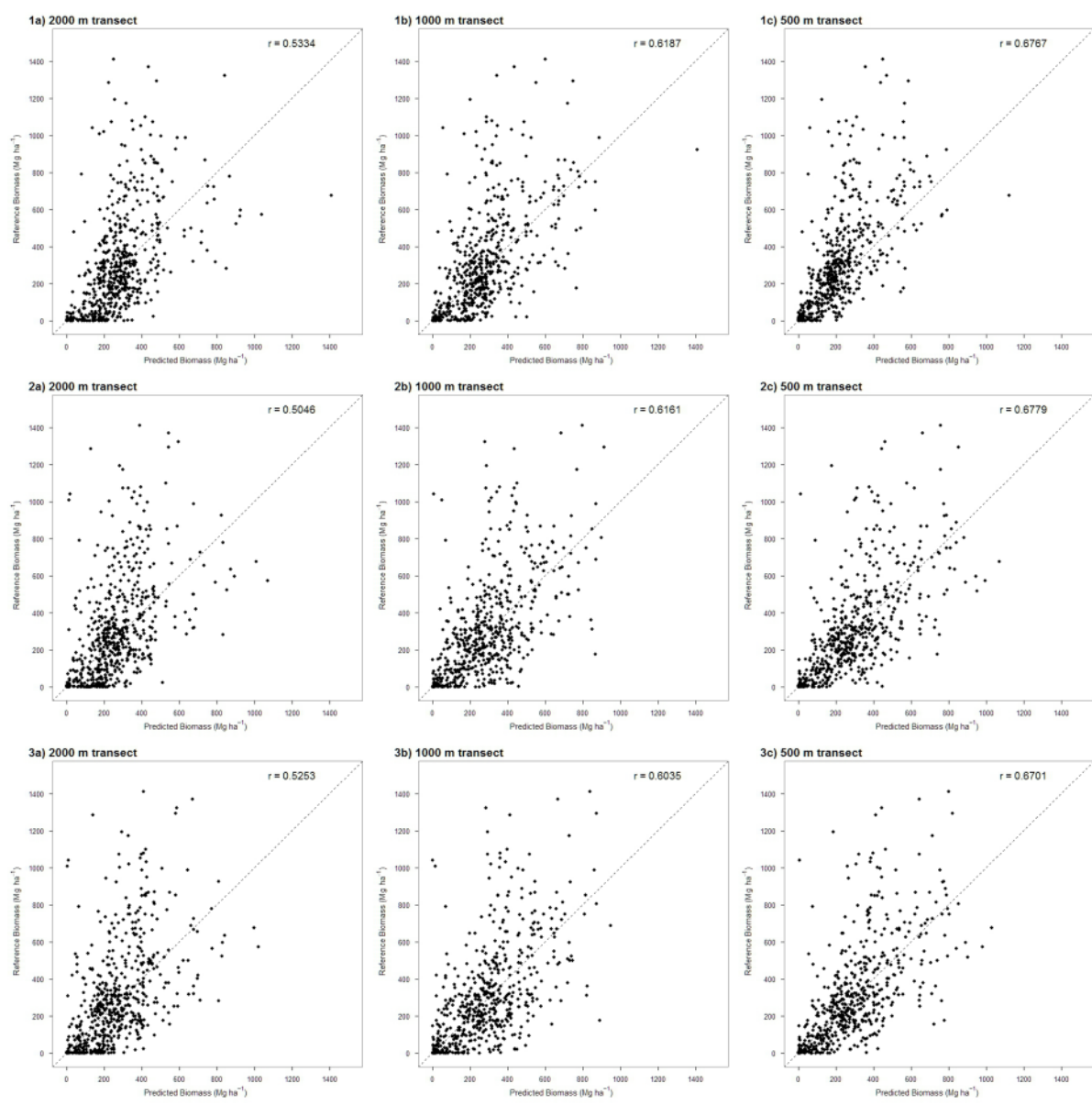


Figure 6

