

Effects of LiDAR point density, sampling size and height threshold on estimation accuracy of crop biophysical parameters

Shezhou Luo,^{1,2} Jing M. Chen,² Cheng Wang,^{1,*} Xiaohuan Xi,¹ Hongcheng Zeng,³
Dailiang Peng,¹ and Dong Li¹

¹Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China

²Department of Geography and Program in Planning, University of Toronto, Toronto, ON M5S 3G3, Canada

³Faculty of Forestry, University of Toronto, 33 Willcocks Street, Toronto, ON M5S 3B3, Canada
[*wangcheng@radi.ac.cn](mailto:wangcheng@radi.ac.cn)

Abstract: Vegetation leaf area index (LAI), height, and aboveground biomass are key biophysical parameters. Corn is an important and globally distributed crop, and reliable estimations of these parameters are essential for corn yield forecasting, health monitoring and ecosystem modeling. Light Detection and Ranging (LiDAR) is considered an effective technology for estimating vegetation biophysical parameters. However, the estimation accuracies of these parameters are affected by multiple factors. In this study, we first estimated corn LAI, height and biomass ($R^2 = 0.80, 0.874$ and 0.838 , respectively) using the original LiDAR data (7.32 points/m^2), and the results showed that LiDAR data could accurately estimate these biophysical parameters. Second, comprehensive research was conducted on the effects of LiDAR point density, sampling size and height threshold on the estimation accuracy of LAI, height and biomass. Our findings indicated that LiDAR point density had an important effect on the estimation accuracy for vegetation biophysical parameters, however, high point density did not always produce highly accurate estimates, and reduced point density could deliver reasonable estimation results. Furthermore, the results showed that sampling size and height threshold were additional key factors that affect the estimation accuracy of biophysical parameters. Therefore, the optimal sampling size and the height threshold should be determined to improve the estimation accuracy of biophysical parameters. Our results also implied that a higher LiDAR point density, larger sampling size and height threshold were required to obtain accurate corn LAI estimation when compared with height and biomass estimations. In general, our results provide valuable guidance for LiDAR data acquisition and estimation of vegetation biophysical parameters using LiDAR data.

©2016 Optical Society of America

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

References and links

1. A. A. Gitelson, A. Viña, T. J. Arkebauer, D. C. Rundquist, G. Keydan, and B. Leavitt, "Remote estimation of leaf area index and green leaf biomass in maize canopies," *Geophys. Res. Lett.* **30**(5), 1248 (2003).
2. M. Chopping, G. G. Moisen, L. Su, A. Laliberte, A. Rango, J. V. Martonchik, and D. P. C. Peters, "Large area mapping of southwestern forest crown cover, canopy height, and biomass using the NASA multiangle imaging spectro-radiometer," *Remote Sens. Environ.* **112**(5), 2051–2063 (2008).
3. J. Heiskanen, "Estimating aboveground tree biomass and leaf area index in a mountain birch forest using ASTER satellite data," *Int. J. Remote Sens.* **27**(6), 1135–1158 (2006).
4. L. Chasmer, C. Hopkinson, P. Treitz, H. McCaughey, A. Barr, and A. Black, "A lidar-based hierarchical approach for assessing MODIS fPAR," *Remote Sens. Environ.* **112**(12), 4344–4357 (2008).
5. I. Jonckheere, S. Fleck, K. Nackaerts, B. Muys, P. Coppin, M. Weiss, and F. Baret, "Review of methods for in situ leaf area index determination. Part I. Theories, sensors and hemispherical photography," *Agric. For. Meteorol.* **121**(1–2), 19–35 (2004).

6. Q. Chen, G. Vaglio Laurin, J. J. Battles, and D. Saah, "Integration of airborne lidar and vegetation types derived from aerial photography for mapping aboveground live biomass," *Remote Sens. Environ.* **121**, 108–117 (2012).
7. J. M. Chen and T. A. Black, "Measuring leaf area index of plant canopies with branch architecture," *Agric. For. Meteorol.* **57**(1–3), 1–12 (1991).
8. J. M. Chen, A. Govind, O. Sonnentag, Y. Zhang, A. Barr, and B. Amiro, "Leaf area index measurements at Fluxnet-Canada forest sites," *Agric. For. Meteorol.* **140**(1–4), 257–268 (2006).
9. S. C. Popescu, K. Zhao, A. Neuschwander, and C. Lin, "Satellite lidar vs. small footprint airborne lidar: Comparing the accuracy of aboveground biomass estimates and forest structure metrics at footprint level," *Remote Sens. Environ.* **115**(11), 2786–2797 (2011).
10. J. B. Drake, R. G. Knox, R. O. Dubayah, D. B. Clark, R. Condit, J. B. Blair, and M. Hofton, "Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships," *Glob. Ecol. Biogeogr.* **12**(2), 147–159 (2003).
11. M. A. Lefsky, D. P. Turner, M. Guzy, and W. B. Cohen, "Combining lidar estimates of aboveground biomass and Landsat estimates of stand age for spatially extensive validation of modeled forest productivity," *Remote Sens. Environ.* **95**(4), 549–558 (2005).
12. R. W. Kulawardhana, S. C. Popescu, and R. A. Feagin, "Fusion of lidar and multispectral data to quantify salt marsh carbon stocks," *Remote Sens. Environ.* **154**, 345–357 (2014).
13. N. Zhang, M. Wang, and N. Wang, "Precision agriculture—a worldwide overview," *Comput. Electron. Agric.* **36**(2–3), 113–132 (2002).
14. S. Enghart, V. Keuck, and F. Siegert, "Aboveground biomass retrieval in tropical forests — the potential of combined X- and L-band SAR data use," *Remote Sens. Environ.* **115**(5), 1260–1271 (2011).
15. A. Peduzzi, R. H. Wynne, T. R. Fox, R. F. Nelson, and V. A. Thomas, "Estimating leaf area index in intensively managed pine plantations using airborne laser scanner data," *For. Ecol. Manage.* **270**, 54–65 (2012).
16. P. Hyde, R. Dubayah, B. Peterson, J. Blair, M. Hofton, C. Hunsaker, R. Knox, and W. Walker, "Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems," *Remote Sens. Environ.* **96**(3–4), 427–437 (2005).
17. F. Zhao, Q. Guo, and M. Kelly, "Allometric equation choice impacts lidar-based forest biomass estimates: a case study from the Sierra National Forest, CA," *Agric. For. Meteorol.* **165**, 64–72 (2012).
18. J. M. Chen and J. Cihlar, "Retrieving leaf area index of boreal conifer forests using Landsat TM images," *Remote Sens. Environ.* **55**(2), 153–162 (1996).
19. C. Atzberger, "Advances in remote sensing of agriculture: context description, existing operational monitoring systems and major information needs," *Remote Sens.* **5**(2), 949–981 (2013).
20. G. M. Foody, D. S. Boyd, and M. E. J. Cutler, "Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions," *Remote Sens. Environ.* **85**(4), 463–474 (2003).
21. A. C. Morel, S. S. Saatchi, Y. Malhi, N. J. Berry, L. Banin, D. Burslem, R. Nilus, and R. C. Ong, "Estimating aboveground biomass in forest and oil palm plantation in Sabah, Malaysian Borneo using ALOS PALSAR data," *For. Ecol. Manage.* **262**(9), 1786–1798 (2011).
22. S. Gao, Z. Niu, N. Huang, and X. Hou, "Estimating the leaf area index, height and biomass of maize using HJ-1 and RADARSAT-2," *Int. J. Appl. Earth Obs. Geoinf.* **24**, 1–8 (2013).
23. J. J. Richardson, L. M. Moskal, and S.-H. Kim, "Modeling approaches to estimate effective leaf area index from aerial discrete-return LIDAR," *Agric. For. Meteorol.* **149**(6–7), 1152–1160 (2009).
24. S. Luo, C. Wang, X. Xi, S. Nie, S. Xia, and W. Yiping, "Forest leaf area index estimation using combined ICESat/GLAS and optical remote sensing image," *J. Inf. Millim. Waves* **34**(2), 243–249 (2015).
25. Y. Qin, S. Li, T.-T. Vu, Z. Niu, and Y. Ban, "Synergistic application of geometric and radiometric features of LIDAR data for urban land cover mapping," *Opt. Express* **23**(11), 13761–13775 (2015).
26. V. Kankare, M. Vastaranta, M. Holopainen, M. Rätty, X. Yu, J. Hyyppä, H. Hyyppä, P. Alho, and R. Viitala, "Retrieval of forest aboveground biomass and stem volume with airborne scanning LiDAR," *Remote Sens.* **5**(5), 2257–2274 (2013).
27. T. Hakala, J. Suomalainen, S. Kaasalainen, and Y. Chen, "Full waveform hyperspectral LiDAR for terrestrial laser scanning," *Opt. Express* **20**(7), 7119–7127 (2012).
28. K. Zhao and S. Popescu, "Lidar-based mapping of leaf area index and its use for validating GLOBECARBON satellite LAI product in a temperate forest of the southern USA," *Remote Sens. Environ.* **113**(8), 1628–1645 (2009).
29. D. Riaño, F. Valladares, S. Condés, and E. Chuvieco, "Estimation of leaf area index and covered ground from airborne laser scanner (Lidar) in two contrasting forests," *Agric. For. Meteorol.* **124**(3–4), 269–275 (2004).
30. J. Jensen, K. Humes, L. Vierling, and A. Hudak, "Discrete return lidar-based prediction of leaf area index in two conifer forests," *Remote Sens. Environ.* **112**(10), 3947–3957 (2008).
31. S. Solberg, "Mapping gap fraction, LAI and defoliation using various ALS penetration variables," *Int. J. Remote Sens.* **31**(5), 1227–1244 (2010).
32. G. Pope and P. Treitz, "Leaf area index (LAI) Estimation in boreal mixedwood forest of Ontario, Canada using light detection and ranging (LiDAR) and worldview-2 imagery," *Remote Sens.* **5**(10), 5040–5063 (2013).
33. N. F. Glenn, L. P. Spaete, T. T. Sankey, D. R. Derryberry, S. P. Hardegree, and J. J. Mitchell, "Errors in LiDAR-derived shrub height and crown area on sloped terrain," *J. Arid Environ.* **75**(4), 377–382 (2011).
34. C. Edson and M. G. Wing, "Airborne light detection and ranging (LiDAR) for individual tree stem location, height, and biomass measurements," *Remote Sens.* **3**(12), 2494–2528 (2011).
35. G. W. Frazer, S. Magnussen, M. A. Wulder, and K. O. Niemann, "Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass," *Remote*

