Distinguishing between live and dead standing tree biomass on the North Rim of Grand Canyon National Park, USA using small-footprint lidar data

Yunsuk Kim a, Zhiqiang Yang b, Warren B. Cohen b,*, Dirk Pflugmacher a, Chris L. Lauver c, John L. Vankat d

Abstract

Accurate estimation of live and dead biomass in forested ecosystems is important for studies of carbon dynamics, biodiversity, wildfire behavior, and for forest management. Lidar remote sensing has been used successfully to estimate live biomass, but studies focusing on dead biomass are rare. We used lidar data, in conjunction with field measurements from 58 plots to distinguish between and map standing live and dead tree biomass in the mixed coniferous forest of the North Rim of Grand Canyon National Park, USA. Udar intensity and canopy volume were key variables for estimating live biomass, whereas for dead biomass, lidar intensity alone was critical for accurate estimation. Regression estimates of both live and dead biomass ranged between 0 and 600 Mg ha⁻¹ with means of 195.08 Mg ha⁻¹ and 65.73 Mg ha⁻¹ respectively. Cross validation with field data resulted in correlation coefficients for predicted vs. observed of 0.85 for live biomass (RMSE=50 Mg ha⁻¹ and %RMSE (RMSE as a percent of the mean) =26). For dead biomass, correlation was 0.79, RMSE was 42 Mg ha⁻¹, and %RMSE was 63. Biomass maps revealed interesting patterns of live and dead standing tree biomass. Live biomass was highest in the ponderosa pine zone, and decreased from south to north through the mixed conifer and spruce-fir forest zones. Dead biomass exhibited a background range of values in these mature forests from zero to 100 Mg ha⁻¹ with lower values in locations having higher live biomass. In areas with high dead biomass values, live biomass was near zero. These areas were associated with recent wildfires, as indicated by fire maps derived from the Monitoring Trends in Burn Severity Project (MTBS). Combining our dead biomass maps with the MTBS map, we demonstrated the complementary power of these two datasets, revealing that MTBS burn intensity class can be described quantitatively in terms of dead biomass. Assuming a background range of dead biomass up to 100 Mg ha⁻¹, it is possible to estimate and map the contribution to the standing dead tree biomass pool associated with recent wildfire.

1. Introduction

Accurate estimation of biomass in forested ecosystems is important for global studies of forest monitoring protocols for Grand Canyon National Park (GCNP), Arizona, USA. The U.S. National Park Service is developing a monitoring program (http://science.nature.nps.gov/im/index.cfm):

"...to provide the minimum infrastructure needed to track the overall condition of natural resources in parks and to provide early warning of situations that require intervention. The scientifically sound information obtained through this systems-based monitoring program will have multiple applications for management decision-making, park planning, research, education, and promoting public understanding of park resources."

Monitoring of forest structure in GCNP is important for quantification and understanding of ongoing changes. Current changes were
partially initiated in approximately 1880, when exclusion of naturally occurring, low intensity surface fires began, leading to decreased mortality of small trees and subsequently increased forest densities, especially involving fire-intolerant and shade-tolerant species (Crocker-Bedford et al., 2005; Fule et al., 2002b, 2003; Mast & Wolf 2004; White & Vankat 1993). Beginning in 1980, forest structure has been altered by management fires (prescribed and wildland fire-use fires) used to reintroduce fire into these forests, as well as by wildfires (National Park Service 2008). Moreover, there is evidence that unburned forests are also changing in structure, as a result of increased tree mortality likely caused by the interaction of competition and insect outbreaks (Vankat et al., 2005) and possibly linked to climate change. Similar changes in forest structure have occurred widely in the western USA (van Mantgem et al., 2009), and effective, efficient monitoring is needed to track and understand these changes.

Aboveground biomass (AGB) has been estimated successfully with remote sensing, especially using lidar data (e.g., Bortolot & Wynne 2005; Lim et al., 2003; Lim and Treitz, 2004; Naesset 2004; Nelson et al., 1988, 2005; Popescu 2007; Popescu et al., 2003, 2004; Sherrill et al., 2008; Van Aardt et al., 2006). Although large footprint waveform lidar has been used to estimate biomass (Drake et al., 2002; Hyde et al., 2005, 2006; Lesčky et al., 1999; Pfugmacher et al., 2008), most studies used discrete return data, as these were more commonly available.

Both plot-based and tree segmentation approaches have been used to estimate biomass from small-footprint discrete lidar. The plot-based approach commonly involves field-measured biomass regressed against derived statistics from plot-level lidar data. The lidar statistics can be from the individual returns or from a canopy height model where lidar return values are interpolated to a certain size raster (e.g., Hyde et al., 2007; Lim et al., 2003; Lim & Treitz, 2004; Neesset, 2004). Several different tree segmentation approaches have been used, such as application of allometric equations to individual trees identified in the lidar dataset (Popescu, 2007).

Explicit estimation of dead biomass has received minimal attention. Sherrill et al. (2008) examined the variables derived from canonical correlation analysis as well as the conventional lidar height variables to see how well these variables correlated with various field measurements, one of which was standing dead tree biomass. Lidar mean and maximum heights were the two variables that showed the highest correlation with standing dead biomass. Bater et al. (2007) estimated the density proportion of dead trees in coastal forests of Vancouver Island, British Columbia, Canada. In their study, the mean extracted from a log-normal distribution of wildlife tree class (or decay class) was highly correlated with the log-transformed coefficient of variation of lidar height.

