Enhanced forest interior estimations utilizing lidar-assisted 3D forest cover map

Gang Shao\(^a\), Basil V. Iannone III\(^b\), Songlin Fei\(^b\),\(^*\)

\(^a\) Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907, USA
\(^b\) School of Forestry and Resource Conservation, University of Florida, Gainesville, FL 32611, USA

Abstract

Accurate estimation of forest interior is essential for large scale sustainable forest management. Conventional 2D forest cover maps only indicate the presence and absence of forests and may introduce quantity and allocation bias in the delineation of forest interior due to the inclusion of early successional/recently disturbed forests. We addressed this issue by utilizing both lidar data and National Land Cover Database (NLCD) to generate a 3D forest cover map in the forested region of western North Carolina, USA. We first classified forest cover as either early or late successional based on vertical structural information (using either a single-height threshold or a variable-height threshold by forest type). Then, we estimated forest interior based on these different forest cover classes and assessed disagreement in quantity and allocation of interior forests as determined using 2D forest cover maps. In addition, we determined spatial relationships between developed areas and forest interiors to evaluate the impact of landscape settings on the estimation of forest interior. Our results indicated that using a single-height threshold, the 3D forest cover map was able to distinguish early vs. late successional forests with a classification accuracy of 96.6%. A similar classification accuracy (94.6%) was achieved when applying variable-height thresholds by forest type. Based on the single-height threshold method, excluding early successional forests (6.4% of all forested area) reduced estimates of forest interior by 10.3%, 9.6% and 10.4% at spatial resolutions of 4.4 ha, 39.7 ha and 234 ha, respectively. Using variable-height threshold by forest type, the estimates of forest interior were approximately 1% less than the estimates using the single-height threshold method due to the slightly decrease of the early successional forest classification (5.7% of all forest area). Our results indicated the forest interior may be overestimated without accounting for successional stage. Moreover, geospatial distance analysis revealed the overestimation of forest interior to be most pronounced in highly-fragmented areas. Our study demonstrated the advantage of considering 3D structural information for the accurate estimation of forest interior, particularly in areas already having fragmented forests. This information can be relatively easily obtained from lidar data. This 3D method will allow for the creation of high-accuracy forest interior maps that can help to answer a variety of ecological questions and improve forest management and conservation.

1. Introduction

Forest interiors (or core forest) are essential for the conservation and restoration of overall forest biodiversity (Chazdon, 2008; Chiabai et al., 2011; Franklin, 1993; Lindenmayer and Franklin, 2002; Morford et al., 2011). These areas are defined by a local dominance of forest coverage (Wickham et al., 2007) and are buffered from the environmental influence of nearby non-forested land or of nearby young, recently disturbed forests. These interior areas provide particular physical and biological conditions required for certain plants and animals (Faaborg, 1995; Foley et al., 2005; Riitters and Wickham, 2012), and tend to resist invasion from many light-limited exotic species (Fei et al., 2008; Kuhman et al., 2010). Therefore, accurate estimation of forest interior is important for forest monitoring and assessing biodiversity, as well as a broad range of other ecological applications (Kerr and Ostrovsky, 2003; McIntyre and Hobbs, 1999).

In recent years, there has been a great deal of research efforts aimed at estimating forest interior and its change at landscape levels (Broadbent et al., 2008; Heilman et al., 2002; Kacholi, 2014; Laurance et al., 2007; McDonald and Urban, 2004; Ritters et al., 2016). These efforts have been broadly based on forest cover maps, which are commonly derived from optical remote sensing imagery through forest

---

\(^*\) Corresponding author.
E-mail address: sfei@purdue.edu (S. Fei).

https://doi.org/10.1016/j.ecolind.2018.06.037
Received 27 November 2017; Received in revised form 23 April 2018; Accepted 18 June 2018
Available online 21 June 2018
1470-160X/ © 2018 Elsevier Ltd. All rights reserved.
biochemistry properties (Friedl et al., 2002; Hansen et al., 2010; Heilman et al., 2002) and canopy phenology (Collins and Woodcock, 1996; Shao et al., 2014; Zhu et al., 2012). Optical-imagery-based forest cover maps provide valuable information for forest interior estimation distinguishing forests from other land covers. However, these maps, being two dimensional (2D), only contain presence-absence information of forest cover without regard to forest successional stage. While early-stage, regenerating forests may eventually provide the physical and biological functions of forest interiors, they currently do not (Chazdon, 2008). Their inclusion when estimating forest interior may lead to overestimates of the amount of forests that exhibit characteristics of forest interior. However, exclusion of early-stage forest from the forest interior estimation can be challenging using optical-imagery-based, 2D forest cover maps, as algorithms relating spectral features to forest age are often complex (Huang et al., 2010) and often require multi-temporal data coverage (Kennedy et al., 2007; Pflugmacher et al., 2012).

