Spatially explicit estimation of aboveground boreal forest biomass in the Yukon River Basin, Alaska

Lei Ji\textsuperscript{a,}\textsuperscript{*}, Bruce K. Wylie\textsuperscript{b}, Dana R. N. Brown\textsuperscript{c,d}, Birgit Peterson\textsuperscript{a}, Heather D. Alexander\textsuperscript{e}, Michelle C. Mack\textsuperscript{f}, Jennifer Rover\textsuperscript{b}, Mark P. Waldrop\textsuperscript{g}, Jack W. McFarland\textsuperscript{g}, Xuexia Chen\textsuperscript{h}, and Neal J. Pastick\textsuperscript{i}

\textsuperscript{a}ASRC InuTeq, contractor to the US Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198, USA; \textsuperscript{b}USGS EROS Center, Sioux Falls, SD 57198, USA; \textsuperscript{c}Department of Biology and Wildlife, University of Alaska Fairbanks, Fairbanks, AK 99775, USA; \textsuperscript{d}Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775, USA; \textsuperscript{e}Department of Biological Sciences, University of Texas at Brownsville, Brownsville, TX 78520, USA; \textsuperscript{f}Department of Biology, University of Florida, Gainesville, FL 32611, USA; \textsuperscript{g}USGS, Menlo Park, CA 94025, USA; \textsuperscript{h}Sigma Space Corporation, VIIRS Characterization Support Team (VCST), Lanham, MD 20706, USA; \textsuperscript{i}Stinger Ghaffarian Technologies, Inc., contractor to USGS EROS Center, Sioux Falls, SD 57198, USA

(Received 25 February 2014; accepted 17 October 2014)

Quantification of aboveground biomass (AGB) in Alaska’s boreal forest is essential to the accurate evaluation of terrestrial carbon stocks and dynamics in northern high-latitude ecosystems. Our goal was to map AGB at 30 m resolution for the boreal forest in the Yukon River Basin of Alaska using Landsat data and ground measurements. We acquired Landsat images to generate a 3-year (2008–2010) composite of top-of-atmosphere reflectance for six bands as well as the brightness temperature (BT). We constructed a multiple regression model using field-observed AGB and Landsat-derived reflectance, BT, and vegetation indices. A basin-wide boreal forest AGB map at 30 m resolution was generated by applying the regression model to the Landsat composite. The fivefold cross-validation with field measurements had a mean absolute error (MAE) of 25.7 Mg ha\textsuperscript{-1} (relative MAE 47.5%) and a mean bias error (MBE) of 4.3 Mg ha\textsuperscript{-1} (relative MBE 7.9%). The boreal forest AGB product was compared with lidar-based vegetation height data; the comparison indicated that there was a significant correlation between the two data sets.

1. Introduction

The boreal forest extends as a broad belt between 50° N and 70° N across Eurasia and North America. As the second biggest terrestrial ecosystem in the world after tropical forests, the boreal forest biome occupies 33% of the forest area and holds 23% of total carbon stores on the Earth’s land surface (Carlson, Wells, and Roberts 2009; Food and Agriculture Organization of the United Nations 2001; Intergovernmental Panel on Climate Change 2000). The Yukon River Basin in central Alaska and northwestern Canada is located in the westernmost extent of the boreal forest biome in North America. Forests and shrublands compose about 70% of land cover in this basin (Brabets, Wang, and Meade 2000). Because biomass is a key biophysical parameter in evaluating and modeling carbon stocks and cycles, an accurate estimation of regional biomass is important for

\textsuperscript{*}Corresponding author. Email: lji@usgs.gov
understanding the ecosystems and their dynamics in the boreal forest of the Yukon River Basin.

Assessing and monitoring regional aboveground biomass (AGB) rely heavily on field investigation and sampling, which in general are affordable and applicable for only a limited number of sites. For an expansive and remote region like the Yukon River Basin, remote-sensing techniques are practical and effective alternatives for spatially explicit estimation of AGB at a regional scale. For Alaska, a statewide forest AGB map is available from the Forest Biomass across the Lower 48 States and Alaska data set (Blackard et al. 2008; http://webmap.ornl.gov/biomass/biomass.html), produced by the US Department of Agriculture (USDA) Forest Service Forest Inventory and Analysis (FIA) programme and the Remote Sensing Applications Centre. This product provides estimates of live forest AGB at 250 m resolution for the conterminous USA, Alaska, and Puerto Rico. Production of this map was based primarily on FIA inventory data collected from 1990 to 2003 and Moderate Resolution Imaging Spectroradiometer (MODIS) data acquired in 2001, climate data, and other geospatial data (land cover and topography) (Blackard et al. 2008). Although this data set offers a regional AGB estimate, it does not capture spatially detailed AGB variations, especially those associated with historical fires. In 2011, we mapped AGB at 30 m resolution in the Yukon Flats ecoregion of interior Alaska based on Landsat data and the field data collected between 2008 and 2010 (Ji et al. 2012). However, the mapped area was only 33,400 km², which is about 2% of Alaska’s total area.