Lidar intensity values are increasingly available. Intensity is the ratio of the power returned to the power emitted and is mainly a function of surface reflectivity at the emitted wavelength (Kaasalainen et al., 2007). It is also a function of the area of the object that returns the pulse, and the proportion of the pulse remaining after previous returns (Brandtberg, 2007). Intensity data are generally not calibrated for differences in receiver gains that are periodically adjusted during acquisition. Gain settings are currently proprietary, and thus are not made available to the end user (Boyd & Hill, 2007; Donoghue et al., 2007; Kaasalainen et al., 2007).

Although others have not explicitly focused on using lidar intensity to estimate dead biomass, they have nonetheless relied on the knowledge that foliage exhibits a higher near-infrared (NIR) return intensity than non-foliage vegetation components to estimate live biomass and related variables. Lim et al. (2003) used an intensity threshold to remove lower NIR intensity returns when estimating live biomass of a northern hardwood forest in Ontario, Canada. In that study, the mean height of the higher intensity returns was the best predictor of basal area, biomass and volume ($R^2 > 0.85$). Hudak et al. (2006) estimated basal area and density of a managed mixed forest in Idaho, USA using lidar height and intensity in combination with spectral bands of the Advanced Land Imager. Donoghue et al. (2007) used intensity data to estimate the proportion of two conifer species in Scotland. Brandtberg et al. (2003) used intensity to identify tree species within a deciduous forest in West Virginia, USA.

The above studies indicate that lidar intensity values have great potential for characterizing forest structure and live biomass, but no studies have explicitly focused on use of lidar intensity for directly estimating and mapping dead biomass. The primary objectives of our study were to: (1) develop regression equations to predict live and dead standing tree biomass from discrete return lidar data; (2) examine the value of lidar intensity for distinguishing between live and dead biomass; and (3) apply derived regression equations to map live and dead biomass across a portion of the North Rim of Grand Canyon National Park.

2. Materials and methods

2.1. Study area

The study area was located in the central portion of the North Rim region of GCNP in northern Arizona, USA (Fig. 1). Elevations range from 1963 to 2797 m from south to north. The lower elevations have ponderosa pine (Pinus ponderosa) forest. Mid-elevations have mixed conifer forest dominated by various combinations of ponderosa pine, Douglas-fir (Pseudotsuga menziesii), white fir (Abies concolor) and quaking aspen (Populus tremuloides). Higher elevations and north aspects of mid-elevations have spruce-fir forest dominated by Engelmann spruce (Picea engelmannii), blue spruce (Picea pungens) and subalpine fir (Abies lasiocarpa).

2.2. Field data

A total of 58 rectangular field plots (20 m x 50 m) within the study area were surveyed (Fig. 1). Twenty-five plots in the south half of the study area were a subset of 456 plots originally sampled as part of a vegetation mapping project in 1935. The locations had been "determined by the good judgment of the mapper" such that "they represent average or typical conditions within the [vegetation] type" (Coffman, 1934). Thirty-three plots in the northern half were a subset of 40 plots originally established in 1984 in clusters to represent the variation in elevation and topography within Thompson Canyon watershed (White & Vankat, 1993). All 58 plots had been surveyed in 2004 and 2005, including recording of plot compass bearings and the species. Diameter at breast height (dbh), live/dead status, and location of standing trees > 10 cm dbh.

In July 2007, the 58 plots used in this study were revisited to collect additional data. Differentially corrected GPS readings were taken at the plot centers, tree heights (m) were measured using an Opti-Logic 400LH hypsometer, and the live/dead status of individual trees was updated. All live trees > 10 cm dbh in the prior surveys were determined to be > 3 m height. The measurements used to represent the variation in elevation and topography within Thompson Canyon watershed, about one-third were dead (Table 1). Four of the 58 plots had been burned in recent crown fires, and all recorded trees were dead. Three of these plots on the east were from the 2000 Outlet fire, a large-scale crown fire that started as a prescribed burn. The dead plot in the south was part of the 2005 Dragon fire in which only spotty crowning occurred. In addition, paper tree location maps were created for all the field plots.

2.3. Biomass estimation

No local tree biomass equations were available for the North Rim. Therefore the generalized national allometric equation of Jenkins et al. (2003) was used to calculate tree biomass. The Jenkins equation (Eq.
The foliage component was subtracted from the remaining trees using Eq. (2), the component biomass equation of Jenkins et al., (2003). This equation represents the average proportion (component ratio) of AGB in foliage, stem bark, stem wood and coarse roots as a function of dbh:

\[ \text{Ratio} = \exp(\alpha_0 + (\alpha_1 / \text{dbh})) \]

where

\[ \alpha_0 \text{ and } \alpha_1 \] are parameters specific to each tree component and species group. For the foliage component, \( \alpha_0 \) is -4.08 and \( \alpha_1 \) is 5.9 for hardwood species and -2.96 and 4.48 for softwood species, respectively. Subtraction of foliage biomass reduced total dead tree AGB by 2-8%, depending on species. Third, dead trees that likely had broken tops were identified and their biomass adjusted downward. This was accomplished using the frequency distribution of the ratio of live tree height (m) to dbh (cm) - the HDR (Fig. 2). The HDR threshold used was 29.5, which was two standard deviations \((2\sigma)\) from the mean HDR on the lower tail of the distribution for live trees (Fig. 2, top). Our

---

**Table 1**

Summary of standing trees ≥10 cm dbh in the 58 plots.