Fortunately, with the increased availability and improved computational technologies, airborne lidar (light detection and ranging) can provide region-wide information on forest structure, which offers opportunities to overcome the problems associated with using optical remote sensing to separate forest succession stages (Lim et al., 2003; Wulder et al., 2012). Lidar can directly observe the three-dimensional (3D) structural information of forests by measuring the distance with the round-trip travel time of a reflected laser pulse. This information can be used to develop a canopy height model (CHM), which in turn can be used to assess forest disturbance and regeneration by relating succession stages to vertical canopy structure (Bonnet et al., 2015; Koukoulas and Blackburn, 2004; Lefsky et al., 1999; Vepakomma et al., 2008). Therefore, CHM can be used in conjunction with 2D forest cover maps to separate forests into areas comprised of early and late successional stages, potentially helping to increase the accuracy of forest interior delineation.

In this study, we demonstrate the utility of lidar-assisted forest cover maps for the regional-level delineation of forest interiors in the forests of western North Carolina, USA. The main objective is to investigate the effects of early-stage forest cover on forest interior delineation. Specifically, we (1) assess the accuracy of lidar-assisted forest interior delineation with the consideration of successional stages and (2) identify the quantity and spatial allocations of estimation bias of forest interior caused by the inclusion of early-stage forest cover.

2. Materials and methods

2.1. Study area

The study area spans over 16 counties (a total area of 17,989 km$^2$) in western North Carolina (Lat. 35.2–36.5°N, Long. 81.3–84.5°W), USA, including parts of the Great Smoky Mountain, Black Mountain, and Kings Mountain ranges of the southern Appalachian Mountains (Fig. 1). The elevation of the study area is 1342–2004 m in the Great Smoky Mountains, 1199–2038 m in the Black Mountains and 498–685 m in the Kings Mountains. The major land-cover type in this region is forest with mainly state and national forests in the northwest and commercial timberlands in the southeast. Red spruce (Picea rubens), Fraser fir (Abies fraseri), American beech (Fagus grandifolia), yellow birch (Betula alleghaniensis) and maple (Acer spp.) trees dominate higher elevations (above 1000 m). At lower elevations, eastern hemlock (Tsuga canadensis) trees dominate moist bottom and shady slopes, while pine (Pinus spp.) and oak (Quercus spp.) trees dominate dry ridges and exposed slopes.

2.2. Generation of a 3D forest cover map

We integrated airborne lidar data and an existing land cover map to generate a forest cover map containing 3D structural information of our study area. Lidar data were used to provide structural information of forests. These data were acquired in March-April of 2005 by the North Carolina Floodplain Mapping Program. The flying altitude of the acquisition was 305 m above ground level with a scan angle of 55°. The average point density was 0.25 pts/m$^2$. Lidar point cloud data were classified into ground and non-ground points using LasTools. A Digital Terrain Model (DTM) was then generated by interpolating the ground points using the natural neighbor algorithm. A Digital Surface Model (DSM) was also calculated by identifying maximum elevation within each pixel. The CHM representing the structural information of the study area’s forests (shown in Fig. 1) was then computed by subtracting the DTM from the DSM.

National Land Cover Data (NLCD) 2006, a 16-class land cover map of the USA, was utilized to generate a conventional 2D forest cover map representing the presence-absence of forests. In our study area, there are three major forest types (deciduous, evergreen, and mixed deciduous-evergreen), which covered 99.7% of the forest areas. Therefore, these three forest types were combined as forest cover, while all other land cover types were grouped as non-forest cover (Ritters and Wickham, 2012). For integration, both CHM and NLCD were set with the same spatial resolution (30 by 30 m) and origin. The resulting 3D forest cover map was then used to delineate forest interior either with or without regard to successional stage.

