Towards scale-invariant aboveground biomass estimation in Savanna ecosystems using small-footprint waveform lidar

Dave Kelbe

Follow this and additional works at: http://scholarworks.rit.edu/theses

Recommended Citation
Towards Scale-invariant Aboveground Biomass Estimation in Savanna Ecosystems using Small-footprint Waveform Lidar

Dave Kelbe
Advisor: Dr. Jan van Aardt
Senior Project, Center for Imaging Science
Rochester Institute of Technology
Summer 2010

1. ABSTRACT

Land degradation is becoming an issue of increasing concern in the savanna ecosystems of southern Africa. As a result, there is a growing need to map structural changes at the fine scale, while retaining the ability to aggregate up to landscape level for analysis across land use gradients. Aboveground biomass (AGB) is an important indicator of vegetation structure and therefore is the ideal variable for estimation from light detection and ranging (lidar) data. To avoid the effects of scale, this paper takes a tree-delineation approach for segmentation of the structurally heterogeneous savanna environment. Diameter at breast height (DBH) measurements collected in-field are then regressed against lidar-derived statistics to estimate DBH on a per tree basis, from which biomass follows naturally by allometry. The result is a spatially explicit biomass map of the savanna environment, believed to be one of the first of its kind, that can be scaled by aggregation of per-tree biomass distributions.

2. INTRODUCTION

Land degradation is quickly becoming a force of global importance.\textsuperscript{1} Most poignant in the delicate savanna ecosystem of South Africa, the effects of increasing populational pressure in rural environments have already resulted in a number of observable degradation processes (Figure 1). Typical processes involve a reduction in grass production, changes in plant species composition, soil erosion, increases in tree cover (bush encroachment), and excessive wood removal.\textsuperscript{2}

Researchers in environmental remote sensing seek to quantify these changes at the small scale, while retaining the ability to compare estimates across larger, landscape scales. This is often done by measuring aboveground biomass (AGB), because AGB is an important indicator of ecosystem health and sustainability, and therefore an important proxy for assessing land cover and land degradation. AGB here is defined literally as the mass of the wood, foliage, and branch components that make up individual trees and shrubs.

Figure 1. The pressures of large rural populations in the former homelands (a) surrounding Kruger National Park have had increasingly degrading effects on ecosystem health. Such impacts include (b) cattle overgrazing and (c) firewood harvesting from trees.
Though aboveground biomass can be measured accurately and precisely in field, such methods typically suffer from high cost and limited geographic coverage, often requiring destructive sampling (modeling) which is not ideal for such studies emphasizing conservation. However, the growing need for spatially explicit biomass mapping has been partially compensated by recent advances in remote sensing technologies. In particular, light detection and ranging (lidar), with the ability to record three-dimensional vegetation structure, has become the modality of choice for measurement of AGB.

Airborne lidar is an active remote sensing technique in which a scanning laser emits pulses that travel through the canopy and reflects off the ground, with each interaction causing a portion of the pulse energy to be reflected back to the sensor (Figure 2). In discrete return systems, only the position and intensity of a discrete number returns are digitized. In waveform systems, however, the full backscattered energy profile is recorded. This provides greater information via advanced processing methods, for the improvement of canopy and ground height estimates and for the generation of denser 3D point clouds.

Figure 2. As the waveform lidar laser pulse travels through the canopy, it interacts with the tree structure, causing reflections of the original emitted pulse. The sensor records the time of arrival of the backscattered, reflected pulse (vertical axis) and its intensity (horizontal axis). These variables can be coupled with the speed of light to determine range distance and therefore three dimensional structure of vegetation. Red cross hairs designate the time locations of canopy interactions.

This research seeks to use lidar technology for scale-invariant AGB estimation. Typical methods in establishing lidar-based models for estimating landscape-level AGB involve regression analysis for relating certain lidar metrics to the spatially coincident in-situ measurements. Upon validation, these regressed models are then applied to the remainder of the lidar data and spatial coverage for prediction. However, most estimation models in previous studies are scale dependent. In other words, models are fitted to objects at a given scale, e.g., forest stands, and therefore must be applied at a scale equal to the size used in model fitting.