<table>
<thead>
<tr>
<th>Scientific name</th>
<th>Condition</th>
<th>dbh (cm)</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abies concolor</td>
<td>Live</td>
<td>20.0</td>
<td>91.7</td>
</tr>
<tr>
<td>Abies lasiocarpa</td>
<td>Dead</td>
<td>22.4</td>
<td>19.1</td>
</tr>
<tr>
<td>Picea spp.</td>
<td>Live</td>
<td>18.5</td>
<td>9.9</td>
</tr>
<tr>
<td></td>
<td>Dead</td>
<td>20.4</td>
<td>8.5</td>
</tr>
<tr>
<td>Pinus ponderosa</td>
<td>Live</td>
<td>22.5</td>
<td>13.6</td>
</tr>
<tr>
<td></td>
<td>Dead</td>
<td>19.5</td>
<td>15.8</td>
</tr>
<tr>
<td>Populus tremuloides</td>
<td>Live</td>
<td>35.1</td>
<td>22.2</td>
</tr>
<tr>
<td></td>
<td>Dead</td>
<td>37.1</td>
<td>24.3</td>
</tr>
<tr>
<td>Pseudotsuga menziesii</td>
<td>Live</td>
<td>25.6</td>
<td>18.6</td>
</tr>
<tr>
<td></td>
<td>Dead</td>
<td>19.2</td>
<td>20.2</td>
</tr>
</tbody>
</table>

Total 2950

---

**Table 2**

Parameter values used for the biomass equation.

<table>
<thead>
<tr>
<th>Species name</th>
<th>( \beta_0 )</th>
<th>( \beta_1 )</th>
<th>( r^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abies concolor</td>
<td>-2.5384</td>
<td>2.4814</td>
<td>0.992</td>
</tr>
<tr>
<td>Abies lasiocarpa</td>
<td>-2.5384</td>
<td>2.4814</td>
<td>0.992</td>
</tr>
<tr>
<td>Picea spp.</td>
<td>-2.0773</td>
<td>2.3323</td>
<td>0.988</td>
</tr>
<tr>
<td>Pinus ponderosa</td>
<td>-2.5356</td>
<td>2.4349</td>
<td>0.987</td>
</tr>
<tr>
<td>Populus tremuloides</td>
<td>-2.2094</td>
<td>2.3867</td>
<td>0.953</td>
</tr>
<tr>
<td>Pseudotsuga menziesii</td>
<td>-2.2304</td>
<td>2.4435</td>
<td>0.992</td>
</tr>
</tbody>
</table>

\( r^2 \) is the model statistic associated with the number of data points generated from published equations for parameter estimation (Jenkins et al., 2003).
assumption was that outliers in live tree distribution were due to errors at the time of recording or were from live trees with broken tops. For the dead trees with HDR < 29.5 (Fig. 2, bottom), biomass was adjusted downward. Downward adjustment was based on calculation of an imaginary full height at the minimum HDR value of 29.5, using the measured dbh. For example, a tree with 60 cm dbh should have at least a 17.7 m height (60 cm x 29.5 = 17.7 m). A dead tree with 60 cm dbh and height of 5.9 m would be, therefore, one-third expected minimum height, and its biomass was reduced by two-thirds using the Jenkins equation, which is based on dbh alone. This calculation did not take tree taper into consideration directly. Rather, we assumed that using the HDR at the lower 20% would account for taper. The number of dead trees adjusted was 162 of 930 above 3 m height.

2.4. Lidar data

Our lidar data were acquired in July 2007 using a Leica ALS50 Phase II laser system mounted in a Cessna Caravan 208, covering the study area of 10.714 ha (Table 3). Raw data were obtained in lidar data exchange (LAS) format version 1.1, including ground classified points and all multiple returns with intensity values in 1064 nm laser energy rescaled to 8-bit (0-255). The vendor provided the digital terrain model with the cell size of 1 m.

The raw elevation data were normalized to tree height values using Fusion software developed by the USDA Forest Service for forest/vegetation applications (McGaughey, 2007). Returns below 3 m were considered non-tree canopy (i.e., the shortest recorded live tree in the field, as described earlier), and a threshold was applied to remove them from the canopy dataset.

2.4.1. Geographic registration of field and lidar data

A difference between the datum from the GPS recorded in the field (WGS84) and the lidar data (NAD83), GPS error, and possible imprecision in the field-determined plot bearings, resulted in misregistration of field plots with respect to the lidar dataset. To register the two datasets more accurately, the plots were manually relocated. First, for each plot, the plot polygon corner coordinates (created from plot center coordinates, and plot size, shape, and orientation) were calculated. Using these corner coordinates as control points, the paper tree map for a given plot was rectified into the lidar geographic space (Fig. 3). The locations of all trees on the tree map were then used to help find the true location of the plot in the lidar point cloud, as follows: individual trees with each plot were identified visually using color ramps for both intensity and height values (ArcMap and Fusion 3D viewer), and the plot polygon was then shifted and rotated so that the tree locations within the map properly aligned with the point cloud. Because the GCNP forests had a relatively open canopy (7-73% cover, with a mean of 50%) and many dead trees, we were able to identify most of the field-recorded trees in the lidar point cloud. Of the 58 plots, 54 were relocated and 4 were left unchanged.

