Title

Simple method for direct crown base height estimation of individual conifer trees using airborne LiDAR data.

Journal

Optics express, 26(10)

ISSN

1094-4087

Authors

Luo, Laiping
Zhai, Qiuping
Su, Yanjun
et al.

Publication Date

2018-05-01

DOI

10.1364/oe.26.00a562

Peer reviewed

Research Article | Vol. 26, No. 10 | 14 May 2018 | OPTICS EXPRESS A562

Optics EXPRESS

Simple method for direct crown base height estimation of individual conifer trees using airborne LiDAR data

LAIPING LUO,[1,2] QIUPING ZHAI,[3,4] YANJUN SU,[2,3,6] QIN MA,[2] MAGGI KELLY,[5] AND QINGHUA GUO[2,3,4,7]

[1]*State Key Lab of Urban Environmental Processes & Digital Simulation, College of Resource Environment and Tourism, Capital Normal University, Beijing 100048, China*
[2]*Sierra Nevada Research Institute, School of Engineering, Sierra Nevada Research Institute, University of California at Merced, Merced, CA 95343,USA*
[3]*State Key Laboratory of Vegetation and Environmental Change, Chinese Academy of Sciences, Institute of Botany, Beijing 100093, China*
[4]*University of Chinese Academy of Sciences, Beijing 100049, China*
[5]*Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720, USA*
[6]*suyanjun1987@gmail.com*
[7]*guo.qinghua@gmail.com*

Abstract: Crown base height (CBH) is an essential tree biophysical parameter for many applications in forest management, forest fuel treatment, wildfire modeling, ecosystem modeling and global climate change studies. Accurate and automatic estimation of CBH for individual trees is still a challenging task. Airborne light detection and ranging (LiDAR) provides reliable and promising data for estimating CBH. Various methods have been developed to calculate CBH indirectly using regression-based means from airborne LiDAR data and field measurements. However, little attention has been paid to directly calculate CBH at the individual tree scale in mixed-species forests without field measurements. In this study, we propose a new method for directly estimating individual-tree CBH from airborne LiDAR data. Our method involves two main strategies: 1) removing noise and understory vegetation for each tree; and 2) estimating CBH by generating percentile ranking profile for each tree and using a spline curve to identify its inflection points. These two strategies lend our method the advantages of no requirement of field measurements and being efficient and effective in mixed-species forests. The proposed method was applied to a mixed conifer forest in the Sierra Nevada, California and was validated by field measurements. The results showed that our method can directly estimate CBH at individual tree level with a root-mean-squared error of 1.62 m, a coefficient of determination of 0.88 and a relative bias of 3.36%. Furthermore, we systematically analyzed the accuracies among different height groups and tree species by comparing with field measurements. Our results implied that taller trees had relatively higher uncertainties than shorter trees. Our findings also show that the accuracy for CBH estimation was the highest for black oak trees, with an RMSE of 0.52 m. The conifer species results were also good with uniformly high R^2 ranging from 0.82 to 0.93. In general, our method has demonstrated high accuracy for individual tree CBH estimation and strong potential for applications in mixed species over large areas.

OCIS codes: (280.3640) Lidar; (100.6890) Three-dimensional image processing.

References and links

1. E. Næsset and T. Økland, "Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve," Remote Sens. Environ. **79**, 105–115 (2002).
2. S. C. Popescu and K. Zhao, "A voxel-based lidar method for estimating crown base height for deciduous and pine trees," Remote Sens. Environ. **112**(3), 767–781 (2008).

#324724 https://doi.org/10.1364/OE.26.00A562
Journal © 2018 Received 23 Feb 2018; revised 11 Apr 2018; accepted 11 Apr 2018; published 27 Apr 2018

Optics EXPRESS

3. J. Vauhkonen, "Estimating crown base height for Scots pine by means of the 3D geometry of airborne laser scanning data," Int. J. Remote Sens. **31**(5), 1213–1226 (2010).
4. L. Korhonen, J. Vauhkonen, A. Virolainen, A. Hovi, and I. Korpela, "Estimation of tree crown volume from airborne lidar data using computational geometry," Int. J. Remote Sens. **34**(20), 7236–7248 (2013).
5. M. A. Lefsky, W. Cohen, S. Acker, G. G. Parker, T. Spies, and D. Harding, "Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests," Remote Sens. Environ. **70**(3), 339–361 (1999).
6. C. J. Gleason and J. Im, "Forest biomass estimation from airborne LiDAR data using machine learning approaches," Remote Sens. Environ. **125**, 80–91 (2012).
7. H.-E. Andersen, R. J. McGaughey, and S. E. Reutebuch, "Estimating forest canopy fuel parameters using LIDAR data," Remote Sens. Environ. **94**(4), 441–449 (2005).
8. S. D. Roberts, T. J. Dean, D. L. Evans, J. W. McCombs, R. L. Harrington, and P. A. Glass, "Estimating individual tree leaf area in loblolly pine plantations using LiDAR-derived measurements of height and crown dimensions," For. Ecol. Manage. **213**(1-3), 54–70 (2005).
9. G. S. Biging and M. Dobbertin, "Evaluation of competition indices in individual tree growth models," For. Sci. **41**, 360–377 (1995).
10. M. Dobbertin, "Tree growth as indicator of tree vitality and of tree reaction to environmental stress: a review," Eur. J. For. Res. **124**(4), 319–333 (2005).
11. S. J. Zarnoch, W. A. Bechtold, and K. Stolte, "Using crown condition variables as indicators of forest health," Can. J. For. Res. **34**(5), 1057–1070 (2004).
12. S. L. Stephens, "Evaluation of the effects of silvicultural and fuels treatments on potential fire behaviour in Sierra Nevada mixed-conifer forests," For. Ecol. Manage. **105**(1-3), 21–35 (1998).
13. M. Kelly, Y. Su, S. Di Tommaso, D. L. Fry, B. M. Collins, S. L. Stephens, and Q. Guo, "Impact of Error in Lidar-Derived Canopy Height and Canopy Base Height on Modeled Wildfire Behavior in the Sierra Nevada, California, USA," Remote Sens. **10**(2), 10 (2017).
14. L. A. Arroyo, C. Pascual, and J. A. Manzanera, "Fire models and methods to map fuel types: the role of remote sensing," For. Ecol. Manage. **256**(6), 1239–1252 (2008).
15. P. Foster, "The potential negative impacts of global climate change on tropical montane cloud forests," Earth Sci. Rev. **55**(1-2), 73–106 (2001).
16. S. Luo, J. M. Chen, C. Wang, X. Xi, H. Zeng, D. Peng, and D. Li, "Effects of LiDAR point density, sampling size and height threshold on estimation accuracy of crop biophysical parameters," Opt. Express **24**(11), 11578–11593 (2016).
17. T. J. Dean, Q. V. Cao, S. D. Roberts, and D. L. Evans, "Measuring heights to crown base and crown median with LiDAR in a mature, even-aged loblolly pine stand," For. Ecol. Manage. **257**(1), 126–133 (2009).
18. J. Hyyppä, H. Hyyppä, D. Leckie, F. Gougeon, X. Yu, and M. Maltamo, "Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests," Int. J. Remote Sens. **29**(5), 1339–1366 (2008).
19. Q. Ma, Y. Su, S. Tao, and Q. Guo, "Quantifying individual tree growth and tree competition using bi-temporal airborne laser scanning data: a case study in the Sierra Nevada Mountains, California," Int. J. Digit. Earth **5**, 1–19 (2017).
20. J. P. Dandois and E. C. Ellis, "High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision," Remote Sens. Environ. **136**, 259–276 (2013).
21. M. Kelly and S. Di Tommaso, "Mapping forests with Lidar provides flexible, accurate data with many uses," Calif. Agric. **69**(1), 14–20 (2015).
22. J. D. Muss, D. J. Mladenoff, and P. A. Townsend, "A pseudo-waveform technique to assess forest structure using discrete lidar data," Remote Sens. Environ. **115**(3), 824–835 (2011).
23. R. E. McRoberts, Q. Chen, and B. F. Walters, "Multivariate inference for forest inventories using auxiliary airborne laser scanning data," For. Ecol. Manage. **401**, 295–303 (2017).
24. C. Alexander, A. H. Korstjens, and R. A. Hill, "Influence of micro-topography and crown characteristics on tree height estimations in tropical forests based on LiDAR canopy height models," Int. J. Appl. Earth Obs. Geoinf. **65**, 105–113 (2018).
25. M. K. Jakubowksi, Q. Guo, B. Collins, S. Stephens, and M. Kelly, "Predicting surface fuel models and fuel metrics using Lidar and CIR imagery in a dense, mountainous forest," Photogramm. Eng. Remote Sensing **79**(1), 37–49 (2013).
26. A. A. Plowright, N. C. Coops, C. M. Chance, S. R. Sheppard, and N. W. Aven, "Multi-scale analysis of relationship between imperviousness and urban tree height using airborne remote sensing," Remote Sens. Environ. **194**, 391–400 (2017).
27. W. Li, Z. Niu, J. Li, H. Chen, S. Gao, M. Wu, and D. Li, "Generating pseudo large footprint waveforms from small footprint full-waveform airborne LiDAR data for the layered retrieval of LAI in orchards," Opt. Express **24**(9), 10142–10156 (2016).
28. H. Tang, M. Brolly, F. Zhao, A. H. Strahler, C. L. Schaaf, S. Ganguly, G. Zhang, and R. Dubayah, "Deriving and validating Leaf Area Index (LAI) at multiple spatial scales through lidar remote sensing: A case study in Sierra National Forest, CA," Remote Sens. Environ. **143**, 131–141 (2014).