However, scale dependence may be avoided by inventoring biomass at the individual tree scale. AGB at scales above tree levels can then be derived simply by integrating the results from tree-level estimation, up to the desired scale. Yet computational limitations and inter-site variability prevent successful tree segmentation over large geographic areas. Therefore, for large area inquiries, it is still desirable to obtain biomass estimates from a plot-level model. The conundrum is solved by using the spatially explicit individual tree biomass as synthesized...
reference data for large-scale modeling. In this way, the effects of scale may be minimized in subsequent model generation.

The objective of this study is to adapt the tree delineation approach outlined in Zhao et al. (2009) to map biomass in the savanna ecosystem using both waveform and discrete return lidar, in a way that minimizes the scale constraints previously outlined. The approach is adapted to also segment small trees, shrubs, and bushes, which are typically found in savanna environments. The watershed algorithm is used to segment a smoothed canopy height model (CHM). The diameter at breast height (DBH) of trees measured in-field are then regressed against lidar-derived metrics of corresponding computer-delineated trees. After validation, the regression model is applied to all segmented trees, from which biomass can be determined based on allometric relationships. Each tree’s component (wood, leaves, branches) biomass is assigned a Gaussian distribution at the spatial location of the lidar-derived trees. Creation of a such a scaleable biomass map provides a means of carbon inventory and monitoring land degradation, and therefore is an appropriate tool for conservation management and policy development (e.g. rural land use limitations, fire wood rotations, livestock carrying capacity, etc.)

3. LITERATURE REVIEW

Success of the individual tree-delineation approach as precursor to mapping AGB relies on choosing and implementing the appropriate methods for division of the feature space into correct segments. Yet traditional, spectral reflectance-based measurements have proven to be an ill-suited tool for analyzing the fine scale structure in both dense and complex canopy environments. However, with the advancement of increasingly sophisticated, small footprint lidar systems, segmentation can be performed on a three-dimensional point cloud of data, a structural data set much more suited to the geometric segmentation problem we are faced with.

Most of the proposed methods for single tree detection seek to find the local maxima in a canopy height model (CHM), from which the surrounding canopy can be delineated through either a watershed or slope-based algorithm. Other methods include local maxima detection or local maxima filtering with fixed or variable window sizes. Reitberger (2009) implemented a three-dimensional segmentation approach in the coniferous Bavarian Forests National Park of Germany, based on normalized cut segmentation. Furthermore, the k-means clustering algorithm was been applied to the structural segmentation task in the alpine conifer forests of Switzerland by Morsdorf (2003).

Yet these methods have been developed for conifer forests, which assume a vertical stem located at the center of a conical tree structure. Deciduous trees, however, have a relatively flat canopy surface making treetops difficult to detect with conventional methods. Chen (2006) implemented a marker-controlled watershed segmentation in an open oak woodland in California and found that such an approach was an ineffective solution to a complicated structural environment. His study was the only of its kind focused in deciduous environments.

The savanna woodland of the protected Kruger National Park and surrounding degraded rangelands of Bushbuckridge present a unique study site for the development of lidar-based tree segmentation techniques. In general, tree structure is characterized by irregularly shaped canopies (an effect which is further accentuated by fire wood harvesting and animal grazing) and many sparse, small trees and shrubs. This research attempts to draw from the available methods to adapt previously developed segmentation techniques to this complicated structural environment.