2.4.2. Lidar-derived explanatory variables

Various explanatory variables (Table 4) were calculated from the first returns of each plot to estimate plot-level biomass. Only first returns were used because preliminary analyses suggested that, for North Rim forests, variables derived from first returns produced better biomass models than those including other returns (e.g., R² values for total biomass were 0.10 higher with first returns only). This was probably because the relatively open canopy forest allowed good penetration of high-density first returns to the forest floor, permitting them to be representative of vertical structure (Thomas et al., 2006).

The variables HtMax, HtMean, HtMed, HtMode, Htvar, and the percentile heights (Ht10 through Ht90) were calculated from all first returns above the 3 m threshold. Cover (Cov) was calculated as the ratio of the number of first returns above 3 m to the total number of first returns, multiplied by 100. To obtain plot-level volume (Vol), we multiplied cover by mean height.

The peak frequency variables (LoIntPk, HiIntPk) were the relative frequency values of the low and high peaks of the intensity density distribution for each plot (Fig. 4). The density distribution of lidar intensity values for each plot was developed with "density" function using a Gaussian kernel in S-Plus v8 (Insightful, 2007). The intensity density distributions for most plots had two distinct peaks, one in the low intensity range (near 20) and one in the high range (near 200). The data for the 6 field plots in Fig. 4 are representative of plots having different proportions of live and dead trees, which ranged from 100% live trees to 0% dead. We hypothesize that lower and higher NIR
specifically, the regression analyses were conducted using: 1) variables calculated from lower intensity first returns (L) for dead biomass, 2) variables calculated from high intensity returns (H) for live biomass, and variables calculated from all first returns across the full intensity range (All) separately for 3) dead and 4) live biomass. No transformations were applied to the data sets.

In addition to stepwise regression models, we explored the possibility of identifying alternative predictor sets that might serve as a standard set of discrete return lidar-derived variables to be used across forest types (Lefsky et al., 2002; Pfugmacher et al., 2008).

Alternative models were defined based on our understanding of relationships between biomass and variables derived from lidar data, e.g. volume, or height for live and dead biomass. We hypothesized that these simplified models are less likely tuned to our specific intensity lidar returns are related to the dead and live materials, respectively. LoIntPK and HiIntPK were two variables used to represent the dead and live tree components, respectively. Of all 58 plots, only one plot, with all dead trees and no foliage above the 3 m threshold (plot 214 in Fig. 4), lacked a peak in the high intensity region of the density plot (above 125). For this plot, relative frequency of intensity 200 was used.

### 2.5. Biomass modeling

To investigate relationships of lower intensity returns with dead biomass and higher intensity returns with live biomass, the lidar data were stratified at intensity value 125 (see Fig. 4). All variables in Table 4, except for LoIntPK and HiIntPK (i.e., height variables, cover, and canopy volume), were calculated using the two intensity strata independently and together for four separate stepwise regression analyses. More specifically, the regression analyses were conducted using: 1) variables calculated from lower intensity first returns (L) for dead biomass, 2) variables calculated from high intensity returns (H) for live biomass, and variables calculated from all first returns across the full intensity range (All) separately for 3) dead and 4) live biomass. No transformations were applied to the data sets.

In addition to stepwise regression models, we explored the possibility of identifying alternative predictor sets that might serve as a standard set of discrete return lidar-derived variables to be used across forest types (Lefsky et al., 2002; Pfugmacher et al., 2008). Alternative models were defined based on our understanding of relationships between biomass and variables derived from lidar data, e.g. volume, or height for live and dead biomass. We hypothesized that these simplified models are less likely tuned to our specific intensity lidar returns are related to the dead and live materials, respectively. LoIntPK and HiIntPK were two variables used to represent the dead and live tree components, respectively. Of all 58 plots, only one plot, with all dead trees and no foliage above the 3 m threshold (plot 214 in Fig. 4), lacked a peak in the high intensity region of the density plot (above 125). For this plot, relative frequency of intensity 200 was used.

### 2.5. Biomass modeling

To investigate relationships of lower intensity returns with dead biomass and higher intensity returns with live biomass, the lidar data were stratified at intensity value 125 (see Fig. 4). All variables in Table 4, except for LoIntPK and HiIntPK (i.e., height variables, cover, and canopy volume), were calculated using the two intensity strata independently and together for four separate stepwise regression analyses. More specifically, the regression analyses were conducted using: 1) variables calculated from lower intensity first returns (L) for dead biomass, 2) variables calculated from high intensity returns (H) for live biomass, and variables calculated from all first returns across the full intensity range (All) separately for 3) dead and 4) live biomass. No transformations were applied to the data sets.

In addition to stepwise regression models, we explored the possibility of identifying alternative predictor sets that might serve as a standard set of discrete return lidar-derived variables to be used across forest types (Lefsky et al., 2002; Pfugmacher et al., 2008). Alternative models were defined based on our understanding of relationships between biomass and variables derived from lidar data, e.g. volume, or height for live and dead biomass. We hypothesized that these simplified models are less likely tuned to our specific intensity lidar returns are related to the dead and live materials, respectively. LoIntPK and HiIntPK were two variables used to represent the dead and live tree components, respectively. Of all 58 plots, only one plot, with all dead trees and no foliage above the 3 m threshold (plot 214 in Fig. 4), lacked a peak in the high intensity region of the density plot (above 125). For this plot, relative frequency of intensity 200 was used.