Optics EXPRESS

29. G. Goldbergs, S. R. Levick, M. Lawes, and A. Edwards, "Hierarchical integration of individual tree and area-based approaches for savanna biomass uncertainty estimation from airborne LiDAR," Remote Sens. Environ. **205**, 141–150 (2018).
30. F. Pimont, J.-L. Dupuy, E. Rigolot, V. Prat, and A. Piboule, "Estimating leaf bulk density distribution in a tree canopy using terrestrial LiDAR and a straightforward calibration procedure," Remote Sens. **7**(6), 7995–8018 (2015).
31. C. Véga, J.-P. Renaud, S. Durrieu, and M. Bouvier, "On the interest of penetration depth, canopy area and volume metrics to improve Lidar-based models of forest parameters," Remote Sens. Environ. **175**, 32–42 (2016).
32. J.-H. Lee, G. S. Biging, and J. B. Fisher, "An individual tree-based automated registration of aerial images to lidar data in a forested area," Photogramm. Eng. Remote Sensing **82**(9), 699–710 (2016).
33. W. Xiao, S. Xu, S. O. Elberink, and G. Vosselman, "Individual tree crown modeling and change detection from airborne lidar data," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **9**(8), 3467–3477 (2016).
34. O. S. Ahmed, S. E. Franklin, M. A. Wulder, and J. C. White, "Characterizing stand-level forest canopy cover and height using landsat time series, samples of airborne LiDAR, and the random forest algorithm," ISPRS J. Photogramm. Remote Sens. **101**, 89–101 (2015).
35. J. Breidenbach, R. E. McRoberts, and R. Astrup, "Empirical coverage of model-based variance estimators for remote sensing assisted estimation of stand-level timber volume," Remote Sens. Environ. **173**, 274–281 (2016).
36. A. S. Maguya, K. Tegel, V. Junttila, T. Kauranne, M. Korhonen, J. Burns, V. Leppanen, and B. Sanz, "Moving Voxel Method for Estimating Canopy Base Height from Airborne Laser Scanner Data," Remote Sens. **7**(7), 8950–8972 (2015).
37. J. Holmgren, Å. Persson, and U. Söderman, "Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images," Int. J. Remote Sens. **29**(5), 1537–1552 (2008).
38. J. Holmgren and Å. Persson, "Identifying species of individual trees using airborne laser scanner," Remote Sens. Environ. **90**(4), 415–423 (2004).
39. S. Solberg, E. Naesset, and O. M. Bollandsas, "Single tree segmentation using airborne laser scanner data in a structurally heterogeneous spruce forest," Photogramm. Eng. Remote Sensing **72**(12), 1369–1378 (2006).
40. Y. Su, Q. Guo, B. M. Collins, D. L. Fry, T. Hu, and M. Kelly, "Forest fuel treatment detection using multi-temporal airborne lidar data and high-resolution aerial imagery: a case study in the Sierra Nevada Mountains, California," Int. J. Remote Sens. **37**(14), 3322–3345 (2016).
41. W. Li, Q. Guo, M. K. Jakubowski, and M. Kelly, "A new method for segmenting individual trees from the lidar point cloud," Photogramm. Eng. Remote Sensing **78**(1), 75–84 (2012).
42. Q. Ma, Y. Su, and Q. Guo, "Comparison of Canopy Cover Estimations From Airborne LiDAR, Aerial Imagery, and Satellite Imagery," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **10**(9), 4225–4236 (2017).
43. M. K. Jakubowski, W. Li, Q. Guo, and M. Kelly, "Delineating individual trees from LiDAR data: A comparison of vector-and raster-based segmentation approaches," Remote Sens. **5**(9), 4163–4186 (2013).
44. A. Matkan, M. Hajeb, B. Mirbagheri, S. Sadeghian, and M. Ahmadi, "Spatial analysis for outlier removal from LiDAR data," Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci **40**(W3), 187–190 (2014).
45. J. R. Ben-Arie, G. J. Hay, R. P. Powers, G. Castilla, and B. St-Onge, "Development of a pit filling algorithm for LiDAR canopy height models," Comput. Geosci. **35**(9), 1940–1949 (2009).
46. Q. Guo, W. Li, H. Yu, and O. Alvarez, "Effects of topographic variability and lidar sampling density on several DEM interpolation methods," Photogramm. Eng. Remote Sensing **76**(6), 701–712 (2010).
47. M.-V. Piazza, L. A. Garibaldi, T. Kitzberger, and E. J. Chaneton, "Impact of introduced herbivores on understory vegetation along a regional moisture gradient in Patagonian beech forests," For. Ecol. Manage. **366**, 11–22 (2016).
48. D. Reed, L. Washburn, A. Rassweiler, R. Miller, T. Bell, and S. Harrer, "Extreme warming challenges sentinel status of kelp forests as indicators of climate change," Nat. Commun. **7**, 13757 (2016).
49. M. H. Zwietering, I. Jongenburger, F. M. Rombouts, and K. van 't Riet, "Modeling of the bacterial growth curve," Appl. Environ. Microbiol. **56**(6), 1875–1881 (1990).
50. J. B. Drake, R. O. Dubayah, D. B. Clark, R. G. Knox, J. B. Blair, M. A. Hofton, R. L. Chazdon, J. F. Weishampel, and S. Prince, "Estimation of tropical forest structural characteristics using large-footprint lidar," Remote Sens. Environ. **79**(2-3), 305–319 (2002).
51. X. Yu, J. Hyyppä, A. Kukko, M. Maltamo, and H. Kaartinen, "Change detection techniques for canopy height growth measurements using airborne laser scanner data," Photogramm. Eng. Remote Sensing **72**(12), 1339–1348 (2006).
52. K. Zhao, J. C. Suarez, M. Garcia, T. Hu, C. Wang, and A. Londo, "Utility of multitemporal lidar for forest and carbon monitoring: Tree growth, biomass dynamics, and carbon flux," Remote Sens. Environ. **204**, 883–897 (2018).
53. X. Lu, Q. Guo, W. Li, and J. Flanagan, "A bottom-up approach to segment individual deciduous trees using leaf-off lidar point cloud data," ISPRS J. Photogramm. Remote Sens. **94**, 1–12 (2014).
54. A. Ferraz, S. Saatchi, C. Mallet, S. Jacquemoud, G. Gonçalves, C. A. Silva, P. Soares, M. Tomé, and L. Pereira, "Airborne lidar estimation of aboveground forest biomass in the absence of field inventory," Remote Sens. **8**(8), 653 (2016).
55. D. G. Schweikert, "An interpolation curve using a spline in tension," Stud. Appl. Math. **45**, 312–317 (1966).