4. MATERIALS

4.1 Study Area

The study area is bounded by (22°8’00’’ S; 30°34’52’’E) and (25°32’48’’S; 32°2’50’’E) in South Africa (Figures 3 and 4) and spans from east to west across a conserved, to increasingly rural and degraded land use gradient. In other words, airborne lidar coverage consists of state owned conservation lands in Kruger National Park, private conservation areas in Sabie Sands game reserve, and communal rangelands in Bushbuckridge Municipality (Figure 4). While the protected conservation lands provide an example of healthy, natural savanna, the rural rangelands of Bushbuckridge are severely degraded by the large populational pressure. The result is that this research has unique implications towards policy development and appropriate land management.
4.2 Field Measurements

Field data for this research were collected from 10 sites in the study area, each 50 x 50 m in size. A total of 36 plots (2-5 meters variable radius) were laid out within each site on a 10 m spacing, resulting in a grid-like pattern. All trees within the specified plot radius were measured for DBH and height. DBH is an important measurement in that it is the input to widely developed allometric equations predicting biomass. In addition, differentially-corrected GPS locations of selected trees were recorded along with their corresponding DBH. Furthermore, in April 2010, 300 trees were measured for DBH and marked with a GPS across the above-mentioned land use gradient.

4.3 Lidar Data

Airborne lidar data were collected in April 2008 by the Carnegie Airborne Observatory Alpha System (Table 1. This system consists of an integrated imaging spectrometer and a small-footprint scanning lidar system. The spectrometer is co-mounted on a stabilized plate with the scanning LiDAR operating at up to 100 kilohertz with full waveform digitization. The system simultaneously collects discrete-return lidar and waveform lidar on a common data system. Each scene pixel consists of an incoming (received) waveform with 256 bands at 1ns (0.15m) intervals. The waveform of the outgoing pulse, associated with each incoming waveform (40 bands with 1 ns spacing), was also provided. The ground height above sea level for each pixel was extracted from a digital elevation model (DEM), which was derived from coincident discrete return lidar data, following the extraction of ground returns using Terrasolid software (V. 008.001). Calibration showed absolute horizontal accuracies of ±0.05 – 0.08 m and vertical accuracies of ±0.06 – 0.14 m for the lidar returns.

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>waveform lidar</td>
<td>1064 nm wavelength</td>
</tr>
<tr>
<td></td>
<td>12-bit dynamic range</td>
</tr>
<tr>
<td></td>
<td>1 nanosecond temporal resolution</td>
</tr>
<tr>
<td></td>
<td>up to 440 slices per laser shot</td>
</tr>
<tr>
<td></td>
<td>up to 100 kHz laser repetition</td>
</tr>
<tr>
<td></td>
<td>.56 mrad beam divergence</td>
</tr>
<tr>
<td></td>
<td>.56 m spatial resolution at 1300 m a.g.l</td>
</tr>
</tbody>
</table>

Table 1. CAO Alpha system specifications.
5. METHODOLOGY AND RESULTS

5.1 Overview

Previous studies have urged caution when drawing conclusions from the full waveform parameters of pulse width and intensity. For example, a wide echo with low amplitude does not necessarily come from vegetation. In fact, the shape of the nth pulse depends on the scattering characteristics of the n-1 previous echoes. Furthermore, the recorded intensity values are yet a fairly noisy and unclearly defined term. The intensity is affected by many factors: target characteristics, lidar system, scan geometry, etc., and requires calibration and signal processing before use. Therefore, rather than attempt to draw conclusions from waveform parameters that have yet to be clearly defined, this research seeks to derive measurements only from the structural point cloud of lidar data. A flow chart of the methodology is outlined in Figure 5. Though the afore-mentioned waveform parameters are not used in this research, waveform lidar data is still used in order to generate a 3D point cloud, denser that what would have resulted from a discrete return system.

The overall methodology is adapted from Zhao et al. (2009). In this study a tree delineation approach is taken to map AGB at the individual tree level. Results can then be aggregated to create a fine resolution biomass map, providing synthesized data for modeling and testing at the landscape level. This paper outlines only the tree segmentation approach for biomass estimation, outlined in the upper dotted rectangle in Figure 5. Further research will develop the scaleable modeling in the bottom half of Figure 5.