### 2.5. Biomass modeling

To investigate relationships of lower intensity returns with dead biomass and higher intensity returns with live biomass, the lidar data were stratified at intensity value 125 (see Fig. 4). All variables in Table 4, except for LoIntPK and HiIntPK (i.e., height variables, cover, and canopy volume), were calculated using the two intensity strata independently and together for four separate stepwise regression analyses. More specifically, the regression analyses were conducted using: 1) variables calculated from lower intensity first returns (L) for dead biomass, 2) variables calculated from high intensity returns (H) for live biomass, and variables calculated from all first returns across the full intensity range (All) separately for 3) dead and 4) live biomass. No transformations were applied to the data sets.

In addition to stepwise regression models, we explored the possibility of identifying alternative predictor sets that might serve as a standard set of discrete return lidar-derived variables to be used across forest types (Lefsky et al., 2002; Pfugmacher et al., 2008). Alternative models were defined based on our understanding of relationships between biomass and variables derived from lidar data, e.g. volume, or height for live and dead biomass. We hypothesized that these simplified models are less likely tuned to our specific
datasets, and are more readily applicable in other forest systems. Initially, our logic followed one of our main objectives: to explore the utility of intensity for estimating biomass. Assuming that lower and higher NIR intensity returns are related to the dead and live biomass respectively, we conceived of a general model that utilized the intensity data to separately estimate live and dead biomass. However, we also selected predictors from all returns across the intensity range to predict both live and dead biomass, which if successful would eliminate the need to stratify the returns by intensity.

Four models, the two best from stepwise regression and the two best from our alternative models, were selected for error assessment using leave-one-out cross validation - two for live and two for dead biomass. Correlation between predicted and observed, root mean square error (RMSE), and bias values were calculated from the 58 independent predictions produced by cross validation.

Using the best models for live and dead biomass, lidar-predicted biomass maps were produced. A 30-by-30 m grid resolution was adopted to approximate the area of the 50 by 20 m sample plots. Cells with predicted negative biomass values were truncated to zero biomass. A few cells having biomass predictions larger than 600 Mg ha$^{-1}$ (up to 630 Mg ha$^{-1}$) were truncated at 600 Mg ha$^{-1}$.

3. Results

We report nine of the regression models we developed, the best for live and dead biomass from stepwise regression and seven from the alternative model set (AMS) (Table 5). For live biomass, the stepwise regression model (model 1) included four predictor variables from the high intensity stratum. Of these, only volume and 20th percentile height appear to have any practical significance in terms of their contributions to the coefficient of determination. Alternative live biomass models included the volume of high intensity returns (model 2), cover and mean height (the two variables from which volume was calculated) from high intensity returns (model 3), and volume of all returns plus peak high intensity count (HilnPk) (model 4). All three of these alternative models were similar in terms of predictive power ($R^2$, RMSE and %RMSE). The stepwise model, not unexpectedly, had a higher $R^2$ (0.76) and lower RMSE (46 Mg ha$^{-1}$) than any of the other three; but the difference in predictive power of the stepwise model was small (approximately 4 Mg ha$^{-1}$ lower RMSE).

Comparing models 2 and 3 reveals that cover and mean height interactions are somewhat important, given that the model based solely on volume (cover times mean height) had a higher $R^2$ and lower errors than the model based on these two separate variables. In relation to model 2, model 4 illustrates two important findings: (1) volume of high intensity returns is more strongly predictive of live biomass than is volume across all intensity values, and (2) if volume from the non-stratified dataset is used, the peak count of high intensity returns is required for more accurate prediction.

The stepwise model for dead biomass (model 5) included four significant variables, the most important of which was the low intensity peak count. This model was not as strong as the ones for live biomass ($R^2 = 0.62$, %RMSE = 56, although it does have a lower absolute RMSE). The alternative dead biomass models included volume of low intensity returns (model 6), cover and mean height of low intensity returns (model 7), peak low intensity count (model 8), and LoIntPk plus volume from all returns (model 9). Highlighting the importance of low intensity peak count for predicting dead biomass are the poor predictive powers of models 6 and 7, which do not include the LoIntPk variable. Moreover, when the LoIntPk variable is used by itself (model 8), it has nearly the predictive strength of the model that includes LoIntPk plus volume of all returns (model 9).

Cross validation was performed on selected models (Table 6). The two stepwise models, as might be expected, performed slightly better than the selected alternative models, with higher correlation coefficients for predicted vs. observed, as well as lower RMSEs (Table 5). Bias was negligible in all cases. For the alternative models, we chose to examine the single variable models (models 2 and 8). These models performed quite well in comparison to the stepwise models, as shown by the validation statistics (Table 6) and the scatterplots of predicted vs. observed (Fig. 5).

The models produced by stepwise regression analysis were applied to the entire study area to create live and dead biomass maps (Fig. 6). Larger amounts of live biomass were observed in the ponderosa pine forest zone in the south, with live biomass gradually decreasing toward the spruce-fir forest zone in the north. Dead biomass was concentrated inside the two wildfire areas. In general, the spatial distributions of live and dead biomass maps were complementary, with areas of high live biomass having low amounts of dead biomass and visa-versa.