- Sens. Environ. **115**(2), 636–649 (2011).
36. O. W. Tsui, N. C. Coops, M. A. Wulder, P. L. Marshall, and A. McCardle, “Using multi-frequency radar and discrete-return LiDAR measurements to estimate above-ground biomass and biomass components in a coastal temperate forest,” *ISPRS J. Photogramm. Remote Sens.* **69**, 121–133 (2012).
 37. J. U. H. Eitel, T. S. Magney, L. A. Vierling, T. T. Brown, and D. R. Huggins, “LiDAR based biomass and crop nitrogen estimates for rapid, non-destructive assessment of wheat nitrogen status,” *Field Crops Res.* **159**, 21–32 (2014).
 38. L. Zhang and T. E. Grift, “A LIDAR-based crop height measurement system for *Miscanthus giganteus*,” *Comput. Electron. Agric.* **85**, 70–76 (2012).
 39. N. Tilly, D. Hoffmeister, Q. Cao, S. Huang, V. Lenz-Wiedemann, Y. Miao, and G. Bareth, “Multitemporal crop surface models: accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice,” *J. Appl. Remote Sens.* **8**(1), 083671 (2014).
 40. I. J. Davenport, R. B. Bradbury, G. Q. A. Anderson, G. R. F. Hayman, J. R. Krebs, D. C. Mason, J. D. Wilson, and N. J. Veck, “Improving bird population models using airborne remote sensing,” *Int. J. Remote Sens.* **21**(13–14), 2705–2717 (2000).
 41. W. Li, Z. Niu, N. Huang, C. Wang, S. Gao, and C. Wu, “Airborne LiDAR technique for estimating biomass components of maize: A case study in Zhangye City, Northwest China,” *Ecol. Indic.* **57**, 486–496 (2015).
 42. C. Wang, M. Menenti, M. P. Stoll, A. Feola, E. Belluco, and M. Marani, “Separation of ground and low vegetation signatures in LiDAR measurements of salt-marsh environments,” *IEEE Trans. Geosci. Rem. Sens.* **47**(7), 2014–2023 (2009).
 43. C. Hopkinson, L. E. Chasmer, G. Sass, I. F. Creed, M. Sitar, W. Kalbfleisch, and P. Treitz, “Vegetation class dependent errors in lidar ground elevation and canopy height estimates in a boreal wetland environment,” *Can. J. Rem. Sens.* **31**(2), 191–206 (2005).
 44. M. K. Jakubowski, Q. Guo, and M. Kelly, “Tradeoffs between lidar pulse density and forest measurement accuracy,” *Remote Sens. Environ.* **130**, 245–253 (2013).
 45. M. Watt, A. Meredith, P. Watt, and A. Gunn, “The influence of LiDAR pulse density on the precision of inventory metrics in young unthinned Douglas-fir stands during initial and subsequent LiDAR acquisitions,” *N. Z. J. For. Sci.* **44**(1), 1–9 (2014).
 46. L. Ruiz, T. Hermosilla, F. Mauro, and M. Godino, “Analysis of the influence of plot size and LiDAR density on forest structure attribute estimates,” *Forests* **5**(5), 936–951 (2014).
 47. T. Gobakken and E. Næsset, “Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data,” *Can. J. For. Res.* **38**(5), 1095–1109 (2008).
 48. M. Magnusson, J. E. S. Fransson, and J. Holmgren, “Effects on estimation accuracy of forest variables using different pulse density of laser data,” *For. Sci.* **53**(6), 619–626 (2007).
 49. J. Strunk, H. Temesgen, H.-E. Andersen, J. P. Flewelling, and L. Madsen, “Effects of lidar pulse density and sample size on a model-assisted approach to estimate forest inventory variables,” *Can. J. Rem. Sens.* **38**(5), 644–654 (2012).
 50. P. Wilkes, S. D. Jones, L. Suarez, A. Haywood, W. Woodgate, M. Soto-Berelev, A. Mellor, and A. K. Skidmore, “Understanding the effects of ALS pulse density for metric retrieval across diverse forest types,” *Photogramm. Eng. Remote Sensing* **81**(8), 625–635 (2015).
 51. K. K. Singh, G. Chen, J. B. McCarter, and R. K. Meentemeyer, “Effects of LiDAR point density and landscape context on estimates of urban forest biomass,” *ISPRS J. Photogramm. Remote Sens.* **101**, 310–322 (2015).
 52. J. Estornell, L. A. Ruiz, B. Velázquez-Martí, and A. Fernández-Sarriá, “Estimation of shrub biomass by airborne LiDAR data in small forest stands,” *For. Ecol. Manage.* **262**(9), 1697–1703 (2011).
 53. X. Li, G. Cheng, S. Liu, Q. Xiao, M. Ma, R. Jin, T. Che, Q. Liu, W. Wang, Y. Qi, J. Wen, H. Li, G. Zhu, J. Guo, Y. Ran, S. Wang, Z. Zhu, J. Zhou, X. Hu, and Z. Xu, “Heihe watershed allied telemetry experimental research (HiWATER): scientific objectives and experimental design,” *Bull. Am. Meteorol. Soc.* **94**(8), 1145–1160 (2013).
 54. Q. Xiao and J. Wen, “HiWATER: Airborne LiDAR-DEM&DSM data production in the middle reaches of the Heihe River Basin,” (Heihe Plan Science Data Center, 2013).
 55. S. Luo, C. Wang, X. Xi, and F. Pan, “Estimating FPAR of maize canopy using airborne discrete-return LiDAR data,” *Opt. Express* **22**(5), 5106–5117 (2014).
 56. A. Jochem, M. Hollaus, M. Rutzinger, and B. Höfle, “Estimation of aboveground biomass in alpine forests: a semi-empirical approach considering canopy transparency derived from airborne LiDAR data,” *Sensors* **11**(1), 278–295 (2010).
 57. S. Luo, C. Wang, F. Pan, X. Xi, G. Li, S. Nie, and S. Xia, “Estimation of wetland vegetation height and leaf area index using airborne laser scanning data,” *Ecol. Indic.* **48**, 550–559 (2015).
 58. E. Næsset and T. Gobakken, “Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser,” *Remote Sens. Environ.* **112**(6), 3079–3090 (2008).
 59. J. Chen, S. Gu, M. Shen, Y. Tang, and B. Matsushita, “Estimating aboveground biomass of grassland having a high canopy cover: an exploratory analysis of in situ hyperspectral data,” *Int. J. Remote Sens.* **30**(24), 6497–6517 (2009).
 60. G. W. Sileshi, “A critical review of forest biomass estimation models, common mistakes and corrective measures,” *For. Ecol. Manage.* **329**, 237–254 (2014).
 61. E. Luedeling and A. Gassner, “Partial least squares regression for analyzing walnut phenology in California,” *Agric. For. Meteorol.* **158–159**, 43–52 (2012).
 62. S. Wold, M. Sjöström, and L. Eriksson, “PLS-regression: a basic tool of chemometrics,” *Chemom. Intell. Lab.*

