International Journal of Digital Earth
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tjde20

Ground peak identification in dense shrub areas using large footprint waveform LiDAR and Landsat images

Wei Zhuang[a] & Giorgos Mountrakis[a]
[a] Department of Environmental Resources Engineering, State University of New York College of Environmental Science and Forestry, Syracuse, NY, USA
Accepted author version posted online: 09 Jul 2014.Published online: 13 Aug 2014.

To cite this article: Wei Zhuang & Giorgos Mountrakis (2014): Ground peak identification in dense shrub areas using large footprint waveform LiDAR and Landsat images, International Journal of Digital Earth, DOI: 10.1080/17538947.2014.942716

To link to this article: http://dx.doi.org/10.1080/17538947.2014.942716

International Journal of Digital Earth, 2014
http://dx.doi.org/10.1080/17538947.2014.942716

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Ground peak identification in dense shrub areas using large footprint waveform LiDAR and Landsat images

Wei Zhuang* and Giorgos Mountrakis

Department of Environmental Resources Engineering, State University of New York College of Environmental Science and Forestry, Syracuse, NY, USA

(*Received 29 January 2014; accepted 4 July 2014*)

Large footprint waveform LiDAR data have been widely used to extract tree heights. These heights are typically estimated by subtracting the top height from the ground. Compared to the top height detection, the identification of the ground peak in a waveform is more challenging. This is particularly evident in ground detection in shrub areas, where the reflection of the shrub canopy may significantly overlap with the ground reflection. To tackle this problem, a novel method based on Partial Curve-Fitting (PCF) of the shrub peak was developed to detect the ground peak. Results indicated that the PCF method improves ground identification by 32–42%, compared to existing methods. To offer further improvement, a Multi-Algorithm Integration Classifier (MAIC) was built to fuse multiple ground peak algorithms and selectively apply the best method for each waveform plot. The PCF ground peak identification method along with the MAIC-based fusion is expected to significantly improve ground detection and shrub height estimation, thus assisting biodiversity, forest succession, and carbon sequestration studies, while offering an early example of future multiple algorithm integration.

Keywords: ground identification; large footprint; waveform LiDAR; shrub; partial fitting; algorithmic fusion

1. Introduction

Canopy height is an important vegetation structure characteristic for numerous ecological studies (Riaño et al. 2007; Skowronski et al. 2007; Boudreau et al. 2008; Martinuzzi et al. 2009). Specifically, accurate canopy height estimation in shrub areas (i.e. shrub heights) is crucial for fuel load management (Saatchi et al. 2007; Skowronski et al. 2007; Jakubowksi et al. 2013) and wildlife habitat characterization (Hyde et al. 2005; Pinotti, Pagotto, and R. Pardini 2012; Smart et al. 2012). However, conventional methods that measure shrub height involve significant fieldwork, and therefore are time-consuming and labor-intensive. To improve efficiency and cost, advanced discrete return LiDAR technologies have been recently applied to estimate shrub heights (Rango et al. 2000; Su and Bork 2007; Wang and Glenn 2009; Lee et al. 2011; Sun et al. 2011; Estornell et al. 2012).

The Root Mean Square Error (RMSE) of shrub height estimation using discrete return LiDAR data was shown to be comparable to that of tree height estimation in mature forests. However, due to the much lower height of shrub areas, the relative accuracy of

*Corresponding author. Email: will.zhuang1982@gmail.com

shrub height estimation is much lower than that of tree height estimation (Clark, Clark, and Roberts 2004; Streutker and Glenn 2006; Popescu and Zhao 2008; Mitchell et al. 2011; Wasser et al. 2013). For example, Su and Bork (2007) investigated the relationship between LiDAR-derived heights and the measured heights of both the *Populus* and the understory shrub layer. They found that the prediction of shrub height had smaller RMSE (i.e. 0.48 m) when compared with the mean height of shrub (i.e. 0.63 m), however, the prediction results only explained a maximum of 21% of the variance in the shrub heights. On the contrary, the prediction of *Populus* tree had an R^2 of 0.93. Glenn et al. (2011) summarized several possible reasons for the inaccurate results in estimating heights for low-height vegetation, using discrete LiDAR: (1) the inability to separate LiDAR data points from vegetation and ground returns (Riaño et al. 2007); (2) the occlusion effect from canopy, which leads to low penetration rate of LiDAR data points (Streutker and Glenn 2006; Gould et al. 2013); and (3) the effect of slope on increasing classification errors of vegetation and ground (Hodgson and Bresnahan 2004; Su and Bork 2007; Spaete et al. 2011).