4. Discussion

4.1. Lidar intensity

We explored the relationship between scaled lidar intensity values and plot-level standing tree biomass in the conifer-dominated forests of the North Rim of GCNP. Results indicated that lidar intensity data can be used to estimate and distinguish between standing live and

<table>
<thead>
<tr>
<th>Response</th>
<th>Parameters</th>
<th>$\beta$</th>
<th>Partial $R^2$</th>
<th>$R^2$</th>
<th>RMSE (Mg ha$^{-1}$)</th>
<th>RMSE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live biomass</td>
<td>Intercept</td>
<td>-23.25</td>
<td>0.76</td>
<td>46.01</td>
<td>23.66</td>
<td></td>
</tr>
<tr>
<td>Stepwise</td>
<td>Vol1</td>
<td>40.87</td>
<td>0.722</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ht20H</td>
<td>-20.92</td>
<td>0.040</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HtMeanI</td>
<td>21.99</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HtMaxI</td>
<td>-2.44</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMS</td>
<td>Intercept</td>
<td>20.26</td>
<td>0.72</td>
<td>49.84</td>
<td>25.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VolI</td>
<td>38.56</td>
<td>0.722</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ht20I</td>
<td>-104.54</td>
<td>0.70</td>
<td>51.50</td>
<td>26.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HtMeanI</td>
<td>6.99</td>
<td>0.567</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HtMaxI</td>
<td>7.05</td>
<td>0.136</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMS</td>
<td>Intercept</td>
<td>-86.31</td>
<td>0.73</td>
<td>49.52</td>
<td>25.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VolI</td>
<td>23.23</td>
<td>0.654</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LoIntPk</td>
<td>1.27</td>
<td>0.071</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dead biomass</td>
<td>Intercept</td>
<td>-112.97</td>
<td>0.62</td>
<td>37.09</td>
<td>56.31</td>
<td></td>
</tr>
<tr>
<td>Stepwise</td>
<td>LntPk</td>
<td>3.58</td>
<td>0.519</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HtMedI</td>
<td>3.62</td>
<td>0.058</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CovI</td>
<td>1.09</td>
<td>0.025</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CovI</td>
<td>0.48</td>
<td>0.010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMS</td>
<td>Intercept</td>
<td>39.19</td>
<td>0.03</td>
<td>59.08</td>
<td>89.68</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VolI</td>
<td>6.91</td>
<td>0.030</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ht20I</td>
<td>24.79</td>
<td>0.02</td>
<td>59.28</td>
<td>89.99</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HtMeanI</td>
<td>1.91</td>
<td>0.014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CovI</td>
<td>0.48</td>
<td>0.010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMS</td>
<td>Intercept</td>
<td>-93.79</td>
<td>0.52</td>
<td>41.63</td>
<td>63.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LntPk</td>
<td>0.52</td>
<td>0.519</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LntPk</td>
<td>3.34</td>
<td>0.519</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VolI</td>
<td>3.03</td>
<td>0.024</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Subscripts: "H" indicates that the model was developed using only the high intensity stratum, "I" indicates the low intensity stratum was used, and "All" indicates the model was based on the non-stratified dataset.
dead biomass above 3 m height in these forests, which contain a relatively high percentage of dead trees.

Among the 58 field plots, almost all of the normalized lidar intensity histograms exhibited a fairly strong bimodal distribution (as in Fig. 4). The lower and upper peaks of these histograms were likely associated with woody and foliage components, respectively. More precisely, given that dead trees lack foliage and live trees consist of foliage and woody components, the low and high intensity peaks should be related to the amount of dead and live standing tree biomass. This might largely explain the importance of intensity in modeling dead and live biomass. That these peaks occurred at nearly the same intensity values in all but one plot suggests that the rescaling, associated with any gain setting adjustments during data acquisition, did not diminish the utility of the intensity data for this purpose.

Our use of intensity data in this study was based on relative frequency distributions of intensity. However, there is more information in the intensity distributions than we exploited. For example, as illustrated in Fig. 7, one can more directly exploit the height distribution of intensity values. Thereby, a third dimension could be added to the canopy volume profile and method of Lefsky et al. (1999), which is based on waveform data, but has been adapted to discrete return lidar by Coops et al. (2007). If the intensity dimension were added to this method, one would have a powerful means of describing distributions of foliage and woody material as a function of height. This would be useful in monitoring and modeling of wildlife habitat, fire fuels, and biogeochemical cycling.

4.2. Live and dead biomass estimation

Our finding that the strongest regression model for live biomass used returns only from the high intensity stratum of the lidar dataset is consistent with Lim et al. (2003), who also found that by using only high intensity returns they could improve the estimation of live hardwood tree biomass in Ontario. Because volume of high intensity returns (model 2, Tables 4 and 5) was nearly identical in predictive power to volume plus three other predictor variables (model 1), we may have identified a single, general predictor for live standing tree biomass from discrete return lidar, This needs to be tested in other forest ecosystems.

In the study of Lim et al. (2003), low intensity returns were described as being associated with echoes from the terrain. In our study, where a 3 m threshold was applied, the low intensity returns were associated with non-foliage components of trees. This enabled us to take advantage of low intensity returns to estimate standing dead tree biomass. Unlike for live biomass, where volume (or cover and mean height) was a critical descriptor, volume (models 6 and 7) was not well correlated to dead biomass. Rather, the critical variable was the low intensity peak count of the intensity frequency distribution. One potential explanation of this is that standing dead biomass is porous to lidar pulses. Why cover would be problematic in this circumstance is obvious. For height, we suspect that high variability in height distributions of live woody material independent of dead biomass amount was confounding. We cannot fully explain why the actual count of the low intensity returns was strongly associated with the amount of dead biomass, but it is clearly quite sensitive to the proportion of echoes from woody material, including dead trees, which would increase with a decrease in foliage. This also needs to be tested in other forest ecosystems.