Our objective in this study was to produce a 30 m resolution boreal forest AGB data set for the Yukon River Basin of Alaska based on Landsat data and field observations acquired from recent years (2008–2010). Specifically, our goals were to (1) create empirical relationships between Landsat-derived spectral indices and AGB derived from field measurements, (2) apply the empirical model to the Landsat images to produce regional AGB estimates and assess the accuracy of the AGB estimates, and (3) analyse the spatial patterns in the regional AGB estimates in relation to vegetation types and fire history. The production of a basin-wide boreal forest AGB data set will provide an important biophysical parameter for the modelling and investigation of Alaska’s ecosystems.

2. Study area and data

2.1. Study area

Our study area is the Alaskan portion of the Yukon River Basin, lying between 60°55′48″ N and 69°04′38″ N and between 141°00′00″ W and 166°11′50″ W (Figure 1). The area of the Alaskan Yukon River Basin is 511,900 km², approximately 61% of the entire basin (Bailey 2005). The Yukon River, along with three major tributaries, the Porcupine River, the Koyukuk River, and the Tanana River, is the major watercourse in the Alaska portion of the drainage basin. The topography of the study area is characterized by the interior highlands, lowlands, and flat plains surrounded by mountains in the north (Brooks Range) and the southeast (Alaska Range). According to the 2001 National Land Cover Database (NLCD 2001) (Homer et al. 2007), the land-cover types in the Yukon River Basin consist of evergreen forest (28.4%), deciduous forest (7.1%), mixed forest (5.9%), shrubs/scrub (28.6%), dwarf shrub (8.6%), sedge/herbaceous/moss (1.6%), woody wetlands (7.3%), emergent herbaceous wetlands (2.5%), barren land (5.4%), perennial ice/snow (1.2%), cultivated crops (0.04%), developed area (0.08%), and open water (2.9%). Forests are
dominated by black spruce (\textit{Picea mariana}), white spruce (\textit{Picea glauca}), quaking aspen (\textit{Populus tremuloides}), paper birch (\textit{Betula neoalaskana}), and balsam poplar (\textit{Populus balsamifera}). The tall and low shrub communities are dominated by willows (\textit{Salix} spp.), alder (\textit{Alnus} spp.), narrow-leaf Labrador tea (\textit{Ledum decumbens}), and ericaceous species. The basin lies in the continental climate zone, with the exception of the Coastal Flats Plains. Based on the climate data from 43 weather stations recorded since the 1980s (http://www.wrcc.dri.edu/), the average annual total precipitation and snowfall are 402 mm and 1593 mm, respectively, and the average monthly temperature is \(-3.9^\circ\text{C}\) in
the interior basin. Wildfires are common within the region. Records from the Wildland Forest Dataset for Alaska (by the Bureau of Land Management, Alaska Fire Service, http://agdc.usgs.gov/data/blm/fire/) indicate that approximately 160,000 km$^2$ of the study area (or about 31% of the basin) was burned by one or more wildfires between 1960 and 2010.

2.2. Field data

The field AGB data were collected separately by two field teams led by the US Geological Survey (USGS) and the University of Florida. Three field campaigns were carried out by the USGS team in the areas near Boot Lake, Canvasback Lake, Hat Lake, Hornet Lake, Nome Creek, and other areas in August and September 2009 and June to September 2010 (Figure 1). The USGS’s field observations had 90 plots, including 70 plots designed with each plot consisting of three 14 m long transects separated by a distance of 16 m between transects and 20 plots designed as 23 m long transects. The field survey by the University of Florida was conducted in the Tanana River Basin in June 2008, June and July 2009, and June 2010. They collected data at 45 field plots located primarily in boreal forest that burned 20–59 years previously and were in various statuses of post-fire succession (Alexander et al. 2012). The design used a 100 m long linear transect for each plot. Both USGS and the University of Florida teams recorded diameter at breast height (DBH) for tall trees (those taller than 1 m) and basal diameter (BD) for small trees and shrubs within each transect. Coarse woody debris (CWD) for each plot was estimated with a line intersect method (measuring the diameter of dead wood) along each transect (Van Wagner 1968).