The algorithm is summarized as follows: First, pre-processing is done on the raw waveform data to extract the location of canopy interactions in 3D space. Height-variable smoothing is then performed on the canopy height model, which is then inverted and segmented with the watershed transform. Validation is performed to optimize tree (and shrub) delineation at the site level. The DBH of field-GPS'ed trees are then regressed against lidar-derived metrics of corresponding trees. Results of the regression model are subsequently applied to predict the DBH for each segmented tree in the lidar data. DBH is used as the metric of interest because of its estimation reliability and ease of input into well developed biomass allometric equations. Finally, a general savanna allometric equation is applied to determine per-tree biomass, which is then apportioned to a fine scale AGB map. Though images and statistics in the following section refer to only one site (Figure 6), tree delineation was performed on all available sites across the land use gradient.

Figure 5. A flow chart summarizing the procedures of the individual tree-based approach for mapping per-tree biomass in order to simulate reference data for modeling and validation of a landscape-level model. In this research only the steps in the tree-delineation approach are presented.
5.2 Pre-processing

The distance resolution of the received lidar waveform $P_r(t)$ is limited by both the time resolution of the detector and the convolution of the outgoing waveform $P_t(t)$, the receiver impulse response $\Gamma(t)$, and the true response distribution of the target, or cross section $\sigma_t(t)$.

$$P_r(t) = P_t(t) * \sigma(t) * \Gamma(t)$$  (1)

The loss of resolution can be retrieved mathematically by deconvolution. The superposition of noise, however, makes direct deconvolution impractical due to noise amplification. Therefore, a frequency-based lowpass filter, followed by a stable time-domain signal deconvolution is implemented based on the Richardson-Lucy algorithm.

This is an iterative process in which the mathematical solution of $\sigma(t)$ can be expressed as

$$\sigma_{i+1}(t) = \sigma_i(t) \left[ \frac{P_r(t)}{h(t) * \sigma_i(t) * h(t)} \right]$$  (2)

where $h(t) = P_t(t) * \Gamma(t)$ and $i$ denotes the iteration. The residual is computed as

$$r_i(t) = P_r(t) - h(t) * \sigma_i(t)$$  (3)

and will converge as the algorithm progresses towards.

The ground peak of the waveform was then registered to the appropriate spatial location on the DEM, and a trigonometric relationship used to perform angular correction of off-nadir scans. At this point the waveform data corresponding to a 50 x 50 meter site, such as that shown in Figure 6, is a matrix of intensity values and can be visualized as shown in Figure 7. This site will be used to demonstrate the algorithm methodology in the following sub sections.

![Figure 6.](image1.png) ![Figure 7.](image2.png)

Figure 6. Spectral image of the example study site; red channel: 450.9 nm, green channel: 535.9 nm, blue channel: 650.9 nm. Figure 7. Matrix of intensity voxels at 0.56m horizontal resolution and 0.15m vertical resolution.

The final step in preprocessing is to discretize the waveform in real coordinate space. A maxima searching algorithm was implemented to detect non-ground peaks of the waveform. A threshold was added to ensure peaks due to noise were not digitized. The intensity of each relevant peak was then attributed to a point in three dimensional space. The original waveform data are a matrix of voxels with resolution in x, y, and z space, of 0.56, 0.56, and 0.15 meters, respectively. The result after processing is a dense, intermittent point cloud irregularly distributed in 3D space, that depicts the locations of interactions of the original laser pulse with the canopy (Figure 8). Although the points are color coded by height for visualization (in Figure 8), each echo is attributed the intensity of the waveform to be used in regression analysis.
5.3 Segmentation

Previously developed segmentation methods focus on structurally homogeneous conifer stands, for which segmentation is a relatively easy task. This study area, however, is defined by structurally diverse deciduous trees and shrubs. Imagine a large irregularly shaped tree crown with protruding branches and valleys. Clustering or segmentation could misinterpret such variability as unique, separate trees. This is especially the case given the presence of many smaller, more compact trees and shrubs located in close proximity. While over-smoothing would succeed in the first case, it would fail in the second instance by merging many unique, smaller scale trees. Conversely, while a lack of adequate smoothing would retain fine-scale bush structure, the process would fail by over-segmenting large, variable tree crowns.
Therefore, an adaptive smoothing method was employed for application to the savanna environment. Canopy height is modeled by measuring the difference between first and last echos of the discrete return lidar.\textsuperscript{23} The unsmoothed canopy height model, shown in Figure 10, becomes the input to the segmentation process.