To distinguish early vs. later-stage forest cover, we applied both single- and variable-height threshold classification approaches. The single height threshold approach separated all forest areas into early and later-stage forests with one height cut-off value, which was determined using a data training procedure described as follows. First, a total of 100 sample points were randomly placed within the forested areas shown on a high resolution (2 m) aerial photo of our study region that was acquired from the National Agriculture Imagery Program (NAIP). These points were then classified as either early- or late-stage forests. The CHM-based canopy heights of the points were then compared to derive a range of potential threshold values to differentiate between early and late stage (i.e., a break occurring between the lower limit of later-stage forests and the upper limit of early-stage forests). Within this range, the height threshold was adjusted by examining the heights of points occurring along the edges of early- and later-stage forests using the NAIP map as a reference. Lastly, for classification validation, two sets of 300 additional points were randomly generated within each of these two-succession forest covers (Congalton, 1991). The finalized later-stage forest cover was then used for 3D-based interior delineation. Our training procedure suggested that the threshold to distinguish early vs. later-stage forests in our study area was 11 m. The classification validation showed 290 out of 300 (96.7%) samples were correctly classified as later-stage forest cover and 286 out of 300 (95.3%) samples were correctly classified in early-stage forests.

As our study area was relatively large and composed of different forest types, we also tested a variable-height threshold approach in our study with different height cut-off values for each of the three NLCD-based forest types (deciduous, evergreen, and mixed deciduous-evergreen) in the succession-stage classification (Fig. 2). Several literature studies defined the variable-height threshold by the average of canopy drip line for forest gap detection (Gaulton and Malthus, 2010; White et al., 2018). This approach requires field sampling for canopy drip line information, which is limited in our study. Instead, we used the mean canopy height information from the 3D forest cover map to estimate the variable height threshold for succession stage classification. In our study area, the deciduous forest had the highest mean canopy height (25.9 m), which was 4.3 and 3.6 m higher than evergreen (21.6 m) and mixed deciduous-evergreen (22.3 m) forests, respectively (Table 1). Therefore, we used the mean-height-weighted equation (Eq. (1)) to calculate the variable-height threshold value for each forest type based on its mean canopy height and the single threshold defined above.
2.3. Forest interior delineation

We employed the forest area density (FAD) method demonstrated by Riitters and Wickham (2012) to estimate forest interiors. The FAD approach delineates forests interior using 2D information of forests presence-absence. This delineation is influenced by two parameters—spatial resolution and dominancy. Spatial resolution is the size of neighborhoods used for the FAD calculation. The selection of this spatial resolution is based on the mobility and arrangement of species and habitats of interest. In this study, forest interior areas were delineated using multi resolution analysis with neighborhood sizes of 4.4 ha (7 by 7 pixels), 39.7 ha (21 by 21 pixels), and 234.0 ha (51 by 51 pixels), which represented three orders of magnitude of measurement resolution. Dominancy is the threshold of the percentage of forest coverage in neighborhoods used to distinguish forest interior from forest edge. The threshold value of dominancy can be chosen based on the range sizes and habitat needs of a particular species (Debinski and Holt, 2000), making this approach adaptable to a wide range of ecological questions (Fahrig, 2003). Here we defined dominancy akin Riitters and Wickham (2012) as 90%. This threshold of delineation is comparable with other landscape scale forest monitoring efforts in the USA (Fischer and Lindenmayer, 2007; McIntyre and Hobbs, 1999).

2.4. Assessment of 3D-based forest interior delineations

To assess our 3D-based interior delineations (both single and variable-height threshold approaches) and investigate the effects of early-stage forests on the delineation results, we used NLCD-derived 2D forest cover, which included both early- and later-stage forests, to delineate forest interiors and compared these 2D-based interior delineations with the prior 3D-based estimations. The 3D- and 2D-based delineations employed the same algorithm and parameters. Relative difference in

<table>
<thead>
<tr>
<th>Area (ha)</th>
<th>MIN Height (m)</th>
<th>MAX Height (m)</th>
<th>Mean Height (m)</th>
<th>STD Height (m)</th>
<th>Single T (m)</th>
<th>Variable T (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deciduous Forest</td>
<td>1215767</td>
<td>0</td>
<td>54.4</td>
<td>25.9</td>
<td>7.1</td>
<td>11</td>
</tr>
<tr>
<td>Evergreen Forest</td>
<td>98145</td>
<td>0</td>
<td>49.8</td>
<td>21.6</td>
<td>7.54</td>
<td>11</td>
</tr>
<tr>
<td>Mixed Deciduous-Evergreen</td>
<td>39713</td>
<td>0</td>
<td>48.5</td>
<td>22.3</td>
<td>7.06</td>
<td>11</td>
</tr>
</tbody>
</table>
percent (RDP) was used to evaluate the differences between 3D- and 2D-based estimates of forest interiors at different spatial resolutions (Bennett and Briggs, 2011):

$$\text{RDP} = \frac{A_{\text{3D}} - A_{\text{2D}}}{A_{\text{3D}}}$$

where $A_{\text{3D}}$ and $A_{\text{2D}}$ are the areas of the 3D- and 2D-based forest interior estimation, respectively. We also assess the differences between single and variable-height threshold approaches by comparing the succession-stage classification, the forest interior estimates and the compositions of each forest type (deciduous, evergreen and mixed deciduous-evergreen) in the forest interior estimates.