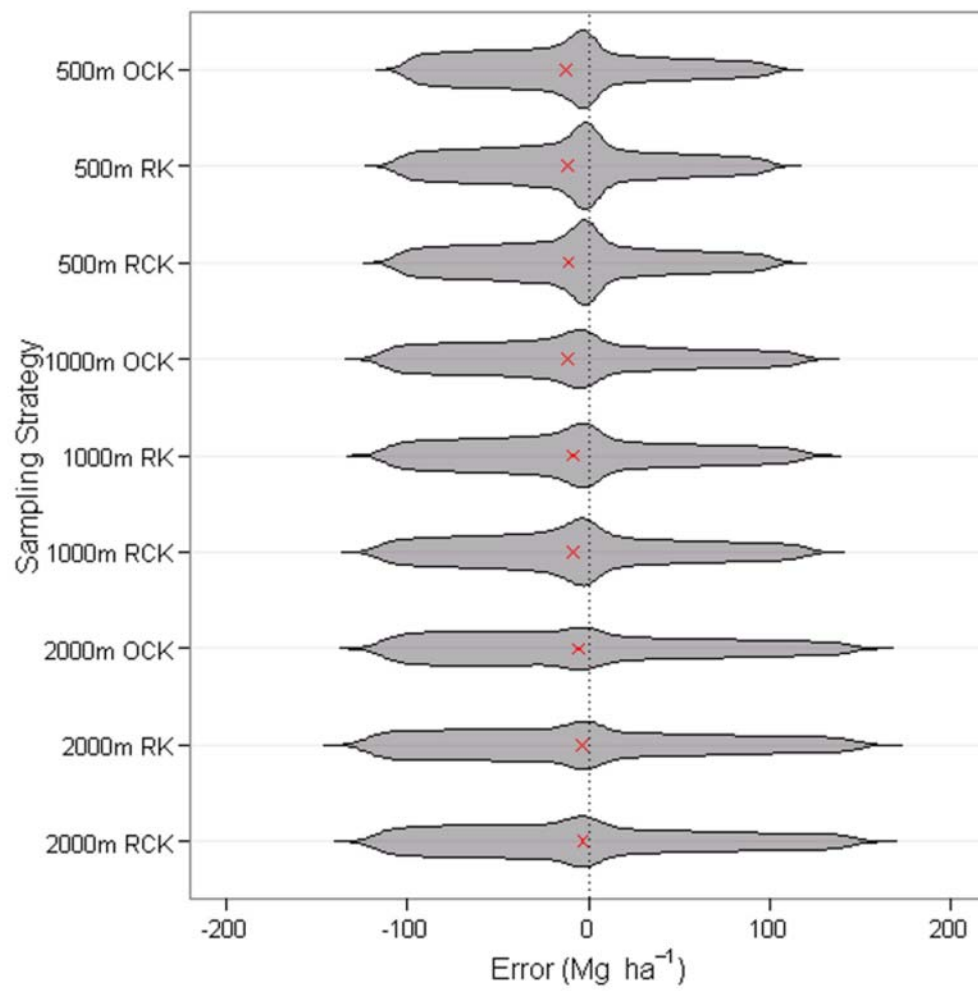
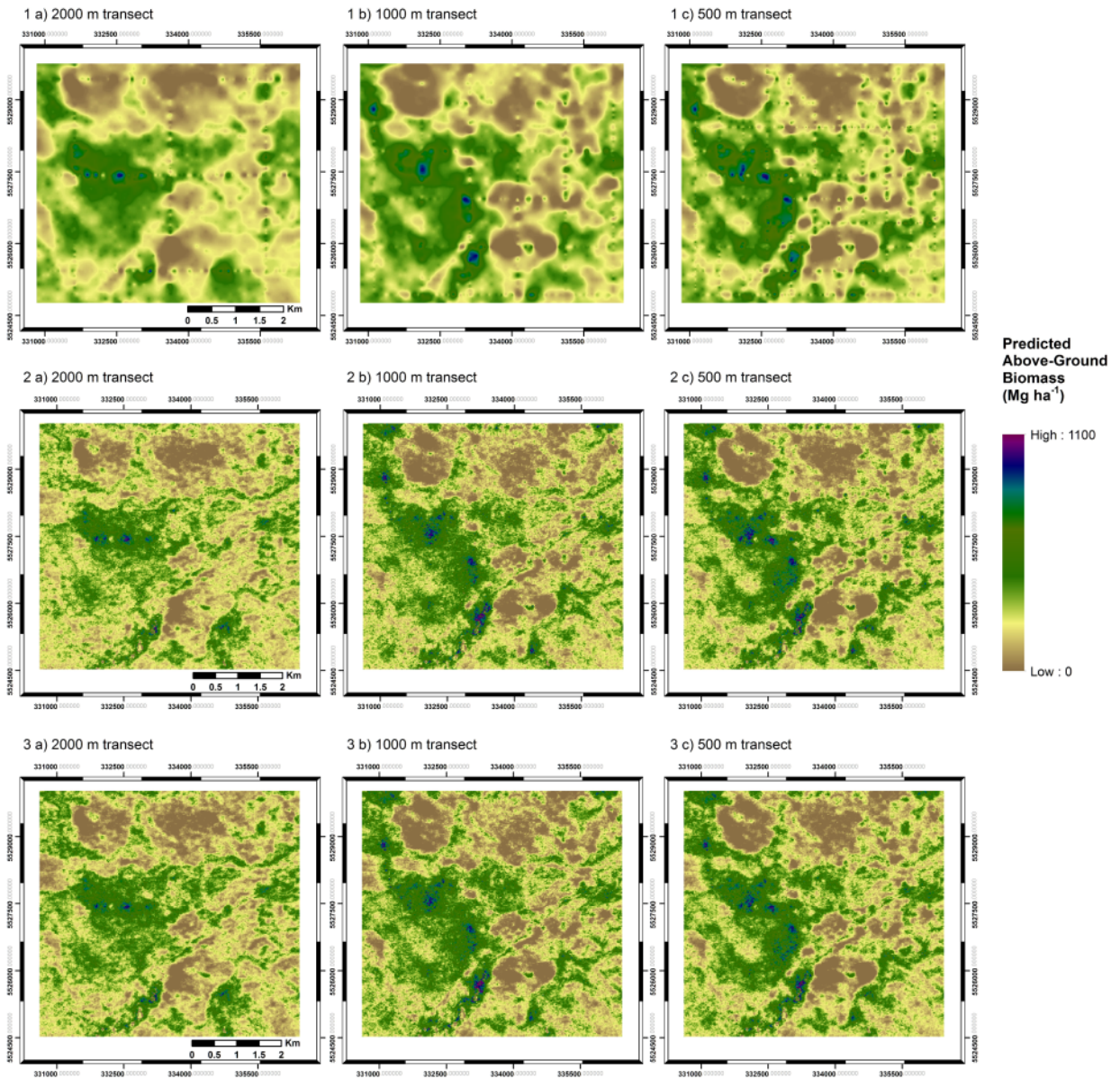