Intensity stratification for modeling live and dead standing tree biomass may be unnecessary. Our models 4 and 9 were based on unstratified data. Even though we selected models 1 and 2 for live biomass, both based on stratified datasets, model 4 (using unstratified data) appeared to be as good a model as models 1 and 2 in terms of predictive strength and errors. Model 5, used to map dead biomass, and model 9, our best alternative model for dead biomass, were both based on non-stratified data.

Biomass calculations from the field data undoubtedly contributed to error in our models. This is especially true for dead biomass, given we did not accurately measure snag characteristics in the field (Harmon et al., 2004; Harmon & Sexton, 1996). However, only marginal shifts in biomass should be expected, which would not change the interpretation of our results.

We were surprised that low intensity volume was not important for estimating dead biomass (as its counterpart, high intensity volume, was for live biomass). Using waveform data of sufficient footprint size (e.g. 10 m), however, we expect that low intensity volume would be important, due to broad photon saturation of the footprint which would likely sample the standing dead wood more effectively than small footprint discrete lidar.

4.3. Utility of live and dead biomass maps

At GCNP, live and dead biomass maps have several potential uses. For example, they could reveal areas of non-forest vegetation such as grasslands and meadows identified as the lack of live and dead trees, as in the southwestern corner of the mixed forest zone (Fig. 6). Monitoring of such areas could identify woody encroachment into non-forested areas. This scenario may occur with climate change, and may cause a decrease in the spatial extent of grasslands and meadows. Standing dead tree biomass (related to snag size and density) in relation to live tree biomass is a key characteristic describing maturity in conifer-dominated forests (Spies & Franklin, 1988), and can be used to assess quality and
Fig. 6. Live (left) and dead (middle) standing tree biomass maps, and a map of dead as a proportion of total biomass (right). Forest cover type was simplified from a National Park Service map, and fire polygons are from the Monitoring Trends in Burn Severity (MTBS) project (www.mtbs.gov). The zero live biomass areas include meadows without trees, or areas with high dead biomass. The southeast corner (without data) is the canyon where no forests exist.
Fig. 7. Intensity data for four plots (left, frequency distributions as a function of height; right, frequency distributions independent of height). Each pair of plots (top, bottom) has similar relative frequency distributions (right), but different distributions as a function of tree height (left).
diversity of habitats for snag-dependent bird species (Lundquist & Mariani, 1991) and a variety of vertebrate species (Raphael, 1991). Snags are also a source of coarse woody debris, which is an important component in forest succession (Harmon et al., 1986). Knowledge about such things as snag density helps make predictions about future site conditions. Given that the dead biomass map depicts the spatial distributions of standing dead trees, this map could be used in investigations of the cause of tree mortality from a variety of factors such as disease, insects, drought, and wildfire. It is also possible to evaluate the effect of spatial patterning of dead trees and their associated causes. Scattered, sparse dead trees indicate a form of minor disturbance, whereas large patches of dead trees indicate more widespread disturbance factors. Spatial pattern information could be useful in developing site-specific forest and fire management plans.

With respect to wildfire, the fact that dead and live biomass appear to be inversely related (note in Fig. 6 where there is exceptionally high dead biomass within the two fire polygons there is a near absence of live biomass) suggests that the proportion of dead biomass relative to total biomass (Fig. 6 right) could accurately reflect fire severity, where known fires have occurred. For example, the Outlet fire appears to have burned more severely, with more widespread proportional crown consumption and tree death than the Dragon fire, which experienced a patchier high intensity burn environment. We can also see that both fires had extensive areas of lesser proportional crown consumption, and thus fire severity. The proportion map shows similar patterns as the dead biomass map; however, an advantage of the proportion map is that the amount of dead biomass is normalized throughout the study area and thus independent of the initial condition.

A potentially valuable use of the dead (or dead as a proportion of total) biomass map is to use it in conjunction with maps of burn severity created by the Monitoring Trends in Burn Severity (MTBS) project (www.mtbs.gov). The MTBS maps use a time series of Landsat data and a spectral index called the normalized burn ratio (NBR), which is calculated using the Landsat bands 4 and 7 in a normalized difference ratio to locate recent fires and estimate fire severity (Key & Benson, 2005; Kokaly et al., 2007). The dNBR, or difference in NBR, is calculated as the difference of prefire and postfire images (Key et al., 2002; Kokaly et al., 2007), but is only an index of burn severity. The severity classes are arbitrarily defined by MTBS analysts and capture the range of severities in 4 classes, for a given forested system. Using our dead biomass map, we describe the amount of dead biomass in each of four MTBS burn severity classes (Fig. 8). The unburned areas inside the MTBS polygon have values consistent with the background value of <100 Mg ha\(^{-1}\). The low severity class has only slightly higher amounts of dead biomass, suggesting that the fire occurred only in the understory in this class. The moderate and high severity classes show increasing amounts of standing dead tree biomass consistent with their labels.