Although the sampling sites were not evenly distributed across the study area because of the limited transportation network, the site selection considered a wide range of environmental conditions, such as different vegetation types, a large range of biomass, and various topographic features. With this sampling strategy, we could reduce the similarity between adjacent sampling locations, minimize spatial autocorrelation issues in statistical analyses, and enhance sample representativeness within the region.

2.3. Remote-sensing data

In this study, we used the Landsat data to map the boreal forest AGB for the Alaskan Yukon River Basin, which needs 72 Landsat path/rows to cover the entire region. As the best quality imagery available for each path/row should be used when using multi-date Landsat data, a fairly large number of Landsat scenes would be required to map AGB for the retire region. To efficiently obtain the best available Landsat data, we acquired Landsat Enhanced Thematic Mapper Plus (ETM+) images from the Web Enabled Landsat Data (WELD) product (version 1.5) (Roy et al. 2010, 2011). WELD provides mosaics of multi-date composites (weekly, monthly, seasonal, and annual) of top-of-atmosphere (TOA) reflectance and brightness temperature (BT) at 30 m resolution for the conterminous USA and Alaska, processed using the terrain-corrected and radiometrically calibrated ETM+ data. To minimize the remaining cloud contamination and effect of filling Scan-Line Corrector-off gaps existing in the WELD data, we used the 2008, 2009, and 2010 annual WELD data to create a 3-year composite (2008–2010) for the Yukon River Basin of Alaska (Figure 1). The compositing algorithm followed the WELD compositing criteria that include non-saturation, non-cloudiness, maximum normalized difference vegetation index (NDVI), and maximum BT (Roy et al. 2011). This algorithm used for the 3-year
compositing procedure, however, may conceal the burns occurring in the years 2008–2010. Therefore, for the areas burned during these 3 years, we used the ‘minimum NDVI value’ method to select pixels from the three annual WELD images. The burned areas were delineated using the Monitoring Trends in Burn Severity (MTBS) product, which provides a burn severity data set at 30 m spatial resolution for all fires (with an area greater than 4 km$^2$) since 1984 in Alaska (http://www.mtbs.gov/). The land-cover types in the basin were identified using the NLCD 2001 product, which was developed with a 20-category classification scheme at 30 m resolution for Alaska. The NLCD open water class was further modified using ETM+ data acquired from 1999 to 2002 (Striegl et al. 2012).

Two light detection and ranging (lidar) data sets were collected at 3800 m above mean terrain with an aircraft-carried Optec ALTM Gemini system operated by Aero-Metric, Inc. One lidar data set was acquired on 14–16 July and 3 September 2009, for an area of about 2600 km$^2$ centred in the Yukon Flats (Figure 1). The data set had a horizontal accuracy of 1.15 m with a nominal point spacing of 2.3 m and a vertical positional accuracy of 0.10 m. Another data set was acquired on 10 October 2010 for an area of about 200 km$^2$ located along Nome Creek (Figure 1), which had a horizontal accuracy of 1.05 m with a nominal point spacing of 2.1 m and a vertical accuracy of 0.067 m. Aero-Metric, Inc., processed the raw data and delivered 2.5 m resolution raster data sets that included the bare-earth digital surface model (DSM) data and first-return DSM data.

3. Methods

3.1. Calculation of field AGB

In this study, the total boreal AGB was defined as the sum of tree and shrub AGB and CWD AGB. We estimated the tree and shrub AGB using field measurements of DBH or BD, using the allometric equations developed for the western North American boreal forest (Bond-Lamberty, Wang, and Gower 2002; Mack et al. 2008; Yarie 2007; Yarie, Kane, and Mack 2007). The biomass estimates of each tree or shrub (taller than 1 m) were then summed for each transect within a plot. We sampled CWD with the line intersect method along each transect and calculated the biomass using Van Wagner’s equation (Van Wagner 1968).