Figure 7



749

750 Figure 8

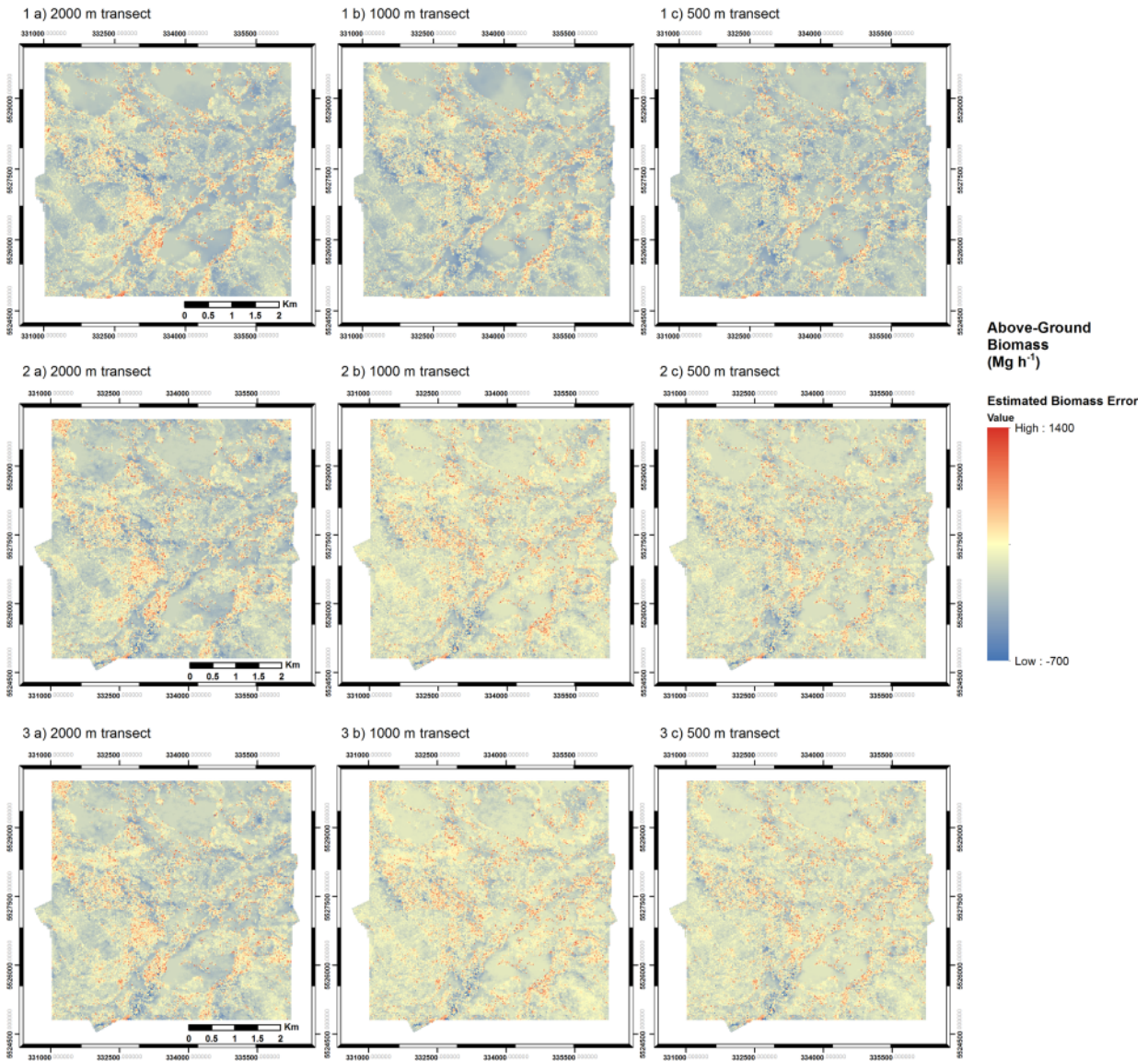


Figure 9