Optics EXPRESS

1. Introduction

Tree crown base height (CBH) is defined as the distance from ground surface to the lowest live branch within a tree crown [1–3]. It is a basic structure parameter for measuring crown volume [4], crown ratio [5, 6], and other crown dimensional parameters [7, 8] both at individual tree and forest stand levels. Thus, there is a need for accurate and efficient estimations of CBH for tree growth observation [9, 10], tree health monitoring [11], forest wildfire management [12, 13], forest ecosystem modeling [14], and global climate change studies [15]. However, obtaining CBH is still a challenging task. Traditional field-based methods can obtain reliable and accurate CBH measurements [16, 17], but these methods are often time-consuming and labor-intensive. Remote sensing technique makes it possible to continuously observe forest parameters with higher efficiency [18, 19]. In particular, light detection and ranging (LiDAR), an active remote sensing technique, has the capability to penetrate forest canopy, and can provide precise estimations of three-dimensional tree structures [20, 21].

Many previous researchers have shown that LiDAR data are effective at estimating forest structural parameters [22–25]. Various algorithms have been successfully developed to estimate tree height [26], leaf area index [27, 28], aboveground biomass [29], canopy cover [7, 30, 31] at both individual tree [29, 32, 33] and forest stand scales [34, 35]. However, studies on LiDAR-based CBH estimation are still insufficient, particularly at individual-tree scale. Most previous researchers investigated the estimation of CBH at plot level [1, 7, 25, 36]. In the recent decade, studies on individual tree-level CBH estimation are increasing to provide more precise forest structural information to support forest management planning and decision making [2, 3, 37].

Currently, methods for LiDAR-based CBH estimation can be divided into two main categories, regression-based method [38, 39] and direct estimations [2, 3]. The regression-based methods require to develop a relationship between field-measured CBH and other tree biophysical parameters derived from LiDAR data, such as tree height and crown width [2, 38, 39]; then the CBH in areas without field measurements can be predicted based on the experimental relationships. The accuracy of CBH estimation using regression-based methods have relied heavily on the reference data sampled in the field. The quality and quantity of field measurements may greatly impact the accuracy and predictability when applying these regression models to other areas [3, 13]. In contrast, direct methods estimate the CBH of individual trees directly from the LiDAR point cloud without the requirement of field measurements [2, 3]. The direct LiDAR methods are free from the uncertainties brought by field measurements, and thus they are more flexible in CBH estimation in forests with different species. Nevertheless, previous studies using direct methods often encountered the problem of overestimation, which was usually caused by the weak capability of laser pulses to penetrate dense canopy and reach the lowest tree branch, particularly in dense forests [2, 3, 25]. To overcome this problem, Popescu and Zhao [2] generated vertical profiles based on the density and intensity of voxelized LiDAR point for each tree, and then calculated CBH as an inflection point of the polynomial curve fitted from the vertical profiles. They obtained a good result with a root mean squared error (RMSE) of 2 m. Vauhkonen [3] presented a 3D geometry-based algorithm to estimate the CBH for scots pine trees (*Pinus sylvestris*) with RMSEs ranging from 1.44 m to 3.56 m. These methods substantially improved the accuracies of CBH, but they can only be applied to limited species in most cases because of the high computational costs associated with the complex procedures of generating vertical profiles or reconstructing the 3D convex of individual trees. Therefore, it is still necessary and important to develop new methods which are simple and computationally efficient, and can be applied on multiple species simultaneously.

The main objective of this study is to develop a simple method to estimate CBH directly from airborne LiDAR data. This method is based on frequency statistics and inflection point detection from the percentile ranking curves generated from LiDAR data. Using the field

measured tree structural information over a mixed-conifer dominated forest in the Sierra Nevada, California as a pilot study, we systematically evaluated our method over different tree species and tree sizes. This proposed method has the potential to apply to other forests and provide useful CBH estimations for forest and wildfire management.

2. Data and methods

2.1 Study area

The study site is located in the Tahoe National Forest of California, U.S.A (Fig. 1), with an area of 92.1 km^2. This area experiences a Mediterranean climate characterized by warm dry summers and cool wet winters. The terrain in this area is very complex, with the elevation ranging from 228 m to 2189 m [40]. The main vegetation in this site is the Sierra Nevada mixed conifer, which consists of white fir (*Abies concolor*, denoted by ABCO), California red fir (*Abies magnifica*, denoted by ABMA), incense cedar (*Calocedrus decurrens*, denoted by CADE), sugar pine (*Pinus lambertiana*, denoted by PILA), ponderosa pine (*Pinus ponderosa*, denoted by PIPO), Douglas-fir (*Pseudotsuga menziesii*, denoted by PSME). The forest includes smaller amounts of broadleaf species, such as Black oak (*Quercus kelloggii*, denoted by QUKE) and Canyon live oak (*Quercus chrysolepis*, denoted by QUCH). The variety of species in this site makes it an ideal study area to evaluate the robustness of our proposed method for CBH estimation.

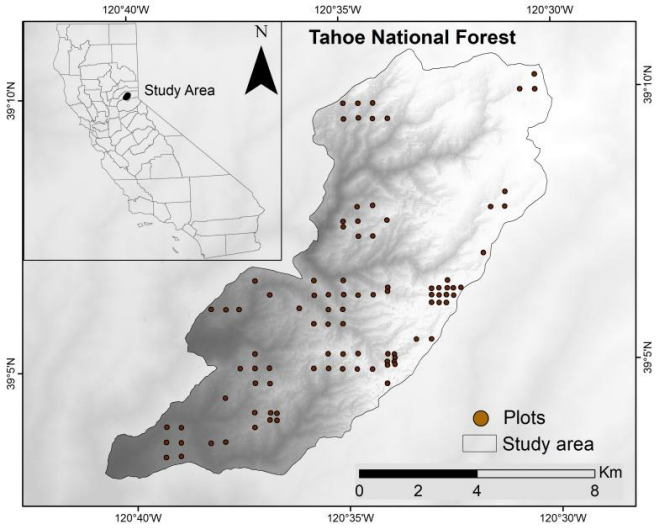

Fig. 1. The map of study area and the locations of field measured plots.

2.2 Field measurements

Field data were obtained during the summer of 2007 in the study site to obtain forest structural information. The first plot was randomly located, and others were regularly placed with a distance of 500m. The plot center was located using a Trimble^{TM} GeoXH GPS, and each plot was a circle with a radius of 12.62 m. Any plot within 12.62m of a road surface or build-ups was moved 25m in an arbitrary cardinal direction to avoid human impacts. The plots located on private lands or inaccessible areas were not measured. With each plot, the location, CBH, diameter at breast height (DBH), and tree height of each tree with a DBH>5 cm were measured and recorded. A total of 94 plots and 637 trees were surveyed in this study site for validation.

Optics EXPRESS

2.3 LiDAR data

The LiDAR data were acquired for the study area on September 2007, using a Cessna twin-engine Skymaster equipped with an Optech GEMINI Airborne Laser Terrain Mapper platform [41]. The laser sensor recorded up to four returns for each laser pulse. The pulse repetition frequency was 100 kHz to 120 kHz. The flight height was about 900 m above the ground. Each flight line had a belt overlap of 67% [19, 42]. The scanning angle of the platform was from −15° to 15°. The spatial precision of the LiDAR data along the vertical and horizontal directions were approximately 10 cm and 10-15 cm, respectively. The average density of the LiDAR data was around 10 points/ m^2.