3.2. Development of regression model and regional AGB mapping

In the statistical analysis, one observation (or sample) consists of AGB measured at plot-level and spectral values extracted from the Landsat WELD images at the locations corresponding to the plot. Although a large number of field plots are located intensively within five areas, i.e. Canvasback Lake, Boot Lake, Hornet Lake, Hat Lake, and Nome Creek (Figure 1), we pre-selected the sampling sites with a great variation of vegetation density and types to minimize spatial autocorrelation. To verify the data independency, we tested spatial autocorrelation using Moran’s $I$ and Geary’s $C$ statistics for each of the five areas (Geary 1954; Moran 1950).

We created a regression model using the spectral variables as predictor variables and the field-derived AGB as the response variable. To minimize potential image misregistration, we extracted the spectral values from a 3 pixel × 3 pixel window and used the average values of the nine pixels in the window. The spectral variables include TOA reflectances in band 1 (0.45–0.52 µm), band 2 (0.52–0.60 µm), band 3 (0.63–0.69 µm), band 4 (0.76–0.90 µm), band 5 (1.55–1.75 µm), and band 7 (2.08–2.35 µm), BT, NDVI International Journal of Remote Sensing 943 Downloaded by [University of Alaska Fairbanks] at 22:19 23 February 2015
(Rouse et al. 1974), enhanced vegetation index (EVI) (Huete, Justice, and Liu 1994), normalized difference infrared index (NDII) (Hardisky, Klemas, and Smart 1983; Hunt and Rock 1989), and normalized difference water index (NDWI) (Ji, Zhang, and Wylie 2009; McFeeters 1996; Xu 2006). For this study, the NDII and the NDWI are defined as

\[
\text{NDII} = \frac{\rho_4 - \rho_5}{\rho_4 + \rho_5},
\]

\[
\text{NDWI} = \frac{\rho_2 - \rho_7}{\rho_2 + \rho_7},
\]

where \(\rho_2, \rho_4, \rho_5,\) and \(\rho_7\) are the TOA reflectances in bands 2, 4, 5, and 7, respectively. We incorporated NDVI, EVI, NDII, and NDWI into the model because these indices are sensitive to land surface characteristics such as vegetation quantity, greenness, and moisture condition, which are generally correlated with forest biomass.

AGB demonstrated a nonlinear relationship against reflectance and spectral indices. To linearize this relationship, we transformed AGB to a natural logarithmic form. The final regression model is expressed as

\[
\ln(\text{AGB}) = b_0 + b_1(\rho_1) + b_2(\rho_2) + b_3(\rho_3) + b_4(\rho_4) + b_5(\rho_5) + b_7(\rho_7) + b_8(\text{BT}) + b_9(\text{NDVI}) + b_{10}(\text{EVI}) + b_{11}(\text{NDII}) + b_{12}(\text{NDWI}) + \varepsilon,
\]

where \(\ln(\text{AGB})\) is natural logarithmic transformed AGB; \(b_0, \ldots, b_{12}\) are the regression coefficients; \(\rho_1, \ldots, \rho_5,\) and \(\rho_7\) are the TOA reflectances in bands 1, 5, and 7, respectively; and \(\varepsilon\) is the random error. We notice that the predictor variables are highly correlated, which may cause the multicollinearity problem in the model. Existence of multicollinearity does not affect the estimation of the dependent variables, although it tends to inflate the variances of the parameter estimates and the variances of the predicted values (Freund and Littell 1991; Rawlings, Pantula, and Dickey 2001; Kroll and Song 2013). Our interest in the study was to predict AGB values rather than to examine and interpret the effect of individual bands or vegetation indices on AGB. To increase the prediction accuracy and improve the model performance, we included all the variables in the regression model.

The regression model was then applied to the Landsat image mosaic (2008–2010 composite) to generate an AGB map for the Yukon River Basin. NLCD 2001 data were used to identify boreal forest including the NLCD classes deciduous forest, evergreen forest, mixed forest, shrublands (dwarf scrub is not included), and woody wetlands. In the final map, the AGB estimates are only shown for the boreal forest, while all other classes were masked out.