We then applied geospatial distance analysis to identify spatial allocations of disagreements between 3D and 2D-based approaches in forest interior estimates. The 2006 NLCD was used to locate the developed areas that served as reference locations (Origins) for this analysis. The Euclidean distances between all forested areas (interior and exterior portions) and the edges of developed areas were calculated as the reference of the overall forest canopy structural distribution. Moreover, the Euclidean distances of 3D-based and 2D-based forest interiors to the edge of developed areas were also calculated for comparison. Then, areas of disagreement between 3D- and 2D-based forest interior estimations and the 3D-based edge forests (total forest cover subtracted from forest interiors) were calculated to further evaluate the spatial patterns of potential bias introduced by including early-stage forests into estimations of forest interiors.

### 3. Results

With the single-height threshold approach, our model detected a total of 1,270,264 ha of later-stage forests, which was about 93.6% of the total forest area (Table 2). The remaining 6.4% of forest cover was classified as early-stage forest cover. Based on the classification validation, the classification accuracy for later- and early-stage forests are 96.7% and 95.3%, respectively. The overall, area-weighted classification accuracy is 96.6%. After the exclusion of early-stage forests from our 3D forest cover map, we estimated the total forest interior in our study area to be 903,965 ha (66.6% of total forest cover), 784,855 ha (57.8% of total forest cover) and 680,714 ha (50.1% of total forest cover) at the spatial resolutions of 4.4 ha, 39.7 ha, and 234.0 ha, respectively (Table 2). Compared to 2D-based models, 3D-based models reduced forest interior estimations at all three spatial resolutions (Fig. 4). The RDP of forest interior reduction was 10.3%, 9.6% and 10.4% for the spatial resolutions of 4.4 ha, 39.7 ha, and 234.0 ha, respectively. The overall, area-weighted classification accuracy decreased from 96.6% to 94.6%. The total area of forest interior estimated with this variable-height threshold approach

![Fig. 3. Later-stage forest cover map of the study area derived from lidar-assisted 3D forest cover map. Pink areas represent early-stage, regenerating forest, i.e., those areas controlled for by our 3D method that could result in biased estimates of forest interior. Aerial photo on the upper right side (NAIP 2005) is provided as a reference. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)](image-url)
were 915884 ha (67.5% of total forest cover), 794632 ha (58.5%) and 688816 ha (50.7%) at the spatial resolutions of 4.4 ha, 39.7 ha, and 234.0 ha, respectively, which were slightly larger compared to the estimates derived from the 2D and 3D single-height threshold based methods, most of the forest interior reduction (forest interiors not recognized by the 3D method) occurred within 500 m of developed areas (Fig. 6d), where were dominated by edge forests (Fig. 6b). The canopy height distributions of forest interior and interior reduction were relatively stable in relation to varying distance to development (Fig. 6c). In contrast, the average canopy height of edge forests was less than that of forest interior and decreased closer to developed areas (Fig. 6b, c).

4. Discussion

Our results revealed that early-stage forest cover had considerable effects on forest interior delineation in the study area and the feasibility of forest succession stage classification, integrating lidar-derived CHM with existing land cover map, NLCD, for overcoming this source of bias. The single-height threshold delineation results of forest interiors showed that by accounting for and eliminating the 6.4% of forest cover classified as early-stage forests reduced estimates of forest interior by approximately 10% at all three spatial resolutions; Similarly, the variable-height threshold delineation results reported an approximately 9% forest interior reduction at all three spatial resolutions with an elimination of 5.7% of forest cover as early-stage forests. Both of these results supports the findings in Riitters and Wickham (2012) that the reduction of forest coverage may fragment forests and lead more loss of forest interior. Conventional 2D forest cover had insufficient information of forest succession and using it solely without consideration of canopy height leads to overestimates of forest interior. By integrating 2D forest cover with forest structural metrics that can easily computed from readily available lidar data such as NEON program (MacLean, 2017) and Open Topography Lidar dataset (Shao et al., 2018), we can greatly improve the estimation of forest interior (both in quantity and spatial distribution) at regional scales.