2.4 CBH estimation methods

The proposed CBH estimation procedure includes three main steps: (1) individual tree segmentation, (2) understory and noise removal, and (3) individual tree CBH calculation. The workflow is shown in Fig. 2 and a graphic example is shown in Fig. 3.

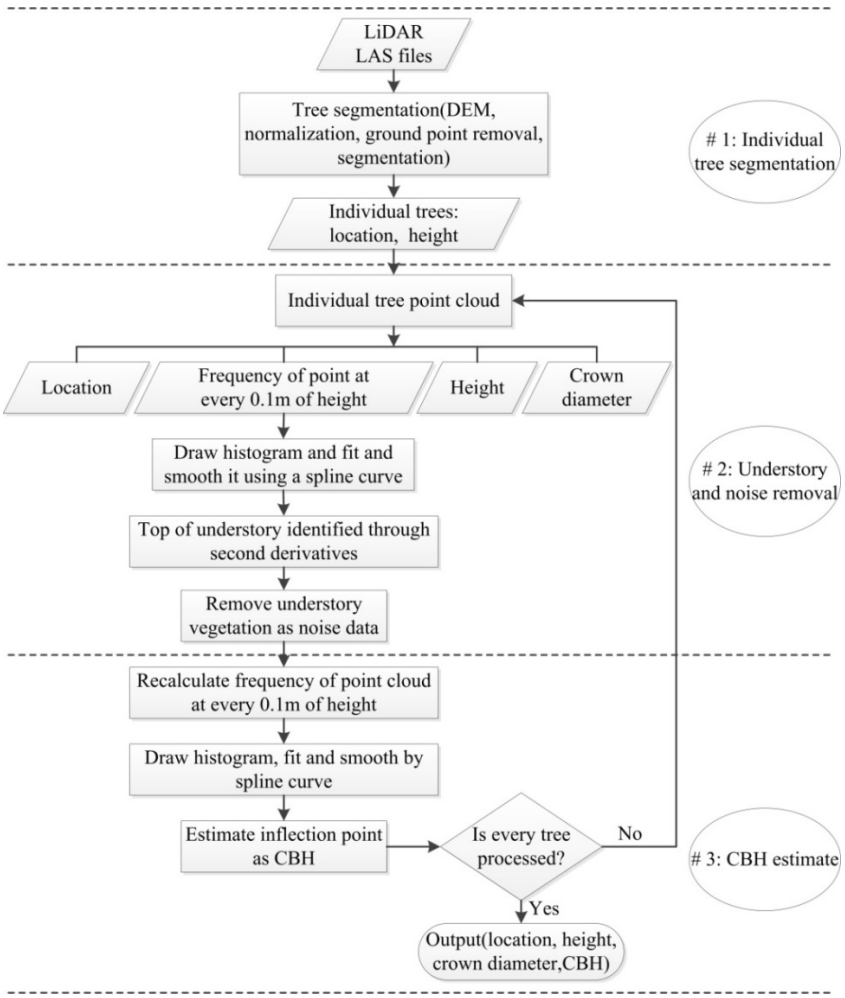

Fig. 2. Flow chart of the method for estimating crown base height (CBH).

Optics EXPRESS

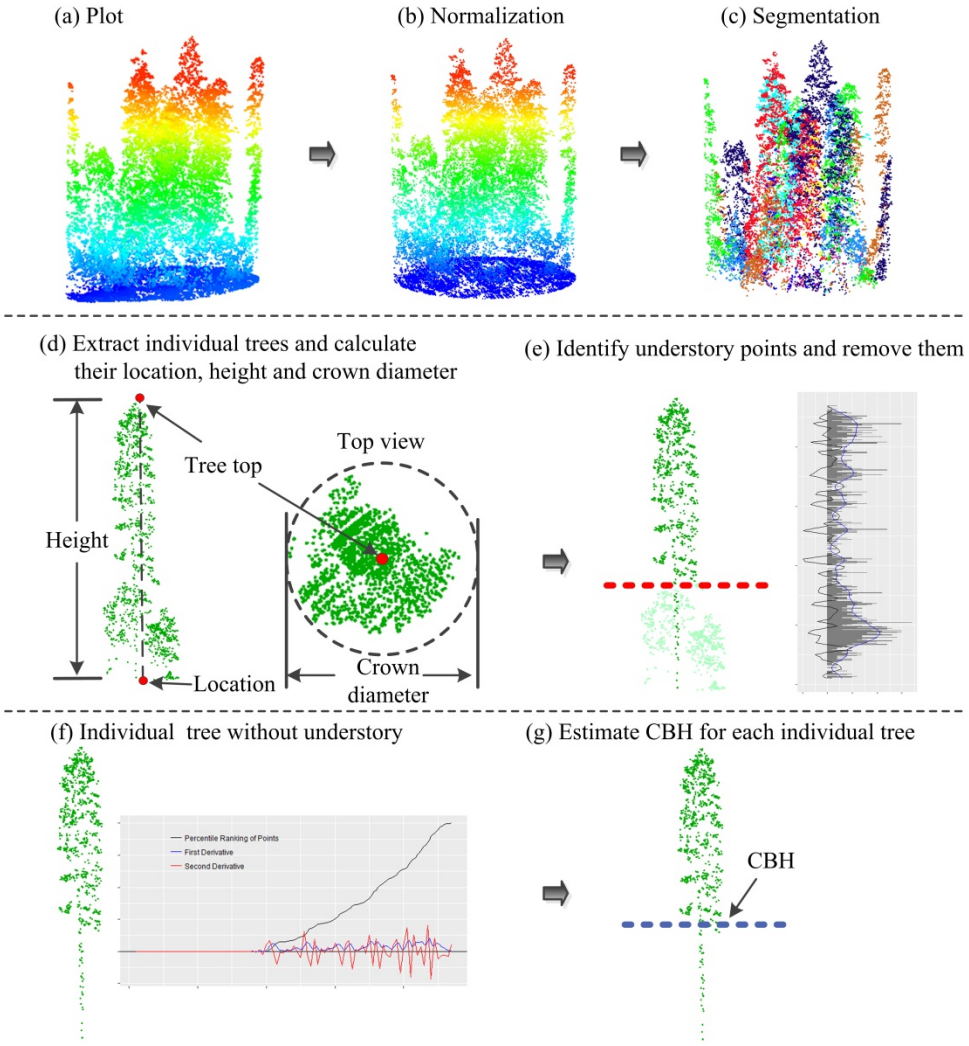

Fig. 3. An example of estimating individual tree CBH from LiDAR point cloud data.

2.4.1 Individual tree segmentation

A top-to-bottom region growing algorithm [41] was used to segment individual trees from the normalized LiDAR point cloud in LAS files. This algorithm segments trees according to the relative distance between trees. Trees may overlap in dense forest, but the space still exists between them at the tops of trees. The algorithm implements the segmentation from the tallest point of a tree to the shortest point of that tree. It has been shown to perform well on individual tree segmentation in complex mixed conifer forests [43]. Three preprocessing steps were conducted to the raw LiDAR point cloud prior to application of this algorithm, which included outlier removal, normalization and ground point removal. Outlier removal is necessary to prepare LiDAR data for individual tree segmentation [44]. We applied a Gaussian filter [45] with 5 times the standard deviation and a 10 * 10 window to detect outliers. Next, we used the ordinary Kriging algorithm to generate a DEM at a resolution of 1 m from ground returns [46]. The elevation of each LiDAR point was normalized to above-ground height by subtracting the DEM pixel value at the same location. Points below 1 m were considered as shrubs according to field survey, and were removed from the CBH estimation. We operated the above-mentioned procedures using the LiDAR360 software

(http://greenvalleyintl.com/). After the tree segmentation procedure, tree location was recorded as the treetop location, and tree height was derived as the above ground height of treetop, and crown diameter was calculated as the diameter of the circle which can contain all the points projected to the ground with the tree top as its center [Fig. 3(d)].