### 3.3. Validation of the AGB model and map

Because of the limited number of field plots relative to the large study area, we applied a fivefold holdout cross-validation to evaluate the accuracies of the regression model and the AGB map. All the samples were randomly split into five subsets and the cross-validation was repeated five times. With each repetition, one subset was held out as the validation set and the remaining subsets were used as the training data. All field AGB values were then compared to the model-estimated AGB values based on the held-out
training samples. The mean absolute error (MAE) and the mean bias error (MBE) were used to compare the field-observed and model-predicted AGB values:

\[
\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |O_i - P_i|,
\]

\[
\text{MBE} = \frac{1}{n} \sum_{i=1}^{n} (O_i - P_i) = \overline{O} - \overline{P},
\]

where \(O_i\) and \(P_i\) are the field-observed and model-predicted values, respectively, for sample \(i\); \(n\) is the number of samples; and \(\overline{O}\) and \(\overline{P}\) are the mean values of all observed and predicted values, respectively. Both MAE and MBE measure the estimation error by comparing the observed and estimated AGB values. MBE measures the systematic error or bias between the observations and estimates, while MAE measures the unsystematic error or random error between the two quantities. We ran the fivefold cross-validation with different combinations of random samples 1000 times and reported the average MAE and MBE values of the 1000 results.

To further assess the accuracy of the AGB map, we used two airborne lidar images acquired from the Yukon Flats and Nome Creek as supplementary validation data sets (Figure 1). In the lidar image, the vegetation height was calculated as the difference between the bare-earth DSM data and the first-return DSM data. The comparison was done by constructing a regression model that related the AGB estimates with the lidar-derived vegetation height. We assume that AGB is positively correlated to tree and shrub height, which thus serves as a relative, rough, and synoptic proxy for canopy biomass. According to the field observations, the primary tree species in interior Alaska have strong linear or nonlinear relationships between tree height and AGB (Mack et al. 2008; Yarie, Kane, and Mack 2007). Additionally, the spatial extent of the lidar data is much larger than field observation sites so the lidar-derived vegetation height is able to assess the consistency and robustness of the estimated AGB in space.

4. Results and discussion

4.1. Regression model and AGB estimation

For the initial model development, we used the field data from 133 plots, with exclusion of two plots located in the northern part of the Boot Lake area, which were burned in early August 2009 during the Big Creek Fire. We tested the spatial autocorrelation of AGB for each clustered sampling area using Moran’s I and Geary’s C statistics (Table 1). All the observed Moran’s I and Geary’s C values were very close to the expected values (except Moran’s I statistic for Hornet Lake), indicating no or very low spatial autocorrelation for each area. Based on the \(p\)-values, which were mostly greater than 0.05, we failed to reject the null hypothesis that there was no spatial autocorrelation among the AGB samples in each area.

As expressed in Equation (3), a multiple regression model was created that regressed the natural logarithmic transformed AGB against the predictor variables, TOA reflectances, BT, and spectral indices. We detected four outliers in the model using influence statistics (e.g. studentized residuals) and eliminated these outliers because they severely biased the regression estimates. The four outliers were located in the grassland where AGB values are very low. The final regression model estimated with 129 samples was
significant ($p$-value < 0.0001) and had fairly high goodness-of-fit (coefficient of determination, $R^2 = 0.703$ and adjusted coefficient of determination, $R^2_{adj} = 0.675$). Figure 2 illustrates the relationship between the observed ln(AGB) and model-estimated ln(AGB), which shows a fairly good agreement between the two variables. Although the model was biased at low values of AGB (i.e. for AGB < 10 Mg ha$^{-1}$), our AGB mapping was primarily focused on the boreal forest.

Table 1. Moran’s $I$ and Geary’s $C$ statistics for testing spatial autocorrelation in AGB observations.