Areas with dead biomass values above 100 Mg ha\(^{-1}\) occurred almost exclusively within these recent fire boundaries. Interestingly, however, areas outside of the burn boundaries have dead biomass levels that are consistent with the areas having lower dead biomass inside the burn boundaries. Because North Rim forests that have not burned in recent crown fires are generally old (Fule et al., 2002a; Warren et al., 1982), standing dead wood is abundant, as in old-growth forests in general (Harmon et al., 1986) and more specifically in the forest types in our study area (Fule et al., 2002a). Our map suggests that the background dead biomass value in forests of the North Rim currently is <100 Mg ha\(^{-1}\). This observation is confirmed with the plot data collected for our study (Fig. 9). Knowing the background value of dead biomass in the forest enables one to estimate the additional contribution of dead standing tree biomass associated with a fire or other wide spread disturbances. For example, if we accept the first 100 Mg ha\(^{-1}\) of dead biomass as background value for this old forest, the additional values present in the moderate and high severity classes can be assumed to be the addition of dead biomass after the fire. Note that, for this particular exploratory analysis, there are temporal gaps between our lidar data acquisition (2007) and the fire incidents (year 2000 for Outlet fire, and 2005 for Dragon fire), which may be a source of error in Fig. 8.

It is important to note that there is a large proportion of dead biomass that we did not measure or map. This dead biomass was observed on the forest floor during fieldwork. According to Fule et al. (2002b), the amount of dead biomass on the floor of this forest is in the range of 3 to 35 Mg ha\(^{-1}\). A study in 1993 reported that the density of woody debris of unburned sites in this region averaged 37.3 Mg ha\(^{-1}\) (Duhnkrack 1993; Fule et al., 2004). Although there are likely statistical relationships between live and dead standing tree biomass and down woody biomass, the degree to which lidar data could directly observe this is likely to be limited. However, for a full accounting of carbon dynamics and wildlife habitat in older forests, it is important to account for this unmapped dead biomass in some way.

5. Summary and conclusion

This study is among the first to examine the potential of small-footprint, discrete return lidar data to distinguish between live and
dead standing tree biomass in mature forests, and to demonstrate the importance of lidar intensity data in accomplishing this. With few exceptions, studies have focused exclusively on estimating live biomass. Consistent with Lim et al. (2003), our results indicate that by taking advantage of lidar intensity data, a significant increase in estimation power for live biomass is possible. For example, when considering a single predictor variable (canopy volume, models 2 and 4, Table 5) the use of only high intensity values increased the R² in our models by 0.07 (0.65 to 0.72) over use of the full range of intensity values. For dead biomass, the peak of the low intensity frequency distribution (LoIntPk) was essential to estimation power. Our model based on LoIntPk alone (model 8) had an R² of 0.52, whereas models without this variable had R² values of 0.03 or less. In this study, we only used frequency distributions of intensity values and stratification of these into two classes. However, other characterizations of intensity are possible and need to be explored across a range of forest systems. For example, intensity as a function of height may prove to be even more important in how we use these data.

We used stepwise regression to select from an array of possible lidar-derived predictor variables, and for comparison, selected alternative sets of variables that we hypothesized should be useful in predicting biomass across a variety forested types. We discovered that single variable predictors (Table 5, models 2 and 8), performed nearly as well as four-variable models selected by stepwise regression (models 1 and 5). For live and dead biomass, respectively, canopy volume of high intensity returns and LoIntPk were the alternative predictors tested (Table 6 and Fig. 5). Whether these single variable parameters (parameterized for local conditions) are useful in other systems needs to be tested. Intensity stratification for modeling live and dead standing tree biomass may be unnecessary. Volume calculated from unstratified data and HilmPk appeared to be as good a model for live biomass estimation as stratified models in terms of predictive strength and errors. Our two best models for dead biomass were based on non-stratified data.

Biomass maps derived from selected regression models depicted the complementary nature of live and dead biomass in the mature forests of our study area on the North Rim of the Grand Canyon National Park (GCNP). Areas with large amounts of live biomass had low amounts of dead biomass, and vise versa. The maps depicted a background range of dead biomass in the mature forests of our study site of zero to 100 Mg ha⁻¹, which was confirmed by our field data. Areas having dead biomass values above the background values were associated almost exclusively with areas recently burned by wildfire.

Patterns of dead biomass were variable between and within the two recent fires mapped by the Monitoring Trends in Burn Severity Project (MTBS). The unburned areas within each MTBS burn polygon had values of dead biomass consistent with the background value outside of the burned polygons. The low severity MTBS burn class had only slightly elevated dead biomass values, indicating mostly understory burning. Significant increases in dead biomass were noted in both the moderate and high severity burn classes. This clearly demonstrates the complementary nature of Landsat-based disturbance maps and lidar-derived biomass maps. Further examination of this complementarity for disturbances from insects and other pathogens, wind, and other factors is warranted.

We have demonstrated that lidar-based maps are useful for examining ecological characteristics and monitoring processes in the forest regions of GCNP. Our focus was on biomass and fire, but extensions to wildlife habitat, carbon dynamics, and climate change are logical and expected.

Acknowledgements

This research was funded by the Ecosystem Processes Program of the Pacific Northwest Research Station within the USDA Forest Service and the Southern Colorado Plateau Network of the National Park Service's Inventory & Monitoring Program. We greatly thank the lidar team of Watershed Sciences, our lidar data provider, for their after-sales service, Robert J. McGaughey and Stephen E. Reutebuch for their assistance with the Fusion software, Professor Manuela Huso for her advice on our statistical inquiries, and Professor Mark Harmon for his advice on calculation of dead tree biomass. We also thank two anonymous reviewers for valuable comments.

References