2.4.2 Understory and noise removal

The understory structure and height of a given forest may vary significantly by environmental conditions and species type [47, 48]. The top of understory is often close to the CBH, which may cause difficulties in identifying crown base [2]. Thus, in order to eliminate the impact of understory on CBH estimation, removing understory and noise is necessary but often neglected in previous studies. In this study, we proposed an automatic method to remove the understory and noise by detecting the top of understory beneath individual trees. Through observation and investigation of LiDAR point clouds and field measurements, we found that LiDAR point density was always lower and even close to zero in the vertical gap between understory and crown base. Therefore, an accurate detection of the vertical gap where the point cloud density reduced sharply is the key to identify the top of understory. In order to characterize the vertical changes of point density, we generated a histogram of the vertical LiDAR point distribution at 0.1 m intervals and generated a smooth curve from this histogram for each individual tree. The 0.1 m interval was selected here because it is fine enough to form a smooth curve which kept the detailed changes in point density. The inflection points of the smoothed curve, which represented points with a sign of the curvature changes [49], were then identified as potential points of understory tree top. Detailed steps for selecting the understory tree top for understory and noise removal are introduced as an example below.

In order to identify the inflection points appropriately, we first applied the spline curve-fitting algorithm to smooth the histogram (the blue curve in Fig. 4). Then we drew a second derivative curve (the red mixed with black curve in Fig. 4) corresponding to the smoothed curve and got multiple inflection points where the second derivative values were zero on each curve (*a, b,…, z,…* points in Fig. 4). These inflection points can be the local maximum, minimum, or stationary points of the point density. The challenging task was how to select the correct inflection point (indicating the top of understory) from the candidate points. This was determined by the following procedures. 1) We highlighted the sub-ranges of second derivative curve which has positive signs (referred as above-zero parts here after), shown as the red curves in Fig. 4. These above-zero parts of curves indicated that the point cloud densities were lower than the adjacent below-zero parts (e.g., point density in *ab* is lower than *bc* in Fig. 4). 2) We calculated the point density for each of the above-zero height region, and chose the region with the lowest point density (in Fig. 4 *mn* has the lowest density at 0.5). 3) We compared the height of *m* with that of *n*, and selected the inflection point with lower height as the top of understory. If the height of *m* is lower than that of *n, m* would be marked as the top of understory, otherwise *n* would be marked as the top of understory (e.g. in Fig. 4, *mn* is the region with the lowest point density and *m* was set to the top of understory). After that, we removed the LiDAR points which were below the selected inflection point (*m* as shown in Fig. 4) as understory and noise for individual trees prior to CBH estimation. This procedure was successful at detecting most understory points. Some points which were too isolated and discrete to be completely detected between the top of understory and the crown base were remained.

Optics EXPRESS

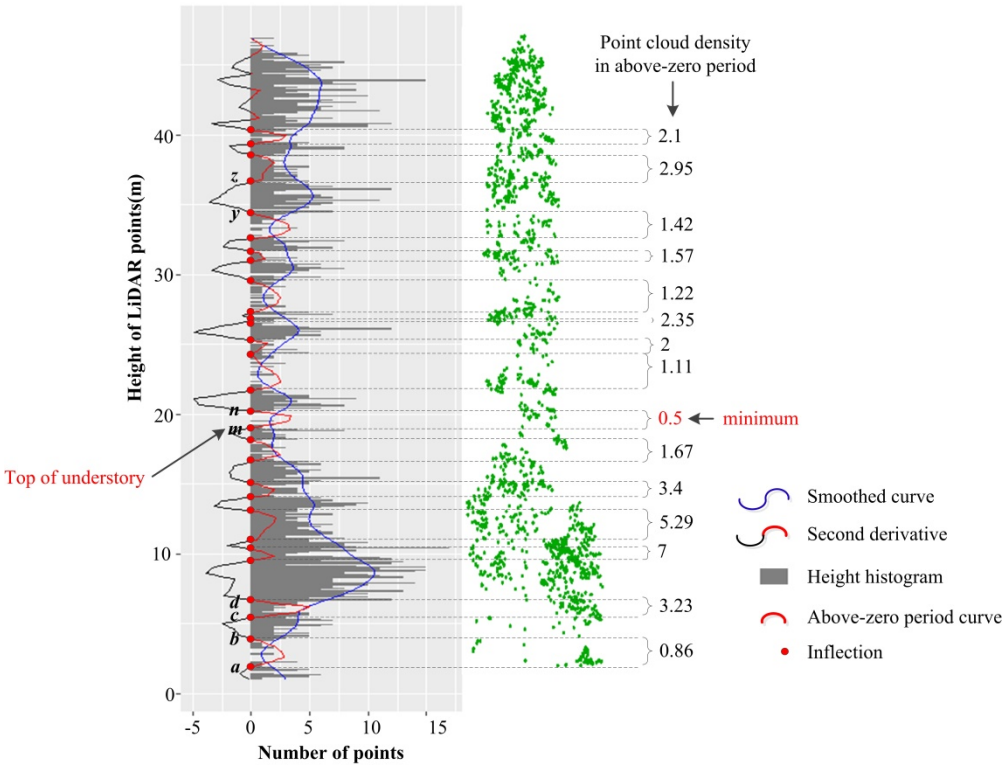

Fig. 4. An example of one individual tree, showing the process of locating the top of understory for subsequential removal.

2.4.3 CBH estimation

A percentile ranking based method was used to estimate the CBH as follows. We first generated a percentile ranking profile per tree using points without understory. This profile was fitted using the spline algorithm (the black curve in Fig. 5). The first derivative curve (the blue curve in Fig. 5) and the second derivative curve (the red curve in Fig. 5) were created, and then located the first derivative value at each inflection point (second derivative equals to zero, e.g. *a,b,c,…, j* in Fig. 5). We then selected the inflection point where the first derivative value is maximum (e.g. *f* in Fig. 5) and the percentile ranking value should be under 0.5 as the estimation of CBH. Figure 5 shows an example of estimating CBH based on the percentile ranking profile. The same procedure was applied to the entire LiDAR point cloud to calculate the CBH for each segmented tree. Noted that the first derivative reveals whether the percentile ranking is increasing or decreasing, and by how much it is increasing or decreasing. The second derivative indicates whether the first derivative is increasing or decreasing. If the second derivative is positive, the first derivative will be increasing (the fitted spline curve is concave). Conversely, if the second derivative is negative, the first derivative will be decreasing (the fitted spline curve is convex). Specially, if the second derivative is equal to zero, there would be an inflection at this point. The inflection occurs where there is an abrupt change on the percentile ranking. The expected inflection as CBH is the one where the first derivative value is maximum and the percentile ranking value should be less than 0.5. The second quartile height has been reported to be a good variable for estimating the forest biomass by [50]. Thus we considered that CBH should be lower than the second quartile height.