<table>
<thead>
<tr>
<th>Area</th>
<th>Number of observations</th>
<th>Statistic</th>
<th>Observed value</th>
<th>Expected value</th>
<th>Standard deviation</th>
<th>z-value</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canvasback Lake</td>
<td>10</td>
<td>Moran’s $I$</td>
<td>−0.175</td>
<td>−0.111</td>
<td>0.160</td>
<td>−0.397</td>
<td>0.6911</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Geary’s $C$</td>
<td>1.056</td>
<td>1.000</td>
<td>0.193</td>
<td>−0.290</td>
<td>0.7718</td>
</tr>
<tr>
<td>Boot Lake</td>
<td>10</td>
<td>Moran’s $I$</td>
<td>−0.157</td>
<td>−0.111</td>
<td>0.099</td>
<td>−0.466</td>
<td>0.6414</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Geary’s $C$</td>
<td>0.855</td>
<td>1.000</td>
<td>0.149</td>
<td>−0.971</td>
<td>0.3317</td>
</tr>
<tr>
<td>Hornet Lake</td>
<td>13</td>
<td>Moran’s $I$</td>
<td>0.271</td>
<td>−0.083</td>
<td>0.118</td>
<td>3.000</td>
<td>0.0027</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Geary’s $C$</td>
<td>0.910</td>
<td>1.000</td>
<td>0.178</td>
<td>−0.504</td>
<td>0.6145</td>
</tr>
<tr>
<td>Hat Lake</td>
<td>17</td>
<td>Moran’s $I$</td>
<td>−0.007</td>
<td>−0.064</td>
<td>0.064</td>
<td>0.868</td>
<td>0.3855</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Geary’s $C$</td>
<td>0.948</td>
<td>1.000</td>
<td>0.069</td>
<td>−0.753</td>
<td>0.4517</td>
</tr>
<tr>
<td>Nome Creek</td>
<td>18</td>
<td>Moran’s $I$</td>
<td>−0.069</td>
<td>−0.059</td>
<td>0.091</td>
<td>−0.110</td>
<td>0.9123</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Geary’s $C$</td>
<td>0.870</td>
<td>1.000</td>
<td>0.110</td>
<td>−1.160</td>
<td>0.2459</td>
</tr>
</tbody>
</table>

Figure 2. Scatterplot of regression-estimated AGB against observed AGB.
The regression model was applied to the Landsat image mosaic (2008–2010 composite) of the Yukon River Basin for estimating the regional AGB (Figure 3). In the AGB map, water and other non-boreal forest pixels were masked out using the NLCD 2001 data. The fire disturbances in the AGB estimates were updated with the MTBS fire severity data up to the year of 2009. Therefore, our mapping product is an estimation of boreal forest AGB in 2009 at 30 m resolution in the Yukon River Basin of Alaska. The digital AGB data product is available for download from the USGS Earth Resources Observations and Science (EROS) Land Cover Applications and Global Change Website (http://lca.usgs.gov/lca/). The basin-wide boreal forest AGB data set provides an important biophysical parameter for the modelling and investigation of Alaska’s ecosystems.

4.2. Accuracy assessment

We performed a fivefold holdout cross-validation for the 129 field plots used for developing the regression model and mapping AGB. The cross-validation generated MAE and MBE values by using Equations (4) and (5). The holdout cross-validation procedure was repeated for 1000 trials, and the averaged MAE and MBE for the 1000 results were 25.7 and 4.3 Mg ha$^{-1}$, respectively. Using the mean AGB value of 54.1 Mg ha$^{-1}$ for the 129 samples as a reference, we calculated the relative errors: relative MAE was 47.5% and relative MBE 7.9%. Relative MAE and relative MAE, respectively, are the ratios of MAE and MBE to the mean AGB value.

To assess the consistency of the AGB map, we compared the AGB estimates with the lidar-derived vegetation height data at 2.5 m resolution. We degraded the lidar-derived vegetation height map to 30 m resolution to match the pixel size of the AGB map. Each
pixel of the 30 m resolution image contains 144 pixels of the 2.5 m resolution image. So within each 30 m pixel, we calculated the mean and maximum values, respectively, of the 144 vegetation height values and generated the mean and maximum canopy height maps at 30 m resolution. Using all 2,830,646 tree and shrub pixels, we created the density scatterplots for the relationships between the AGB estimates and the mean/maximum vegetation height (Figures 4(a) and (b)). In the scatterplots, the AGB values are correlated nonlinearly with both mean and maximum canopy heights. We noticed that at the low canopy height end (mean and maximum heights less than 1 m), the AGB values are widely scattered, ranging from < 1 Mg ha\(^{-1}\) to about 50 Mg ha\(^{-1}\). This phenomenon is probably caused by poor correlation between the biomass of CWD and undersampled low trees/shrubs and the vegetation height derived by the lidar data.