![Canopy height model before smoothing.](image1)

Figure 10. Canopy height model before smoothing. Note that the scale on the z axis is exaggerated to highlight the peaks and valleys in large tree crowns. Segmentation of this surface would yield significant commission errors or false positives.

The input was then smoothed based on a height-variable mean filter. The square filter should have a dimension determined by the predicted crown size based on the height of the tree. Crown size, defined as the average crown diameter along two perpendicular directions,\textsuperscript{9} is found to have larger variability when a tree is higher. Therefore, in order to avoid violating the assumption of homoscedasticity, a nonlinear power model was fitted, as shown in Equation 4.

\[
Height = 0.4 \times CrownSize^{1.3} \tag{4}
\]

As shown in Figure 11, such smoothing has the effect of reducing the spurious local maxima in variable tree crowns that could contribute to over segmentation.

![Height-variable smoothing represses irrelevant local maxima and minima in the CHM, while retaining the structure of small trees and shrubs.](image2)

Figure 11. Height-variable smoothing represses irrelevant local maxima and minima in the CHM, while retaining the structure of small trees and shrubs. Here the smoothed CHM is overlain with red crosses, indicating the location of segmented trees.
The smoothed CHM was then segmented using the watershed transform. The process of watershed segmentation can be illustrated in terms of flooding simulations. Imagine inverting the smoothed CHM as in Figure 12. The result is a series of watershed regions, much like river and valley topography, into which a drop of water placed on any point of the surface will flow. In other words, consider flooding this inverted surface from its minima. We prevent the waters from different sources from merging by building up dams between catchment basins. Individual segments are delineated by dams or watershed lines, and each delineated catchment basin corresponds to one tree.

This result is shown in Figure 13, where the dams are represented by grey watershed lines separating different segments and the delineated trees were coded by unique colors. Figure 14 is included so that the reader can affirm that regions such as that in the upper right hand corner are not in fact over-segmented. To be clear, though the segment area may appear to be larger than the canopy of the corresponding tree, in practice this is simply due to the nature of the algorithm, and there simply may be no canopy lidar points outside of the central area of the segment. The dark blue background represents the locations of the CHM surface that are close enough to the DEM to be considered ground or grass. Therefore, for computation, these were removed during segmentation. The centers of lidar-derived trees were found by locating the position of the maximum height for each segment. In both figures, the locations of lidar-derived tree centers are included, in addition to the GPS locations of field-measured trees.
Initially, a clustering scheme adapted from the iterative k-means approach\textsuperscript{25} was applied to refine the two-dimensional segmentation returned from the watershed transform. Cluster analysis is a well known statistical tool for dividing feature spaces into similar areas and has been effectively used for lidar-based tree delineation.\textsuperscript{13} Based on the knowledge that taller trees will have greater internal variability than small shrubs, the adapted algorithm was optimized for the savanna environment using additional parameters to adjust the cluster distance metric based on the height and number of points in any given cluster. However, since the initial segmentation provided consistent, satisfactory results, this step was omitted.

Though the 3D clustering scheme was not implemented in the final tree detection algorithm, the results obtained by the watershed transform were used to delineate the point cloud into tree clusters based on solely two dimensional segmentation (Figures 15 and 16). These tree-delineated point clouds were used to extract additional lidar metrics for regression with the field-measured DBH.
Figure 15. Nadir view of the segmented point cloud derived from waveform lidar data and watershed segmentation of the CHM.

Figure 16. Isotropic view of segmented point cloud derived from waveform lidar data and watershed segmentation of the CHM.