- Syst. **58**(2), 109–130 (2001).
63. J. Anderson, L. Plourde, M. Martin, B. Braswell, M. Smith, R. Dubayah, M. Hofton, and J. Blair, “Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest,” *Remote Sens. Environ.* **112**(4), 1856–1870 (2008).
-

1. Introduction

Vegetation height, leaf area index (LAI) and aboveground biomass (AGB) are key biophysical parameters for vegetation growth, yield forecasting, health monitoring, climate change and ecosystem modeling [1–3]. The accuracy of these models primarily depends on the accuracy of the key model input parameters [4]. Hence, accurate and reliable estimations of vegetation height, LAI and biomass are highly important [4–6]. LAI is defined as one half of the total leaf area per unit ground surface area [7], and LAI is an important structural parameter of vegetation ecosystems that controls many biological and physical processes, such as photosynthesis, respiration, transpiration and rainfall interception [8]. Aboveground biomass of vegetation is the total dry weight of living organic matter per unit area above the ground surface [9,10]. Vegetation biomass plays a critical role in global change and carbon cycle modeling and has been extensively applied to estimate vegetation productivity and terrestrial carbon stocks [11,12]. Crop height, LAI and biomass provide useful information that assists farmers in decision-making for irrigation, fertilizer application, and insect and weed infestation control [13]. Corn is one of the most important crops and is wide-spread in China. Therefore, the estimations of corn height, LAI and biomass have great significance in corn growth monitoring, yield forecasting and ecological modeling.

Traditional direct measurements are the reliable methods for obtaining vegetation parameters (such as height, LAI and biomass) [5,14,15]. However, direct methods are labor-intensive and time-consuming and frequently require destructive sampling [16,17]. Thus, direct methods are often costly and infeasible for large-area field measurements. Remote sensing technology provides a promising method for acquiring vegetation parameters over large areas [18] and is widely used in agriculture [19]. Using passive optical remotely sensed or radio detection and ranging (Radar) data, numerous investigations have been conducted on estimations of vegetation LAI [3,18], vegetation biomass [14,20,21] and vegetation height [2,22]. These vegetation parameters have been estimated primarily through empirical relationships between field observations and vegetation indices (VIs) calculated from remotely sensed data, such as the normalized difference vegetation index (NDVI), simple ratio (SR) and enhanced vegetation index (EVI). However, traditional remote sensing data encounter limitations in estimation of vegetation parameters due to the saturation problem of VIs, which is common in dense vegetation cover or high LAI areas [23,24]. Therefore, accurate estimation of vegetation height, LAI and biomass is a challenging task with passive optical remotely sensed data.

Light Detection and Ranging (LiDAR) is an active remote sensing technology [25]. Laser pulses emitted from LiDAR systems are able to penetrate into the vegetation canopy and acquire accurate three-dimensional structural information for the vegetation canopy. Therefore, LiDAR systems offer great potential for estimating vegetation parameters with high accuracy [26,27]. Many previous studies have shown that airborne LiDAR data are a highly effective data resource for estimating forest biophysical parameters and can reliably estimate forest LAI [23,28–32], height [33,34] and biomass [35,36]. Some research on estimating crop parameter was carried out using terrestrial laser scanning (TLS) [37–39], however, only a few studies have been conducted on estimations of crop biophysical parameters using airborne discrete-return LiDAR data [40,41]. For airborne discrete-return LiDAR, this scarcity can be attributed to three major factors. First, corn crops are quite short in height relative to forest vegetation. Because of limitations in the LiDAR techniques, the time-of-flight between the first and the last return is too short to detect. As a consequence, the elevation of the canopy and the ground cannot be accurately determined [42]. Second, crop vegetation is usually dense and uniform, and laser pulses have difficulty penetrating the canopy to the ground. The accuracies of DEMs (digital elevation models), DSMs (digital surface models) and LiDAR-derived vegetation parameters are low for short vegetation areas

[43]. Thus, LiDAR data with high accuracy and high laser/LiDAR point density are essential for acquiring reliable vegetation parameters and terrain characteristics. Because high point density means decreasing the flight altitude or speed or increasing the flight overlap, these changes will increase the flight time [44,45]. Thus, high point density incurs high LiDAR data acquisition costs, especially over large study areas. Therefore, knowledge of the appropriate LiDAR point density is important to make informed tradeoffs between the costs of data acquisition and the accuracy of information derived from LiDAR data. Certain previous researchers have investigated the effect of airborne LiDAR point density on estimates of forest vegetation parameters [45–50]. Jakubowski et al. [44] studied the effects of 17 different LiDAR point densities on forest structural parameters, and the results showed that reduced LiDAR point density could reliably estimate forest structural parameters. Similarly, Singh et al. [51] found that reduced point density could estimate biomass without compromising accuracy, and this observation is especially significant for reducing research costs over large area forest studies. However, to the best of our knowledge, no study has been conducted to explore the effects of airborne LiDAR point density on estimates of crop parameters. In addition to LiDAR point density, the sampling size of LiDAR data and the height threshold for separating ground returns and vegetation returns are also key factors that affect the estimation accuracy of crop biophysical parameters. Similarly, these two factors have been investigated for estimation of forest biophysical parameters [23,28,52]. However, we do not know whether the effects of sampling size and height threshold should be considered or whether there is an optimal sampling size and height threshold when estimating low corn LAI, height and biomass using discrete-return LiDAR data. Nonetheless, little research has been performed on corn vegetation. Moreover, for the same LiDAR data and study area, the differences of the optimal sampling size and height threshold for estimating vegetation LAI, height and biomass have not yet been conducted. Thus, it is still necessary to perform comprehensive research on the effects of LiDAR point density, sampling size and height threshold on the estimates of corn biophysical parameters.