Because of the nonlinear relationships, we created three exponential models for the empirical relationships between the AGB and the mean and maximum heights:

\[
\ln(\text{AGB}) = 2.3454 - 1.5164\left(1 - e^{-1.4584H_{\text{mean}}}\right),
\]

\[
\ln(\text{AGB}) = 2.222 - 2.2067\left(1 - e^{-0.1014H_{\text{max}}}\right),
\]

\[
\ln(\text{AGB}) = 2.2516 - 1.8692\left(1 - e^{-0.4463H_{\text{mean}}-0.0704H_{\text{max}}}\right),
\]

where \(H_{\text{mean}}\) and \(H_{\text{max}}\) are the maximum and mean canopy heights, respectively. The statistical test of the three models showed that all models were significant (\(p\)-value < 0.0001). Comparing the goodness-of-fit and parsimony of the three models, the first model (Equation (6)) has a pseudo-\(R^2\) of 0.384, root mean squared error (RMSE) of 0.528, and Akaike’s information criteria (AIC) of \(-1.61 \times 10^7\); the second model (Equation (7)) has pseudo-\(R^2\) of 0.381, RMSE of 0.497, and AIC of \(-1.62 \times 10^7\); the last model (Equation (8)) has pseudo-\(R^2\) of 0.397, RMSE of 0.492, and AIC of \(-1.62 \times 10^7\). Therefore, the last model, which has two independent variables, had a slightly higher pseudo-\(R^2\), lower RMSE, and similar AIC value to the other models. Given the significant relationship between AGB estimates and the lidar-estimated canopy height,

![Figure 4](image_url)

Figure 4. Density scatterplot of the Landsat-derived AGB estimates versus lidar-derived mean canopy height (a) and maximum canopy height (b).
we considered that our total AGB map was fairly consistent with the vegetation height patterns throughout the study area.

The accuracy assessment of the model and the product indicates a certain level of uncertainty in the AGB estimates. Several factors may contribute to the uncertainties. First, optical remote-sensing techniques have limitations for accurately estimating AGB because optical sensors mainly observe upper canopies that reflect only two-dimensional vegetation structures. A large portion of the uncertainties are probably related to forest understory components, primarily shrubs, young trees, and CWD. Second, the plot-level AGB for trees and shrubs is estimated using allometric functions (Bond-Lamberty, Wang, and Gower 2002; Mack et al. 2008; Yarie 2007; Yarie, Kane, and Mack 2007), and the CWD biomass is calculated using an approximation equation (Van Wagner 1968). The errors in the allometric functions will propagate to the regression model and further to the AGB product. Third, each field plot is represented by the transect where the trees and shrubs are measured to estimate the AGB for the entire plot. The heterogeneity of the vegetation type and structure within each transect or each field plot will affect the reliability of the AGB estimates. Fourth, possible minor misregistration in the Landsat data could cause alignment issues between the pixels and the field plots. Last, errors could be induced by the image quality, including cloud contamination, multi-date compositing, Scan-Line Corrector-off effect, and image resampling. Because all the field plots are located in the central and eastern portion of the basin, the accuracy of the AGB estimates is less reliable in the western extend of the basin, especially in the Yukon–Kuskokwim Delta and the Yukon Lowlands area. Higher uncertainties may also exist at high mountains on the north and south edges of the basin (e.g. Brooks Range and Alaska Range).

4.3. Regional AGB distribution and pattern

The statistics of AGB estimates for boreal forest and all different vegetation types in the Yukon River Basin are presented in Table 2. Based on the 5% and 95% quantiles, 90% of the study area has estimated AGB values between 2.2 and 90.0 Mg ha\(^{-1}\). Comparing among different vegetation types, all forests have higher mean AGB values (44.6–51.6 Mg ha\(^{-1}\)), woody wetlands and burned areas (burned between 1990 and 2009) have lower mean AGB values (22.6 and 23.8 Mg ha\(^{-1}\), respectively), and the lowest AGB estimates are found in shrubland (11.9 Mg ha\(^{-1}\)). Figure 5 is the frequency chart of the AGB estimates for the different cover types in the basin. Deciduous, evergreen, and mixed forests have similar AGB distributions, but evergreen forest has a larger area than deciduous and mixed forests. Shrublands encompass a large area, but the estimated AGB values in the shrublands are the lowest among all vegetation types. Because boreal forest covers an area of 399,019 km\(^2\) (or 75.7%) in the Yukon River Basin of Alaska, we estimate that the total boreal forest AGB is 1.15 × 10\(^9\) Mg in the basin (Table 2). If we convert the AGB to total carbon by multiplying AGB by a standard conversion factor of 0.5 (Houghton 2005; Myneni et al. 2001), we can estimate that total carbon (C) measures 5.8 × 10\(^8\) Mg C in the boreal forest of the basin. We estimate that 46.2% of the total AGB in the boreal forest is from the evergreen forest, which makes up 28.9% of the boreal forest. Although the shrubland area is very large (30.3% of the forest), its total AGB is relatively small (11.7%). In general, the AGB pattern in the basin follows the distribution of the land-cover types and burned areas.