Optics EXPRESS

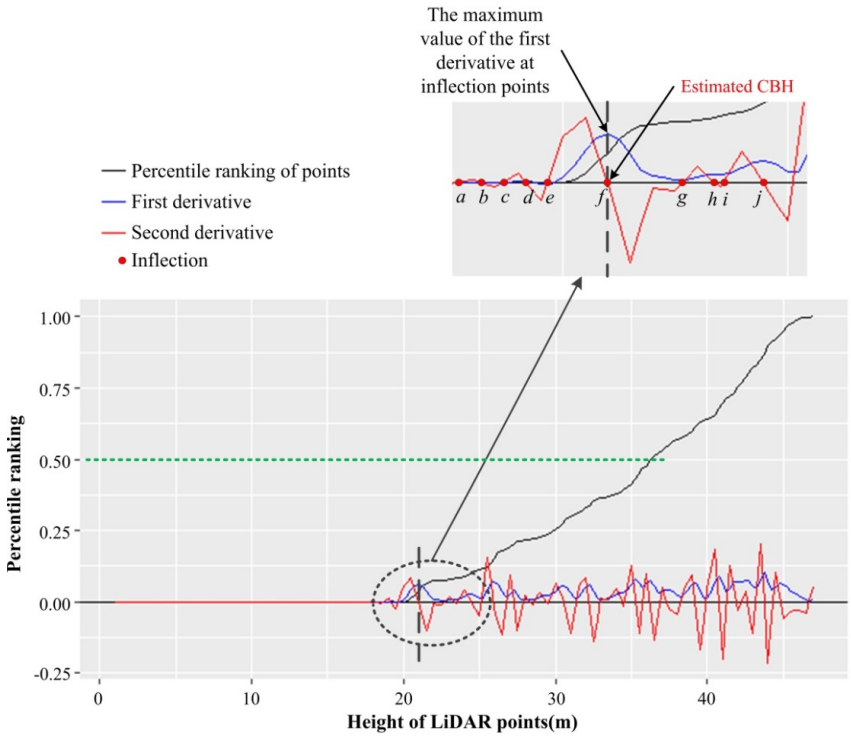

Fig. 5. An example of estimating CBH based on the percentile ranking profile of LiDAR points for an individual tree.

2.4.4 Tree-to-tree matching method

The matching between field measured and LiDAR-derived trees is the preliminary step for validation of the LiDAR estimation results. Tree parameters, such as tree location, tree height and tree crown size, were considered in the tree-to-tree matching procedure of this study. Tree location was considered as an important parameter of the matching in previous studies [2, 51, 52]. However, due to the uncertainties of GPS measurements under forest canopy and the bias of LiDAR-derived tree location, the matching would be difficult and problematic if only using tree location. Therefore, as shown in previous studies, a manual check was required to pair more trees [2]. To make the matching more automatic and to reduce the labor costs, tree height [51, 52] was also used to pair the trees. Moreover, by trial and error, we found that the diameter of tree crown was an appropriate threshold to refine the distance between paired trees in LiDAR estimates and field measurements. Thus, we designed a strategy which combined the tree location, tree height and crown size, to improve the accuracy and efficiency in tree-to-tree matching, as shown in an example below.

For each single tree measured in the field (a_i), we first selected the potential pairing trees b_j from the LiDAR-derived tree segments by their spatial location. Any LiDAR-derived trees with a distance [d_{ij} in Eq. (1)] smaller than the diameter of tree crown was marked as potential paired trees (b_j). Then, we further refined the selection using a comprehensive difference index [D_{ij} in Eq. (2)], which considered the location distance and height difference in a weighted combination. The weight for the height difference was set as 0.5, as suggested [52]. Finally, the tree with the smallest D_{ij} among all the potential LiDAR-derived pairing trees was matched to the field measured tree a_i. Any tree in the field measurements that failed to find potential pairing trees within crown diameter distance was labeled as unmatched and excluded from validation.

$$d_{ij} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \quad (1)$$

$$D_{ij} = d_{ij} + w \times \left(h_i - h_j\right) \quad (2)$$

where $(x_i,y_i)$is the location of a field measured tree a_i to be paired, (x_j,y_j) is the position of a LiDAR-derived tree b_j, d_{ij} denotes the distance between a_i and b_j, D_{ij} denotes the comprehensive difference index of tree matching, h_i denotes the height of a_i, h_j denotes the height of b_j, w is set to 0.5 as the weight of height difference [52].

2.4.5 Accuracy assessment of estimating CBH

After the CBHs of all the trees were estimated, the coefficient of determination (R^2) and root mean squared error (RMSE) [Eq. (3) and Eq. (4)] were used to assess the accuracy of CBH estimation. The bias [Eq. (5)] in LiDAR CBH estimation calculated as the percentage of difference between averaged LiDAR-derived and field measured CBH comparing to field measurement were also used to quantify the magnitude of over/under estimations.

$$R^2 = 1 - \sum_{i=1}^{n} \left(x_i - y_i\right)^2 / \sum_{i=1}^{n} \left(x_i - x_m\right)^2 \quad (3)$$

$$RMSE = \sqrt{\sum_{i=1}^{n} \left(x_i - y_i\right)^2 / n} \quad (4)$$

$$bias = \frac{y_m - x_m}{x_m} \times 100\% \quad (5)$$

where n represents the number of trees, x_i represents the ith CBH value of a field-measured tree, y_i represents the ith CBH value of a LiDAR-derived tree, x_m is the mean CBH value of all field-measured trees, y_m represents the mean CBH value of all LiDAR-derived trees.

3. Results

3.1 Individual tree segmentation

A total of 943 trees were automatically segmented from LiDAR point data in the study area. A visual examination of the preliminary segmentation results were conducted by three researchers with experiences in LiDAR remote sensing. By a majority vote, we observed 57 omission errors and 66 commission errors. The overall accuracy was 0.87, which was satisfactory comparing to previous studies conducted in similar forest conditions [41]. The under-segmented trees (omission errors) were mainly found in dense plots where some trees stood closely to each other and were mistakenly treated as a single tree by the algorithm [53]. To correct those errors, we manually separated the point clouds of under-segmented tree into two or three trees by visually examining the point clouds at different angles in the LiDAR360 software. Finally, we manually segmented the 57 under-segmented trees into 143 trees and generated a total of 1,029 trees for CBH estimation.

3.2 Field measurement results and tree-to-tree matching

A total of 637 trees were measured in the field, and their tree height, DBH, CBH, and tree location were recorded. It should be noted that the field survey was conducted in a quick plot-driving strategy, and only trees with a DBH larger than 15 cm were measured. Therefore, the number of measured trees was smaller than the actual number of trees as well as the number derived from LiDAR data. Among the 637 measured trees, 381 of them could not find a matched tree from LiDAR-data within the crown distance. Thus, only 256 trees in the field measurements were successfully paired with LiDAR-derived ones. The matching accuracy is

Optics EXPRESS

likely influenced by GPS errors and the difference in the definition of tree location between these two data sources. For example, the tree location was determined by the tree top in the LiDAR data, whereas it was recorded as the base of tree trunk location in the field measurements. Therefore, the deviation between tree top and base of tree trunk can also lead to the mismatch. Overall, there were 40.2% of field-measured trees and 24.9% of LiDAR-derived trees were matched successfully in this study.

Table 1 shows the summary of field measurements and tree-to-tree matching results. For field-measured trees, the average CBH was 7.63 m with a standard deviation of 4.51 m, the average tree height was 23.97 m with a standard deviation of 9.31 m and the average DBH was 44.22 cm with a standard deviation of 21.38 cm. For LiDAR-derived trees, the average crown diameter was 8.98 m with a standard deviation of 3.14 m; the average tree height was 22.35 m with a standard deviation of 9.29 m. The average height of LiDAR-derived trees was 1.62 m (6.76%) lower than that of field-measured trees, which indicated that LiDAR-derived tree height was slightly lower than the field-measured ones likely for two reasons: 1) it is often difficult for LiDAR laser beam to hit the exact tops of trees [51], and 2) it is often difficult to measure tree height accurately in the field [54].

Table 2 shows the summary of the field measurements for matched trees. The amount of ABCO, ABMA, CADE, PILA, PIPO, PSME, QUKE is 102 (39.84%), 9 (3.52%), 20 (7.81%), 33 (12.89%), 47 (18.36%), 41 (16.02%) and 4 (1.56%), respectively. Mean tree heights across tree species varied from less than 7 m for PILA to over 27 m for ABMA. Meanwhile, mean DBHs varied from less than 16 cm for PILA to over 55 cm for QUKE.

Table 1. Summary of the field measurements and LiDAR-derived tree structure parameters for the matched trees.