In this study, we explore the effects of LiDAR point density, sampling size and height threshold on the estimation accuracy of corn crop biophysical parameters (height, LAI and aboveground biomass). The specific objectives of this study are to: (1) extract various metrics from the airborne discrete-return LiDAR data, (2) estimate corn biophysical parameters using the metrics derived from LiDAR data, and (3) assess the effectiveness of different LiDAR point densities, sampling sizes and height thresholds on the estimation accuracy of corn biophysical parameters.

2. Materials and methods

2.1 Study area

The study area is located in Zhangye City in Gansu Province of northwest China (38°50'N-38°59'N, 100°20'E-100°28'E). This study is a component of the experimental Heihe Watershed Allied Telemetry Experimental Research (HiWATER) project and the scientific objectives of this project can be found in Li et al. [53]. The mean annual air temperature and precipitation are approximately 7.3 °C and 129 mm, respectively. The terrain in our study area is flat with a mean elevation of 1403 m above sea level. The main crops in this study area are corn, wheat and vegetables.

2.2 Field observations

In this study, the field experiments were conducted on July 13th and 16th of 2012. The size of the field plot was a 5 m × 5 m square. In each plot, the height values of 11 corn plants were measured using a measuring tape. Figure 1 illustrates the selected 11 corn plants for height measurements in a field plot. We placed two tapes in the field plot along and perpendicular to the ridge, respectively. Along the ridge, the measurement started from 0 m of the tape, and six plants were measured with about 1 m interval. However, perpendicular to the ridge, the measurement started from 0.5 m of the tape, and five plants were measured with about 1 m

interval. The mean of the corn height was calculated using these 11 height values, and this value was used as the field-measured height of the plot. Moreover, the number of plants in each field plot was counted to calculate the biomass of each plot. The LAI values were measured using the LAI-2200 Plant Canopy Analyzer (Li-COR, Inc., Lincoln, NE, USA) at the ground level. In the study area, we selected a range of corn plants with different heights and clipped them at the ground level. The height of each plant was measured in the laboratory, and all plants were oven dried at 75 °C until their weights reached a constant value. The empirical relationship (Eq. (1)) between the plant heights and AGB values was established, and the R^2 was 0.95 with RMSE of 15.8 g/plant. The single plant biomass was calculated using the field-measured height of each field plot. The biomass of each field plot was calculated by multiplying the number of plants for each plot and the biomass of a single plant for each plot. Therefore, the above-ground biomass density (g/m^2) was obtained using the ratio of the plot biomass to the plot area (i.e., the plot biomass divided by 25 m^2). For each plot, the center coordinate was determined using a Real Time Kinematic (RTK) Global Positioning System (GPS)-Trimble GPS (Trimble Navigation Ltd.). A total of 42 plots were measured for height, LAI and biomass. The summary statistics of the field-measured biophysical parameters for the 42 field plots are listed in Table 1.

$$B = 129.72 * H - 131.16 \quad (1)$$

where B and H are the biomass (g/plant) and height (m) of a single plant, respectively.

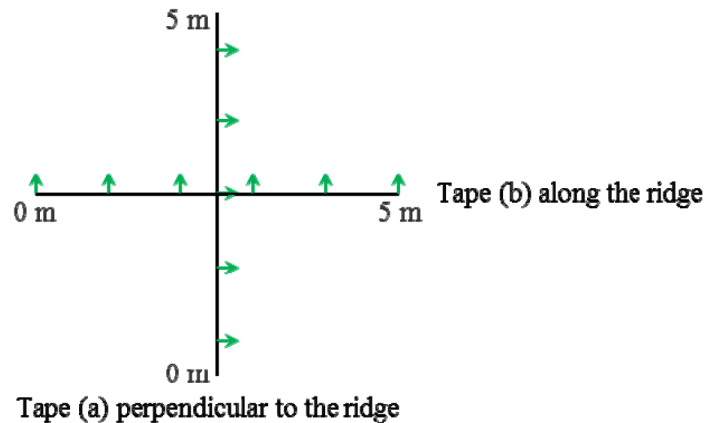


Fig. 1. Schematic diagram illustrating the selected 11 corn plants for height measurements.

Table 1. Statistics of field-measured biophysical parameters at the plot level ($n = 42$).

Biophysical parameter	Minimum	Maximum	Mean	Standard deviation
LAI	1.28	4.69	2.75	0.84
Height (m)	1.059	3.091	2.08	0.555
Biomass (g/m^2)	204.83	3039.82	1465.49	779.56

2.3 LiDAR data acquisition and processing

The LiDAR data used in this study were provided by HiWATER [54]. Airborne discrete-return LiDAR data were collected on July 19, 2012 using a Leica Airborne Laser Scanner (ALS70) system, and the specific acquisition parameters of the LiDAR data in this study are shown in Table 2. The final LiDAR data sets were produced in a LiDAR data exchange format (LAS) and included x , y , z coordinates, return numbers and intensity values. For the study area, the average point density was 7.32 points/m^2 with an average post spacing of 0.37 m. A total of ten different point densities were applied to estimate the biophysical parameters. First, raw laser point clouds were pre-processed [55], and LiDAR data were subsequently

classified into ground and non-ground points using the LiDAR data post processing software (TerraScan, TerraSolid, Ltd., Finland).

Table 2. LiDAR data acquisition parameters used in this study.

Parameter	Specification
Flying altitude	1300 m
Flying speed	60 m/s
Scan angle	$\pm 18^\circ$
Flight-line overlap	60%
Laser wavelength	1064 nm

To study the effect of different LiDAR point densities on estimates of biophysical parameters, the original LiDAR point density was thinned using Idaho State University's BCAL LiDAR Tools (<http://bcal.boisestate.edu/tools/lidar>). In this study, the original LiDAR point density was reduced from 100% to 5% via a random selection method. The major advantage of the method was that it can maintain a uniform distribution of LiDAR data, which was also used by previous researchers [47,51]. In the same manner, the thinned LiDAR data were classified into ground and non-ground points. Table 3 shows the statistics for LiDAR point density and post spacing at 11 different point densities (i.e., ten reduced and one original point densities). For classified LiDAR data, DEMs with a 1.0 m grid size were produced using a triangulated irregular network (TIN) interpolation method. The normalized LiDAR data (i.e., relative height of LiDAR data) were derived by subtracting the DEM elevations from the laser point elevations.

Table 3. Statistics of LiDAR point density and post spacing at the plot level (n = 42).

Percentage of original point density	Point density (points/m ²)			Post spacing (m)		
	Minimum	Maximum	Mean	Minimum	Maximum	Mean
5	0.11	0.85	0.37	1.08	3.02	1.64
10	0.21	1.66	0.74	0.78	2.18	1.16
20	0.42	3.33	1.47	0.55	1.54	0.82
30	0.64	5.02	2.21	0.45	1.25	0.67
40	0.85	6.69	2.94	0.39	1.08	0.58
50	1.10	8.35	3.69	0.35	0.95	0.52
60	1.31	10.01	4.42	0.32	0.87	0.48
70	1.52	11.68	5.16	0.29	0.81	0.44
80	1.73	13.38	5.89	0.27	0.76	0.41
90	1.95	15.04	6.63	0.26	0.72	0.39
100	2.16	16.70	7.32	0.24	0.68	0.37

2.4 Predictor variables from LiDAR data

Common LiDAR metrics used to estimate vegetation biophysical parameters are the maximum height and mean height of vegetation returns, percentiles of LiDAR heights and proportion of canopy returns. None of the LiDAR metrics are always optimal variables for estimating vegetation biophysical parameters for the different vegetation types, LiDAR data, environments and study areas. Therefore, in this study, we calculated a variety of LiDAR metrics to establish the optimal predictive model for vegetation biophysical parameters (Table 4).