The regional patterns of AGB and carbon distributions in the Yukon River Basin could be analysed with the coarse-resolution AGB map (Blackard et al. 2008). However, the fragmented landscape, which contains many small lakes, wetlands, riparian vegetation,
Table 2. Statistics and distribution of the AGB estimates in the Yukon River Basin.

<table>
<thead>
<tr>
<th>Vegetation cover type</th>
<th>Area (km²) and percentage of boreal forest area (%)</th>
<th>Total AGB (Mg) and percentage of boreal forest total (%)</th>
<th>AGB statistics (Mg ha⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Mode</td>
</tr>
<tr>
<td>Boreal forest</td>
<td>399,019 (100%)</td>
<td>1.2 × 10⁹ (100%)</td>
<td>39.5</td>
</tr>
<tr>
<td>Deciduous forest</td>
<td>32,593 (8.2%)</td>
<td>1.5 × 10⁸ (12.6%)</td>
<td>44.6</td>
</tr>
<tr>
<td>Evergreen forest</td>
<td>115,490 (28.9%)</td>
<td>5.3 × 10⁸ (46.2%)</td>
<td>46.0</td>
</tr>
<tr>
<td>Mixed forest</td>
<td>26,107 (6.5%)</td>
<td>1.4 × 10⁸ (11.7%)</td>
<td>51.6</td>
</tr>
<tr>
<td>Shrubland</td>
<td>120,888 (30.3%)</td>
<td>1.4 × 10⁸ (12.5%)</td>
<td>11.9</td>
</tr>
<tr>
<td>Woody wetlands</td>
<td>29,990 (7.5%)</td>
<td>6.7 × 10⁷ (5.8%)</td>
<td>22.6</td>
</tr>
<tr>
<td>Burned area (1990–2009)</td>
<td>73,948 (18.5%)</td>
<td>1.3 × 10⁸ (11.2%)</td>
<td>23.8</td>
</tr>
</tbody>
</table>
fire burns, and post-fire succession communities, would be captured more appropriately with high- or intermediate-resolution images. Accurately mapping the heterogeneous land cover at a local scale is a better approach for quantifying population features in the entire study area than coarse-resolution maps. Comparing the 30 m-resolution AGB map with the earlier 250 m AGB map, the former shows significant improvement in detecting the spatial details, especially the distinct biomass breaks associated with fires. Capturing the spatial heterogeneity in the complicated landscape provides a more accurate and reliable estimation for the regional population than the coarse-resolution data.

5. Conclusions
We mapped the AGB at a 30 m resolution for the boreal forest in the Yukon River Basin of Alaska, which included deciduous forest, evergreen forest, mixed forest, shrublands, and woody wetlands. We used a regression method to estimate regional AGB based on the Landsat at-sensor reflectance and spectral indices and the field measurements including tree, shrub, and CWD biomass. The digital AGB data product is available for download from the USGS EROS Land Cover Applications and Global Change Website (http://lca.usgs.gov/lca/). An accuracy assessment of the AGB map indicated that the MAE and MBE were 25.7 and 4.3 Mg ha$^{-1}$, respectively. The AGB estimates also correlated fairly well with the vegetation heights derived from airborne lidar data sets acquired in the central basin. The regional AGB estimates show that the mean AGB value for the boreal forest is 39.5 Mg ha$^{-1}$, and 90% of the study area has AGB values between 2.2 and 19.0 Mg ha$^{-1}$. The total AGB estimated for the boreal forest portion of the region is $1.15 \times 10^9$ Mg. The spatial AGB patterns in the basin are controlled primarily by land-cover types and disturbances related to historical fires.

Figure 5. Frequency chart of the AGB estimates for the boreal forest and other vegetation types in the Yukon River Basin, Alaska.
Acknowledgements
We thank Mark Winterstein (University of Fairbanks) and Kurtis Nelson (USGS EROS Center) for assistance with field data collection and Jason Stoker (USGS) for processing the lidar data. We also thank Eugene Fosnight for reviewing the manuscript and providing valuable comments.

Disclosure statement
Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government.

Funding
This study was funded by the US Geological Survey Climate Effects Network, Land Change Science, LandCarbon, and Global Change Research & Development Programmes. The work by L. Ji and B. Peterson was performed under USGS contract G08PC91508. The work by N. Pastick was performed under USGS contract G10PC00044.

References