	Min.	Max.	Range	Average	Standard deviation
			Field data		
DBH (cm)	19.8	144.4	124.6	44.22	21.38
Height (m)	8.19	55.75	47.56	23.97	9.31
CBH (m)	1.4	24.1	22.7	7.63	4.51
			LiDAR-derived data		
Crown Diameter (m)	3.05	18.53	15.48	8.98	3.14
Height (m)	7.86	52.5	44.64	22.35	9.29
CBH (m)	1.41	24.56	23.15	8	4.52

Table 2. Summary of the field measurements for the matched trees categorized by tree species. Tree height, CBH and DBH are presented as mean values ± standard deviation.

Abbr.	Tree species	Number of trees (percent)	Height(m)	CBH(m)	DBH(cm)
ABCO	White fir (Abies concolor)	102 (39.84%)	24.06 ± 9.32	7.95 ± 4.45	44.07 ± 21.11
ABMA	Red fir (Abies magnifica)	9 (3.52%)	27.44 ± 11.63	8.64 ± 5.73	51.5 ± 25.59
CADE	Incense cedar (Calocedrus decurrens)	20 (7.81%)	15.31 ± 5.92	5.08 ± 3.09	39.09 ± 20.05
PILA	Sugar pine (Pinus lambertiana)	33 (12.89%)	7.45 ± 9.47	4.5 ± 5.4	16.97 ± 20.86
PIPO	Ponderosa pine (Pinus ponderosa)	47 (18.36%)	22.43 ± 10.72	6.91 ± 4.64	45.26 ± 23.96
PSME	Douglas-fir (Pseudotsuga menziesii)	41 (16.02%)	23.26 ± 6.91	7.45 ± 3.41	40.75 ± 16.58
QUKE	Black oak (Quercus kelloggii)	4 (1.56%)	15.78 ± 5.85	4.6 ± 0.63	55.15 ± 40.72

Optics EXPRESS

3.3 CBH estimation and accuracy assessment

The comparison of CBH estimates between LiDAR-derived data and field measurements are shown in Table 1. The minimum and maximum of LiDAR-derived CBH values were 1.41 m and 24.56 m, respectively, with a range of 23.15 m. The average CBH value was 8 m with a standard deviation of 4.52 m. The RMSE, R^2, and bias were 0.88, 1.62 and 3.36%, respectively, which indicated that LiDAR estimated CBH matched well with field-measured results, albeit with a slight overestimation (Fig. 6).

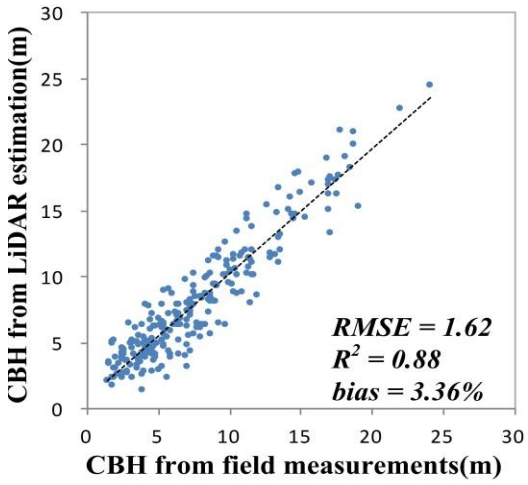

Fig. 6. Comparison of CBH from field measurements and the proposed LiDAR-based method.

The accuracies of LiDAR-based CBH estimates were among four tree height groups (<15m, 15 to 25m, 25 to 25m, and >35m) are presented in Fig. 7, respectively. Results revealed that the RMSEs of groups where tree height below 35 m were nearly the same, slightly varied from 1.55 m to 1.59 m. Yet the group of trees with the height over 35 m showed a larger RMSE at 2 m.

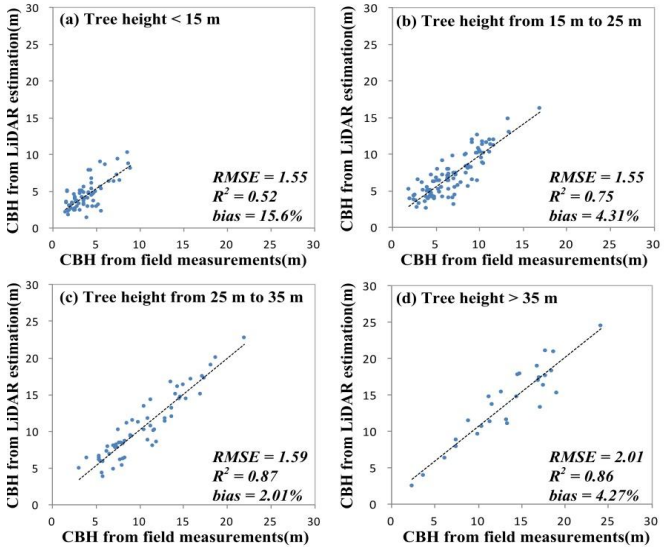

Fig. 7. The CBH from LiDAR estimates compared to the CBH from field measurements, for different tree height groups.

Optics EXPRESS

Figure 8 shows the accuracies among different species. The R^2 in CBH estimation ranged from 0.82 to 0.93, which indicates that the LiDAR-derived results can explain the majority of the variation (>80%) in field-measured CBH regardless of the species type. The RMSE varied substantially among species, with the largest RMSE value in ABMA at 2.31 and lowest in QUKE at 0.52. Accounting for differences in the average CBH among species, we found that the highest bias in CADE at 26.9%, and lowest in PIPO at 1.3%.

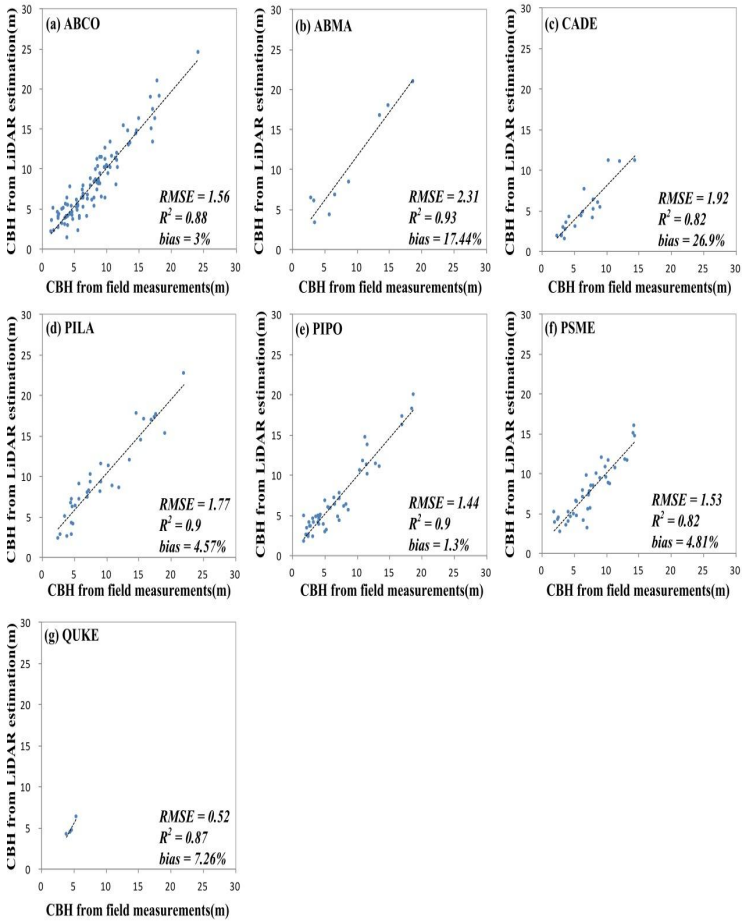

Fig. 8. CBH estimates from LiDAR compared to those from field measurements for seven tree species (i.e., ABCO, ABMA, CADE, PILA, PIPO, PSME, and QUKE).