A number of studies have indicated that the sampling size for the LiDAR data [52,56] and the height thresholds for discriminating vegetation points from ground points [28,57] have a great effect on the estimation accuracies of vegetation parameters. As a consequence, we also conducted a study to assess the effects of data sampling sizes and height thresholds on the estimation accuracies of vegetation biophysical parameters. We used a range of radius plots

(from 1.5 m to 6.5 m incremented by 0.5 m) to extract the normalized laser points that correspond to the field-measured plots. Subsequently, for all extracted plots, LiDAR metrics were calculated using a range of height thresholds (from 0.0 m to 0.3 m incremented by 0.05 m). In this study, the R_cover metric, which represents the fractional canopy cover, was calculated using Eq. (2).

$$R_cover = \frac{N_{canopy}(h)}{N_{all}} \quad (2)$$

where R_cover is fractional canopy cover with sampling radius r at a height threshold of h , N_{canopy} is the number of canopy returns where LiDAR height is greater than h m, N_{all} is the total number of returns.

Table 4. LiDAR-derived metrics for estimating biophysical parameters.

LiDAR metrics	Description
H_max	Maximum of LiDAR height
H_mean	Mean of LiDAR height
H_sd	Standard deviation of LiDAR height
H_var	Variance of LiDAR height
H_cv	Coefficient of variation of LiDAR height
H_p (10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 99)	percentile of LiDAR height
R_cover	Percent of canopy returns (canopy returns / total returns), a description of fractional canopy cover

2.5 Estimations of biophysical parameters

In our study, the biophysical parameters of corn were estimated based on empirical relationships. To investigate the nonlinear relationships between biophysical parameters and LiDAR metrics, all LiDAR metrics and biophysical parameters were log transformed. The selection of the optimal predictive variable is an important step for estimating biophysical parameters. In this study, the performances of the LiDAR-derived metrics were tested using simple linear regression analysis and partial least squares (PLS) regression. PLS regression was performed using the PLS extension module for SPSS Statistics. PLS regression is a multivariate statistical method, closely related to principal components regression (PCR) [58]. PLS regression can effectively overcome multicollinearity problems of LiDAR metrics that are often faced in multiple linear regression [59,60]. The Variable Importance in the Projection (VIP) values are calculated as a weighted sum of squares of the PLS loadings for each variable. VIP values measure the contribution of each predictor variable in fitting the PLS model, and variables with VIP values > 1 were considered in our study [61,62]. All prediction models were developed based on the regression relationships between field-observed biophysical parameters (42 plots) and all LiDAR-derived metrics (Table 4) with a range of LiDAR point densities (from 100% to 5% of original point density), sampling sizes (from 1.5 m to 6.5 m incremented by 0.5 m) and height thresholds (from 0.0 m to 0.3 m incremented by 0.05 m). We compared the predictive results of the developed models to determine the optimal predictor variable, sampling size and height threshold for estimating each of the biophysical parameters (height, LAI and biomass).

2.6 Accuracy assessment

In this study, the estimation accuracy of the biophysical parameters was assessed based on commonly used statistical indicators, such as the coefficient of determination (R^2), the adjusted coefficient of determination ($adj.R^2$) and the root mean squared error (RMSE) and relative RMSE (RMSE_r) [Eq. (3)]. For all predictive models, these indicators were produced using SPSS statistical software (version 20). Because all field-measured data were used to

establish regression models, no additional data were available to validate the predictive values of the developed models. Therefore, the leave-one-out cross-validation (LOOCV) method was used to validate the predictive power of the models. The LOOCV is an efficient method for assessing the generalization ability of the predictive model [60]. The RMSE from the LOOCV ($RMSE_{cv}$) was calculated using Eq. (4), and a low $RMSE_{cv}$ value indicates good predictive power of the model [30,63].

$$RMSE_r = \frac{RMSE}{\bar{y}} \quad (3)$$

where \bar{y} is the mean of the field-observed biomass.

$$RMSE_{cv} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (4)$$

where y_i is the field-observed biomass of sample i , \hat{y}_i is the predicted-biomass of sample i , and n is the total number of samples.

3. Results and discussion

3.1 Estimations of biophysical parameters

First, we estimated corn height, LAI and aboveground biomass using the original LiDAR point density to explore the potential of LiDAR data for estimating these biophysical parameters. We performed a simple linear regression analysis and PLS regression analysis between the field-observed biophysical parameters (42 plots) and all LiDAR metrics. In fitting the PLS model, only variables with VIP values > 1 were used (Fig. 2). For the single variable, the optimal predictor variables for estimating LAI, height and biomass were R_{cover} , H_{mean} and H_{mean} , respectively. The best predictive results for a range of sampling sizes and height thresholds are shown in Table 5. Figure 3 indicates the relationships of the predicted biophysical parameters against field-observed corn biophysical parameters for the best prediction models using the highest LiDAR point density (7.32 points/m²). The R^2 values for the best predictive model of LAI, height and biomass were 0.80 ($RMSE_r = 13.7\%$, $p < 0.0001$), 0.874 ($RMSE_r = 9.6\%$, $p < 0.0001$) and 0.838 ($RMSE_r = 21.7\%$, $p < 0.0001$), respectively. The PLS regression analysis method improved the predictive accuracies of corn LAI, height and biomass, and the R^2 improved by 1.5%, 4.9% and 4%, respectively, compared with the single linear regression, although the improvements were marginal. In addition, we found that the optimal sampling size (radius) and height threshold were different for estimating the different biophysical parameters. The optimal sampling radius and height threshold were 4.0 m and 0.15 m, respectively, for the LAI estimation, but for the height and biomass estimation, 3.0 m and 0.10 m were the optimal values. Both the optimal sampling radius and height threshold for the LAI estimation were greater compared with those of the height and biomass estimations. In this study, all R^2 values for the estimation models of three biophysical parameters were greater than or equal to 0.788, which showed strong relationships between the field-observed physical parameters and LiDAR-derived metrics. Moreover, for all predictive models, the $RMSE_{cv}$ values derived from LOOCV cross-validation were in agreement with their $RMSE$ values, which showed that the estimation models of biophysical parameters developed using LiDAR data had good predictive ability. In general, discrete-return LiDAR data were able to reliably estimate the corn LAI, height and aboveground biomass in our study. Therefore, LiDAR data have great potential for estimating biophysical parameters of crops.