In this study, we applied both single- and variable-height threshold approaches to classify forest cover into early- and later-stage forests, which showed minimal differences in the classification results. One possible reason for the limited differences observed is because the major of the forest in our study area are deciduous forest (over 90%). Therefore, different height threshold for different forest type may only affect about 10% of the classification results. Additionally, the horizontal resolution of our 3D forest map was 30 × 30 m, which is relatively coarse for forest species or forest type separation as one pixel may include several trees with different types. The variable-height threshold approach slightly improved the classification accuracy of early-stage forest but reduced more of the accuracy of later-stage classification by overestimation. Therefore, the overall accuracy of variable-height threshold approach (94.6%) was lower than the single-height threshold approach (96.6%). White et al. (2018) presented a similar result that variable-height threshold approach had a lower classification accuracy compared to single-height threshold approach, when characterizing forest canopy gaps in the coastal temperate rainforests in British Columbia, Canada. These results may imply that the single-height threshold approach is sufficient for a simple—2-class—forest succession-stage classification with an acceptable accuracy (96.6%). To improve the classification accuracy by using variable-height threshold, it may require higher resolution forest maps and better thresholding algorithms to find the optimal height cut-off values for each forest type.

Though the single-height threshold value (11 m) for succession stage classification that is specific for our study area and is higher than

![Maps of 3D forest interiors (single-height threshold approach) in the study area, including estimates of canopy height (first row), and maps showing areas that were determined to be forest interior for 2D but not 3D method (red) at spatial resolutions of 4.4 ha (a), 39.7 ha (b) and 234.0 ha (c) (second row). Developed areas (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)](image-url)
thresholds (2–5 m) reported in some previous studies (Homer et al., 2004; Kane et al., 2011; St-Onge et al., 2014; Vepakomma et al., 2008), it is close to the value (10 m) reported by St-Onge and Vepakomma (2004) and the value of 11 m reported by Zhang (2008), both of which had a similar point density of lidar data as our study. Threshold determination is mainly dependent on the quality of lidar data and CHM spatial resolution. High density, small spatial scale, lidar data is able to create high resolution CHM, in which each individual tree is presented by more than one pixel. In this case, the threshold can be set at individual tree level based on the knowledge of local stand conditions. This threshold, however, may still need to be defined differently for different stand types within the study area. In contrast, at lower lidar return densities, CHM resolution may be low enough that a single pixel can cover more than a single tree. When this occurs, the height values of forest pixels represent the average height of forest canopy. For these instances, a threshold based on data training information will likely provide much better results across a broad range of spatial scales than possible from user-defined criteria.

Forest interior is a key indicator of deforestation and forest degradation, which are global challenges for forest conservation and restoration (Herold and Johns, 2007). Our study offers an easy approach to obtain the spatial patterns of forest interiors. In our study area, the majority of forest interior were located between 200 and 1200 m away from developed areas, and forest interior reductions were mainly within 500 m from developed areas (Fig. 4d). The interior reductions showed similar structural patterns as forest interiors indicating these reduction areas were ‘new’ edge forests converted from forest interiors due to recent disturbances. This result supports the findings of Huang et al. (2015) that industry logging and urban growth, i.e., the major form of forest disturbance, mostly occur near already developed areas in this region. Miettinen et al. (2011) found similar spatial patterns of forest fragmentation in Southeast Asia, where the deforestation in open and developed areas was four times higher than the average deforestation between 2000 and 2010. These findings indicate that locations of potential deforestation and forest degradation may have particular focal areas, but that where these areas occur may vary among different regions likely due to variability in topographic, social and economic factors. Accurate forest structure information including forest interior coverage is essential to capture such differences.

5. Conclusion

This study demonstrated the feasibility and importance of including information on vertical structure when estimating regional abundance of forest interior. The exclusion of 6% early-stage forest lands from the analysis can results in a reduction of about 10% in estimates of forest interior across all spatial resolutions compared to traditional 2D forest cover based estimations. With the increasing availability of lidar data,
Fig. 6. The canopy height distribution of total forest cover (a), edge forests (b), 3D-based forest interiors (c) and for areas determined to be forest interior by the 2D but not 3D-based estimations (d) at spatial resolution of 39.7 ha. The heat maps were canopy height distributions standardized based on the canopy height distribution of 3D-based forest interiors (c).

3D forest cover map can provide explicit spatial information of forest canopy conditions and succession stages for the improved delineation of forest interiors, allowing forest managers and decision makers to devise effective strategies to manage forests sustainably.

Acknowledgement

This study was funded in part by NSF Macrosystems Biology grant DEB1638702.

References


Kerr, J.T., Ostrovsky, M., 2003. From space to species: ecological applications for remote