4. Discussion

Estimating CBH was the core objective of this study. The accuracy of our LiDAR-based CBH estimation is competitive when compared to previous LiDAR-based studies [2, 3, 38, 39]. Our results show that the proposed method in this study can estimate CBH at tree level more accurately in mixed-species forests as compared to previous studies (Table 3). For all trees, including five conifer species and two broadleaf species, we achieved an RMSE of around 1.6 m. In a similar study by Popescu and Zhao (2008) which estimated CBH of both deciduous and pine trees, the averaged RMSE was approximately 2 m [2]. Vauhkonen (2010) gained an RMSE value ranging from 1.44 m to 3.56 m, but only for Scots pine [3]. In addition, in this study the coefficients of determination (R^2 = 0.88) for CBH between field-measured trees and LiDAR-derived trees was slightly higher than those from Popescu et al. [2] (e.g. R^2 of 0.75) and Vauhkonen [3] (e.g. R^2 of 0.8). Overall, the results indicate that our method performed

equally or slightly better than other LiDAR-based methods and also had a wider application in mixed-species forest for individual tree CBH estimation.

Table 3. Summary of results from previous CBH estimation method comparing with our method for individual trees

Method	Reference	Forest type	Understory removal	Sampling data	R^2	RMSE (m)
Vertical profile and regression method	Solberg et al. (2006)	Spruce and deciduous trees	No	Yes	NM[a]	3.5
Vertical profile on voxel	Popescu and Zhao (2008)	Pine and deciduous trees	No	No	Around 0.75	Around 2
The 3D geometry	Vauhkonen (2010)	Scots pine	No	No	0.71-0.84	1.44-3.56
Detecting inflection based on percentile ranking files	This study	Mixed-species including seven species	Yes	No	0.88	1.62

[a] NM denotes not mentioned

The strong flexibility of our method was partly due to the fact that it estimates CBH directly from airborne LiDAR data with no requirement for field measurements. Studies such as [38,39] relied heavily on filed measurements to calculate CBH, and consequently constrained the transferability of the method across species. The direct method proposed in our study, on the other hand, can be used to multiple tree species, and thus has the potential for large-scale applications. The other advantage of this method lies in the removal of understory noise from LiDAR points prior to CBH estimation. The existence of shrubs and small trees in the understory is very common in many forests, but often neglected in CBH estimation. Popescu and Zhao [2] mentioned that the understory vegetation can make the vertical profile of LiDAR point density too noisy to detect CBH accurately. Moreover, the usage of spline curve in CBH estimation also contributed to the robustness of our method. Spline curve function generates a special piecewise curve, which is a simple but very useful tool to identify inflection points [55]. In this study, the spline curve was used to detect the abrupt changes in LiDAR point density in the vertical direction, which has been demonstrated to be helpful for both understory tree top detection and CBH determination. Overall, the three features of this proposed method which include the independence of prior knowledge, the removal of understory noise, as well as the usage of spline curve have contributed to its flexibility, accuracy, and efficiency in LiDAR-based CBH estimation. It also demonstrated strong capability to apply over large areas for mixed-species forests.

To investigate the influences of the tree biophysical parameters on LiDAR-based CBH estimation, we systematically analyzed the accuracies among different height groups (Fig. 7) and tree species (Fig. 8) by comparing with field measurements. Our method would be slightly affected as the tree height exceeds or under a certain value. A possible reason for the larger uncertainty is that taller tree has longer crown length which could make the laser pulses difficult to reach the base of the tree crown. Results also indicated that bias in the group of tree height below 15 m was 15.6% which was a little higher than that of the other three groups. The reason could be that there were more noises which were difficult to be detected by our method under the crown of shorter trees. With respect to the influence of tree species, we found that the accuracies varied substantially among species. Similar results were reported by [2], which mentioned that the correlation between field-measurement data and LiDAR-derived data for deciduous trees is lower than that for pines.

Although our method performed well and estimated CBH efficiently and precisely, it has some uncertainties. The first uncertainty was caused by individual tree segmentation, which is an indispensable step of LiDAR-based forest structural parameter estimation at individual tree level [53]. The results of tree segmentation inevitably influence the accuracy of the CBH estimation. We selected the top-to-bottom region-growing method developed by Li et al. (2012), owing to its excellent performance and comprehensibility in point cloud based tree segmentation [41]. Our results showed that 91% of trees were detected and 87% of them were segmented correctly. These results are comparable to the results reported in the study of Li et al. [41], which reported the accuracy of tree segmentation at 94%. Tree segmentation methods will perform differently in forests with different densities and in mixed species as well as using LiDAR data with different qualities and settings. Some researchers also indicated that the accuracy of tree segmentation over 80% was feasible for the forest applications [53]. We also noticed that many new methods have been developed to improve the tree segmentation, but evaluating their uncertainties is not the focus of this study. Therefore, instead of testing more sophisticated segmentation method, we manually corrected the under-segmented trees as suggested in previous studies [2]. The errors in tree-to-tree matching between LiDAR-derived and field measured trees may also impact the reliability of validation. These biases were mainly caused by uncertainties in GPS location and tree height estimation. In order to reduce these impacts on validation, we designed a comprehensive index to find the best match and excluded samples with large location bias [2, 8]. Similar strategies were also used in previous studies, Popescu and Zhao (2008) only used 12% of field measurement to reduce tree-to-tree matching uncertainties. Lastly, the method proposed in this study also has some limitations. To better delineate the vertical profile and detect the inflections, our method requires LiDAR data with relatively high point density. Further research is necessary to discuss the effect of point cloud density on the accuracy of the CBH estimation using similar method. In addition, our method was based on spline curve, besides its simplicity compared with other sophisticated method [2, 41, 42], and it may have the problem of overfitting. Overall, this method performed well in the Sierra Nevada mixed-conifer dominated forests with the RSME less than 2 m. More studies are needed to evaluate the accuracy of this method over other forests with different densities and species.

5. Conclusions

This study developed and tested a new CBH estimation method, which includes two important strategies. The first is the removal of noise and understory vegetation for each tree, in essence to clean up the vertical profile of a tree and make accurate CBH estimation possible. The second is estimating CBH by generating a percentile ranking profile of LiDAR points for each tree and using a spline curve to identify the inflection points based on the generated percentile ranking profile. The method proposed in this study was applied in a mixed-conifer dominated forest in Sierra Nevada, California, USA. Satisfying results have been achieved using this method with a slight overestimation of 3.36%, and the RMSE and R^2 value for all the trees were 1.62 m and 0.88, respectively. Accuracy for CBH estimation was the highest for black oak (QUKE) trees, with an RMSE of 0.52 m. The conifer species results were also good with uniformly high R^2. For the most abundant conifers in the forest (i.e. ABCO, PIPO, PSME) results were good: ABCO (RMSE = 1.56), PIPO (RMSE = 1.44) and PSME (RMSE = 1.53). In general, taller trees have relatively higher uncertainties than shorter trees. Overall, this direct LiDAR-based method proposed in this study has demonstrated high accuracy for individual tree CBH estimation with no requirement of prior knowledge, and strong potential for applications in mixed species over large areas.

Funding

National Key R&D Program of China (2017YFC0503905, 2016YFC0500202), the Frontier Science Key Programs of the Chinese Academy of Sciences (QYZDY-SSW-SMC011);

Optics EXPRESS

National Natural Science Foundation of China (Grant No. 41471363, 31270563, 41471363, and 31741016); Beijing Natural Science Foundation (Grant No.4164085), Beijing Postdoctoral Research Foundation, Beijing Municipal Education Commission Research Program (Grant No.KM201511418002); and the open funding project of State Key Laboratory of Virtual Reality Technology and Systems (Grant No.011177220010020).

Acknowledgments

We acknowledge the contribution of the Sierra Nevada Adaptive Management Project, an interagency project supported by the USDA Forest Service Region 5, USDA Forest Service Pacific Southwest Research Station, US Fish and Wildlife Service, California Department of Water Resources, California Department of Fish and Wildlife, California Department of Forestry and Fire Protection, and the Sierra Nevada Conservancy.