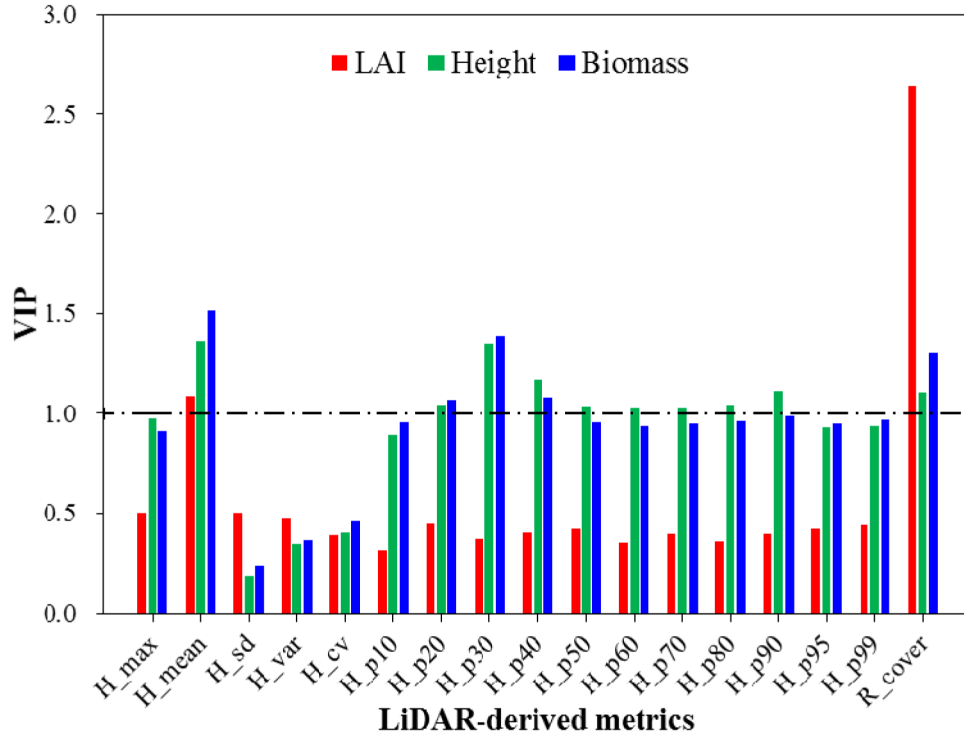


Fig. 2. VIP values of LiDAR-derived metrics in fitting the PLS model.

Table 5. LAI, height and aboveground biomass estimation accuracies using single linear and PLS regression methods from the original LiDAR point density data (n = 42).

Regression method	R ²	adj.R ²	RMSE	RMSE _{cv}	RMSE _r (%)	Sampling radius (m)	Height threshold (m)
Single regression	0.788	0.782	0.388	0.40	14.1	4.0	0.15
PLS regression	0.800	0.795	0.376	0.389	13.7	4.0	0.15
Single regression	0.833	0.829	0.230 (m)	0.237 (m)	11.1	3.0	0.10
PLS regression	0.874	0.871	0.199 (m)	0.205 (m)	9.6	3.0	0.10
Single regression	0.806	0.801	347.74 (g/m ²)	356.98 (g/m ²)	23.7	3.0	0.10
PLS regression	0.838	0.834	317.76 (g/m ²)	325.07 (g/m ²)	21.7	3.0	0.10

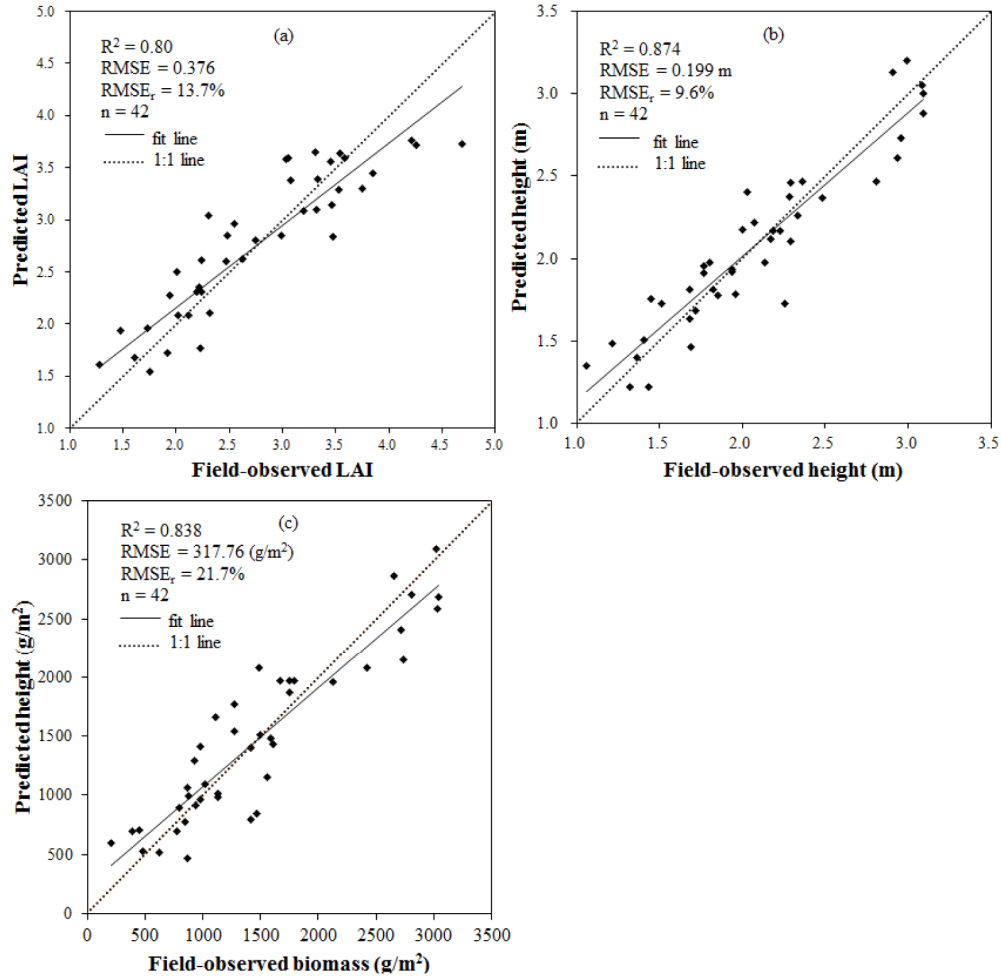


Fig. 3. Scatterplots of the predicted biophysical parameters against field-observed corn biophysical parameters and regression lines of the best prediction models for (a) LAI, (b) height, and (c) AGB. Solid lines indicate the best-fit regression line. Dotted lines denote the 1:1 line.

3.2 Effects of LiDAR point density on estimates of biophysical parameters

To investigate the effects of LiDAR point densities on the estimates of biophysical parameters, we estimated the corn LAI, height and biomass based on different sampling sizes (from 1.5 m to 6.5 m in radius) and height thresholds (from 0.0 m to 0.3 m incremented by 0.05 m). In this study, we used a single LiDAR variable to study the effects of LiDAR point densities on estimates of biophysical parameters, although multiple LiDAR variables produced higher estimation accuracy for corn height and biomass. Eleven different LiDAR point densities (from 0.37 to 7.32 points/m²) (Table 3) were used to estimate the corn biophysical parameters. Figure 4 shows the changes of R² and RMSE of the estimation models for three biophysical parameters across different LiDAR point densities. We found that, for three biophysical parameter models, the R² values did not monotonically increase or decrease with increasing LiDAR point density (Fig. 4). When the lowest point density (0.37 points/m²) was used, R² values of 0.492, 0.720 and 0.677 were obtained for estimating LAI, height and biomass, respectively (Table 6). Nevertheless, all three biophysical parameters achieved a relatively high estimation accuracy (R² = 0.607, 0.807 and 0.778, respectively)

when using 10% of the original LiDAR point density (0.74 points/m²). For the LAI estimation, the highest accuracy was obtained using the highest LiDAR point density (7.32 points/m²). A rapid decline in the estimation accuracy of corn LAI could be observed when the LiDAR point density was reduced to values below 2.94 points/m². Nevertheless, the decrease of point density from 7.32 points/m² to 2.94 points/m² did not considerably reduce the estimation accuracy of LAI. The maximum difference of R² (29.6%) for the LAI estimation model was noted between LiDAR point densities of 7.32 points/m² and 0.37 points/m² (Table 6). For the height and biomass estimations, however, the highest accuracies were obtained using 70% of the original point density (5.16 points/m²). Moreover, the reduced LiDAR pulse densities, except for 0.37 points/m², had less effect on the estimation accuracy of corn height and biomass. The results showed that the LiDAR point density could be reduced from 7.32 points/m² to 0.74 points/m² to estimate corn height and biomass, and reasonable results were still produced. The maximum differences of R² for the height and biomass models were 12% and 15.2%, respectively, which occurred between densities of 5.16 points/m² and 0.37 points/m² (Table 6). We found that a relatively higher LiDAR point density was required to acquire accurate LAI estimates, but accurate height and biomass estimates could be produced using a lower LiDAR point density. In general, the reduced LiDAR point density could be used to estimate the biophysical parameters with reasonable results or without markedly reducing the estimation accuracy. This result was similar to the findings of Jakubowski et al. [44]. From Fig. 4, we also found that simply increasing LiDAR point density might be sometimes ineffective for improving the estimation accuracy of biophysical parameters. In addition, the costs of data acquisition significantly increase as the point density increases, especially for large-area LiDAR data collection. Therefore, LiDAR data should be collected according to the research cost, study coverage and research purposes to obtain an acceptable trade-off between the point density and cost of data acquisition.