5.4 Validation

As noted previously, the degree of smoothing affects the success of tree delineation in terms of false positives and negatives. In the first case, if inadequate smoothing is applied, small peaks and troughs in large, variable tree canopies will lead to errors of commission. This is a problem of over-segmentation, as illustrated in Figure 17d. The second case is the result of smoothing at too coarse a factor, in which neighboring trees will merge and therefore form a tree group instead of single trees (Figure 17b). An appropriate balance (Figure 17c) between undersegmentation and oversegmentation is crucial, since DBH is tied to tree segment parameters, and the regression results could be skewed by inadequate segmentation of trees.

Figure 17. From left to right: a) CHM overlaid with reference and lidar-derived tree locations. b) Undersegmentation, due to excessive smoothing, merges small neighboring trees into one large tree. c) Adequate segmentation d) Oversegmentation due to deciduous tree structure and inadequate smoothing.

Therefore, a validation stage was added to ensure that the segmentation results were optimized when compared to the measured field data. Validation was performed by visually comparing the segmentation results to the canopy height model and locations of the field-GPS’ed tree locations. In this site, 9 out of 11 trees were segmented correctly, while two trees were undersegmented with only one lidar segment for three in-field GPS’ed trees. Though it is commendable to achieve valid segmentation results for all structures, tree delineation accuracy of larger trees is the most important, as these large trees are the dominant contributors to AGB. Therefore, it was assumed that these omission errors of small shrubs were less troublesome than omission/comission errors of the large tree canopies.
5.5 Regression with Field-Measured DBH

The next step involved estimation of DBH for all segmented trees by developing a model in which lidar parameters of GPS’ed trees are regressed against known in situ DBH measurements. DBH estimation is desired as an intermediary step to biomass for several reasons. First, DBH is the most stable predictor for regression models which seek to determine tree biomass, when compared to those that incorporate tree height and crown diameter. Second, robust allometric equations which relate biomass and DBH in the South African savanna environment are readily available, and require no additional parameters. Finally, the estimation of biomass directly from lidar variables is ill-advised because the “truth” biomass of corresponding trees is found by the dbh-dependent allometric equations, and therefore invokes some circular reasoning and compounds modeling errors. Therefore, DBH is the most stable and scalable variable to estimate. A simple allometric equation can then be applied to derive AGB, with lidar-derived DBH as the independent variable.

Trees of various species, structures, and a wide range of heights were marked in field with a differential GPS (Figure 18). In each case the DBH was measured using a tape measure (Figure 19).

Figures 18-19. GPS’ed trees varied in size and structure, from < 1m to > 10m. In each case a tape measure was used to measure diameter at breast height (DBH).

For each GPS’ed tree, lidar-derived statistics of the corresponding data segment were regressed in Statistical Analysis Software (SAS) to find the most appropriate model. Variables extracted from the 3D point cloud included tree height, canopy size in the x, y, and z direction, and the number of discretized point per canopy. Furthermore, though it was noted that intensity is uncalibrated, the total intensity of all returns in the lidar point cloud was included as an additional independent variable, in addition to the total intensity of the original, unprocessed waveforms corresponding to the tree canopy. The latter variables were included based on the mindset of data mining, which seeks to

The number of discretized points in the canopy was found to be the best predictor of dbh, with an $R^2$ value of 0.82. Therefore, this model was used to estimate the DBH on a per tree basis of all segmented trees, based on the delineated point clouds.

5.6 Fine Resolution AGB Map

The goal of this effort was the creation of a spatially explicit AGB map. However, AGB information is difficult, time-consuming and often destructive to collect if the direct method is applied: cut down the tree, strip off its leaves and fruit, chop up the wood, weigh the components, etc. This approach is impractical for a whole plot, so typically a shortcut -allometry- is applied. The generalized, species-independent allometric equations of Nickless et al. (in print) were used. The biomass equation is reproduced in Equation 5 and has an $R^2$ of 0.98, with 443 samples ranging in DBH from 0.3 cm to 33 cm.

$$AGB = \exp(-3.47 + 2.83 \times \ln(DBH))$$ (5)
where \( AGB \) is the aboveground biomass of a tree with a given diameter at breast height (\( DBH \)).