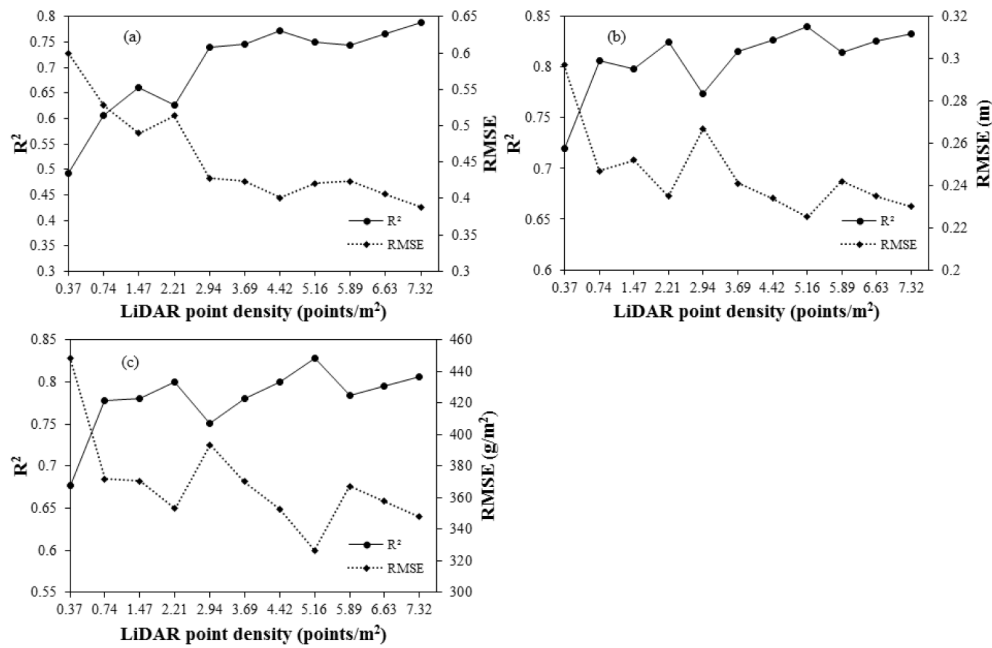


Fig. 4. Changes in R² and RMSE with point densities for (a) LAI, (b) height, and (c) AGB.

Table 6. Estimation accuracies of LAI, height and biomass with varying point density.

p ^a	d ^b	LAI				Height				Biomass			
		R ²	adj.R ²	RMSE	RMSE _r	R ²	adj.R ²	RMSE	RMSE _r	R ²	adj.R ²	RMSE	RMSE _r
5	0.37	0.492	0.479	0.599	21.8	0.720	0.710	0.297	14.3	0.677	0.669	448.75	30.6
10	0.74	0.607	0.507	0.528	19.2	0.807	0.803	0.247	11.9	0.778	0.772	371.91	25.4
20	1.47	0.661	0.653	0.490	17.8	0.798	0.793	0.252	12.1	0.780	0.774	370.52	25.3
30	2.21	0.626	0.617	0.514	18.7	0.825	0.820	0.235	11.3	0.800	0.795	353.20	24.1
40	2.94	0.741	0.734	0.428	15.6	0.774	0.768	0.267	12.8	0.751	0.745	393.51	26.9
50	3.69	0.746	0.740	0.424	15.4	0.816	0.812	0.241	11.6	0.780	0.775	370.06	25.3
60	4.42	0.773	0.767	0.401	14.6	0.827	0.823	0.234	11.3	0.800	0.795	352.79	24.1
70	5.16	0.750	0.743	0.421	15.3	0.840	0.836	0.225	10.8	0.829	0.824	326.66	22.3
80	5.89	0.745	0.739	0.424	15.4	0.815	0.810	0.242	11.6	0.784	0.778	367.18	25.1
90	6.63	0.767	0.761	0.406	14.8	0.826	0.821	0.235	11.3	0.795	0.790	357.60	24.4
100	7.32	0.788	0.782	0.388	14.1	0.833	0.829	0.230	11.1	0.806	0.801	347.74	23.7

^a p represents percentage of original point density; ^b d represents point density (points/m²)

3.3 Effects of LiDAR data sampling size on estimates of biophysical parameters

We estimated corn LAI, height and biomass using the original LiDAR point density with a range of sampling sizes (from 1.5 m to 6.5 m in radius). Figure 5 shows the effects of LiDAR data sampling size on the estimation accuracies of the biophysical parameters. From 1.5 m to 6.5 m sampling sizes, the lowest accuracy was found at 1.5 m for all three biophysical parameter models. However, the optimal LiDAR data sampling sizes for estimating corn LAI, height and biomass were 4.0 m, 3.0 m and 3.0 m in radius, respectively, and the greatest differences in R² for the estimation models of the three vegetation parameters were 11.2%, 11.5% and 11.4%, respectively. As clearly shown in Fig. 5, the R² values always increased before reaching a peak, and subsequently began to decline. When the sampling size was smaller, there were insufficient ground returns to produce accurate DEMs, which would therefore affect the estimation accuracy of corn physical parameters. While a larger sampling size was used, the estimates of corn physical parameters could not effectively characterize the actual values of the field-measured plots. In short, the estimation accuracy of corn biophysical parameters using LiDAR data was affected by sampling sizes, and similar results were found by previous researchers [28,52]. Moreover, our findings indicated that the optimal sampling radius for the LAI estimation was greater than those of the height and biomass estimations. Therefore, the optimal sampling size differed due to different biophysical parameters, vegetation types, geographical settings and LiDAR data. Therefore, to improve the estimation accuracy of biophysical parameters from LiDAR data, the optimal sampling size should be determined based on data from the study area.

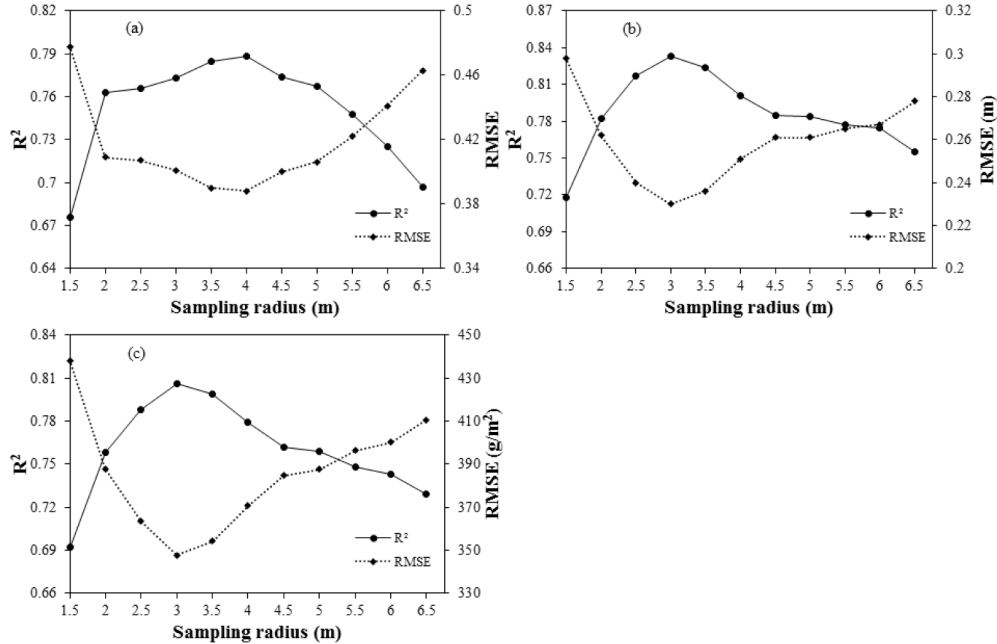


Fig. 5. Effects of varying sampling sizes on the estimation accuracy of biophysical parameters (a) LAI, (b) height, and (c) AGB. The estimation models were constructed based on the original LiDAR point density of 7.32 points/m².

3.4 Effects of height threshold on estimates of biophysical parameters

Corn LAI, height and biomass were estimated using the original LiDAR point density (7.32 points/m²) based on a range of height thresholds (from 0.0 m to 0.3 m). Figure 6 shows the effects of height thresholds on the estimation accuracy of the biophysical parameters. The optimal height thresholds for estimating LAI, height and biomass were 0.15 m, 0.10 m and 0.10 m, respectively. Compared with the estimations of height and biomass, the LAI estimation needed a larger height threshold to obtain the highest estimation accuracy. The maximum difference in R² for the LAI estimation models was 10.3%, which occurred between 0.15 m and 0.30 m height thresholds. However, the maximum differences in R² for the height and biomass estimation models were 2.4% and 3.9%, respectively, which occurred between 0.10 m and 0.30 m height thresholds. The height threshold was used to separate ground and vegetation returns. The optimal height threshold improved estimation accuracies of corn LAI, height and biomass. Therefore, the effects of height threshold should be considered when estimating corn biophysical parameters using discrete-return LiDAR data. If the height threshold was smaller, the effects of underlying ground cover could not be effectively reduced. If the height threshold was larger, too many vegetation returns would be classified as ground points. Thus, a suitable height threshold is essential and can improve the estimation accuracies of the biophysical parameters. Similar to the optimal sampling size for LiDAR data, the optimal height threshold also differed for different vegetation types, LiDAR data, environments and study areas. Therefore, the optimal height threshold in our study cannot be directly used for other vegetation types or study areas, but the methods proposed in this study can be applied to other similar studies.

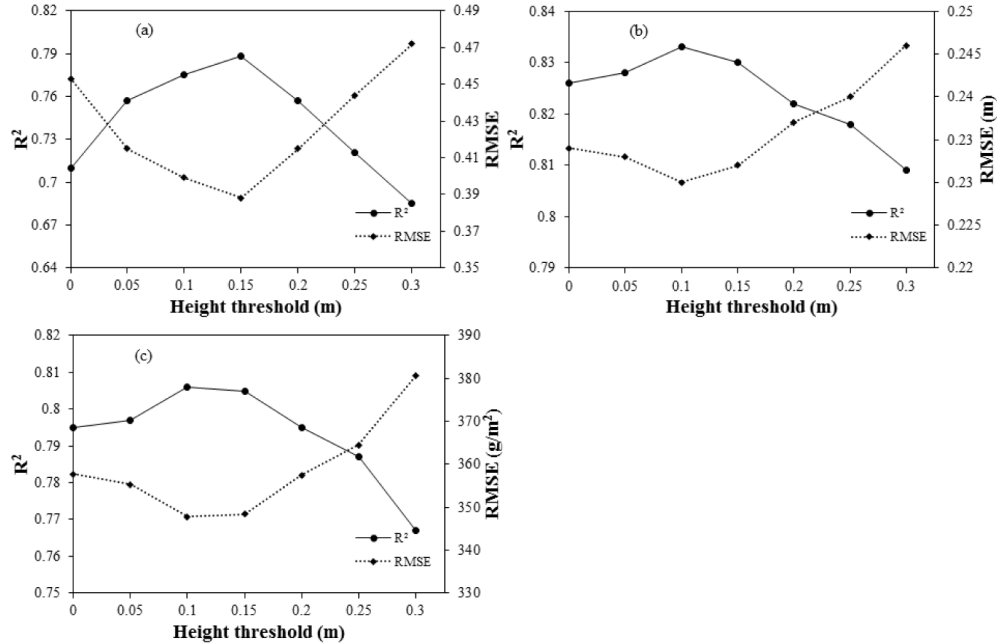


Fig. 6. Effects of varying height thresholds on the estimation accuracy of biophysical parameters (a) LAI, (b) height, and (c) AGB. The estimation models were constructed based on the original LiDAR point density of 7.32 points/m².

4. Conclusions

In this study, we investigated the effects of LiDAR point density, sampling size and height threshold on the estimation accuracy of corn biophysical parameters (LAI, height and biomass). First, the potential of discrete-return LiDAR data for estimating corn biophysical parameters was explored. All R^2 values for the developed models were greater than or equal to 0.788. In addition, we validated all estimation models using the LOOCV cross-validation method, and the results indicated good predictive ability for the models. For the single variable, the optimal predictor variables for estimating LAI, height and biomass were R_{cover} , H_{mean} and H_{mean} , respectively. A study was conducted on the effect of varying LiDAR point density (from 0.37 to 7.32 points/m²) on the estimation accuracy of corn LAI, height and biomass. For corn height and biomass estimations, the highest LiDAR point density (7.32 points/m²) did not produce the best estimation accuracies, but good accuracy was observed at a point density of 5.16 points/m². Although the best accuracy of LAI estimation was produced at the highest point density, the LAI estimation accuracy did not always improve with increasing point density. The effects of the LiDAR data sampling size and height threshold on estimation accuracy were assessed using the original point density. The results showed that the sampling size and height threshold have marked effects on estimation accuracy of three parameters. Moreover, the optimal sampling size and height threshold for estimating LAI were greater than those for height and biomass.

In general, the estimation accuracies for corn LAI, height and biomass were affected not only by LiDAR point density but also by LiDAR data sampling size and height threshold. The results indicated that the estimation accuracy did not increase significantly or could even decrease if the LiDAR point density exceeded a certain level. Thus, a high LiDAR point density did not always produce a high estimation accuracy for the corn biophysical parameters, and a low point density could be used to reliably estimate the biophysical parameters. Therefore, a trade-off should be expected between point density and estimation accuracy to achieve a cost-efficient result. The optimal sampling sizes and height thresholds

were varied for estimation of different biophysical parameters; these values should be determined based on collected LiDAR data and field-observed data and cannot be used directly in other studies. However, the methods proposed in this study, including the methods of corn biophysical parameters estimation using LiDAR data, and the determination of the optimal LiDAR data sampling size and height threshold, can be used for similar studies. Overall, our results provide reliable information for researchers studying crops or similar vegetation types when using LiDAR data.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (Nos. 41371350 and 41271428); the International Postdoctoral Exchange Fellowship Program 2014 by the Office of China Postdoctoral Council (20140042). We thank Andreas Behrendt and the two anonymous reviewers for their thoughtful comments and suggestions on the manuscript.