Each segmented tree in the site area was then apportioned a total biomass value based on Equation 5. To develop a spatially explicit biomass map, component biomass was estimated based on the ratio equations plotted in Figure 20, which estimate the percentage of total AGB due to branches, wood, bark, and foliage.

![Figure 20. Ratio equations based on tree DBH estimate the AGB due to branches, wood, bark, and foliage.](image)

The wood and bark biomass were assigned to the pixel at the location of the tree bole, and the branch and foliage biomass were distributed using a Gaussian distribution centered at the bole, with a standard deviation related to the radius of the tree canopy.

6. DISCUSSION

The results for this site are shown in Figure 21. Note that the large tree in the upper left hand corner seems to have less biomass than some smaller trees. This is due to the fact that the overall tree biomass is distributed over a larger canopy area. This conclusion can be confirmed by referencing Figure 22. In this figure, the biomass is only assigned to the pixel at the bole of the segmented tree. Here it is clear, as expected, that larger trees do indeed have greater biomass.

The reader should remember that the application of this biomass map is intended for computation and not visualization. In fact, though the resolution of the biomass map is 0.56 meters, equivalent to the captured data and imagery, this fine resolution was chosen only for convenience in that it is identical to the input data. Furthermore, though the component ratios seek to apportion biomass using a Gaussian distribution within the tree canopy, in reality it is not advisable to draw conclusions from the biomass at scales below the tree level.

However, these Gaussian biomass distributions can now effectively be aggregated to the scale of choice, without concern for scale-dependence of models or estimates.
Future work will use this procedure for spatially explicit individual-tree biomass mapping to analyze biomass patterns at the landscape level. This biomass map will serve as a reference data for estimating biomass at levels above the individual tree and for potentially developing a new landscape-level model based not on individual tree delineation but on the raw lidar data itself.

There are several sources of error which we hope will be minimized in future work. The sample size for regression analysis of the DBH was based only on the 15 GPS’ed trees in the site shown and therefore any outliers could have a significant impact on the regression. Future work will use all GPS’ed trees in the 10 sites provided, in addition to approximately 300 tree-level measurements collected in April 2010. This will bolster the sample size for regression and statistically improve the regression results.

Currently, all vegetation below a height of 0.5 meters were excluded from the analysis. However, by comparing the tree heights measured in-field and from the lidar data, it was found that true height is often underestimated by lidar on the order of one meter. In the future we wish to more clearly define the line between lidar data that is included or excluded for analysis. Specifically, we wish to exclude shrubs and bushes under one meter which contribute little to overall biomass, but contribute a great deal of variability in terms of their DBH.

Furthermore, the field data, which serves as “truth,” contain some errors in the way effective DBH of multi-stemmed trees are calculated. For example, a tree with 5 stems of 5 cm DBH each is calculated as $5 \times 5 = 25$ cm DBH. This multi-stemmed tree does not have the same biomass as a single-stemmed tree of 25 cm DBH. Thus, field measurements which sum individual stem diameters lead to overestimation of DBH for small, multi-stemmed trees and shrubs. A more appropriate equation for multi-stem trees should improve the fit of the regression.

Finally, we must remember that though DBH is the most stable predictor for biomass estimation, its estimation is somewhat removed from the lidar statistics, and therefore remains an inherently difficult parameter to estimate based on height and number of points alone. The results were therefore deemed appropriate, given this challenge.

7. CONCLUSIONS
This research used a segmentation approach to apportion tree biomass in a spatially explicit map. Watershed segmentation was performed to regress segment-level lidar statistics with field measured DBH on a per-tree basis. The result was a tree-level inventory of AGB, from which larger scale estimates may be found by integration up to the desired scale. Future work will use this map as reference data for generation of lidar models at the landscape level. From this, ecosystem health, structure and land cover may be quantified across large scales in the increasingly degraded savanna environment of South Africa.
REFERENCES