Compared to discrete return LiDAR, waveform LiDAR sensors with large footprint (10–80 m), such as Geoscience Laser Altimeter System (GLAS) and Land, Vegetation, and Ice Sensor (LVIS), have several advantages (Blair, Rabine, and Hofton 1999; Drake et al. 2002a; Mallet and Bretar 2009; Chen 2010a): (1) stronger laser energy that increases the ground reflection; (2) larger area coverage with reduced budget; (3) adaptive feature extraction for different applications; and (4) ability to capture the canopy top due to the larger possibility of the signal return from the large footprint pulse. Large footprint waveform LiDAR data have been widely used in extracting canopy height for large-scale research (Means et al. 1999; Drake et al. 2002b; Anderson et al. 2006; Sun et al. 2008; Wang et al. 2010). Canopy height is calculated as the distance between the canopy top and the ground return found in a waveform (Chen 2010b).The canopy top is identified as the location where the intensity becomes larger than a threshold related to the signal background noise (Drake et al. 2002b; Hyde et al. 2005).The ground is identified as the centroid of the lowest Gaussian function in the Gaussian Decomposition (GD) of a waveform (Hofton, Minster, and Blair 2000; Sun et al. 2003; Hyde et al. 2005). Another method, the FIltering and Clustering Method (FICA; Zhuang and Mountrakis, in review) integrates a second derivative filter and the *k*-means clustering algorithm to identify accurately and computationally efficiently the lowest peak in a waveform.

LVIS is a large footprint waveform LiDAR sensor designed by NASA Goddard center (Blair, Rabine, and Hofton 1999). Typical LVIS products provided by NASA include elevation of the highest detected return, elevation of the centroid of the lowest Gaussian function following a GD method, and Relative Heights (RHs) calculated as the distance between the location at certain percentile of the total energy and the lowest centroid in the full waveform (Blair, Hofton, and Rabine 2006). The height of the lowest Gaussian function is assigned as the ground elevation. Consequently, RH100 (the height at which 100% of the return energy is received) is used as the maximum height within a plot in many studies related to forest structure characterization and biophysical parameters estimation (e.g. Goetz et al. 2007; Anderson et al. 2008). However, little research has been done to specifically target shrub height estimation using waveform LiDAR data. Although the centroid of the lowest Gaussian function has been successfully used as the ground peak location in mature forests, it may not be applicable for ground peak identification in shrub areas because the reflection of the short shrub canopy and the reflection of the ground merge into a single large peak. As a result, the ‘ground peak’

detected in the waveform could be the location of the shrub canopy rather than the actual ground. Hence, a more effective method is needed to accurately detect the ground peak in dense shrub areas, using waveform LiDAR data.

In this study, we propose a novel method for ground identification using waveform LiDAR specifically targeting dense shrub areas. The proposed method detects the ground peak based on a Partial Curve-Fitting (PCF) of the shrub signal of the waveform. The underlying idea is to first extract the shrub signal modeled by a Gaussian function and then work on the residual signal to identify a possible ground location. The objectives of this study are to (1) test the hypothesis that the PCF method can achieve better accuracy in ground identification than that of Gaussian decomposition in shrub waveforms; (2) develop a Multi-Algorithm Integration Classifier (MAIC) using Landsat/LVIS data that allows the fusion of multiple ground peak detection algorithms, in our case PCF and GD methods.

2. Study area

Our study area includes the central region of New York (NY) State surrounding the city of Syracuse (Figure 1). A significant number of farmlands has been abandoned in NY since the beginning of the twentieth century (Nyland et al. 1986; Smith, Marks, and Gardescu 1993), leading to a large amount of old fields with varying ages (Ramankutty and Foley 1999). The abandoned fields have been clear cut and are now characterized by herbs and shrubs (Swanson et al. 2011). Therefore, shrub, common vegetation at early forest successional stage, can be widely found throughout the study area.

3. Remote sensing data

Four different types of data were incorporated in this study: large footprint LVIS waveform, small footprint discrete return LiDAR data, aerial photos, and Landsat TM images. Details about these data are described in the following subsections.

3.1. Waveform LiDAR data

LVIS Waveform LiDAR data were acquired on 24–26 August 2009, during leaf-on season (Figure 1). LVIS is a waveform sensor that records a full waveform from the intercepted surfaces. In this study, the vertical resolution was 0.3 m with a nominally 20-m diameter footprint. The laser bandwidth was 1064 nm, and the waveform contained a total of 432 bins. The horizontal accuracy was 1–2 m. In addition to the center location of the pulse, elevation of the first and the last bin was included in the waveform data. The horizontal coordinate system was UTM 18N with NAD83 datum, and the vertical datum was transformed into NAVD88.

3.2. Discrete return LiDAR data

The airborne discrete return LiDAR data were acquired in April 2010, using a Leica ALS60 Airborne Laser Scanner during leaf-off so that the ground LiDAR points could be clearly identified. The footprint size was 0.37 m. The horizontal and vertical coordinate systems of the data were converted to the same as LVIS data. The point density was 2.1 points per square meter. Vertical RMSE of the discrete return data was 0.033 m based on 24 ground control points; the horizontal accuracy was assumed to be 1 m based on the

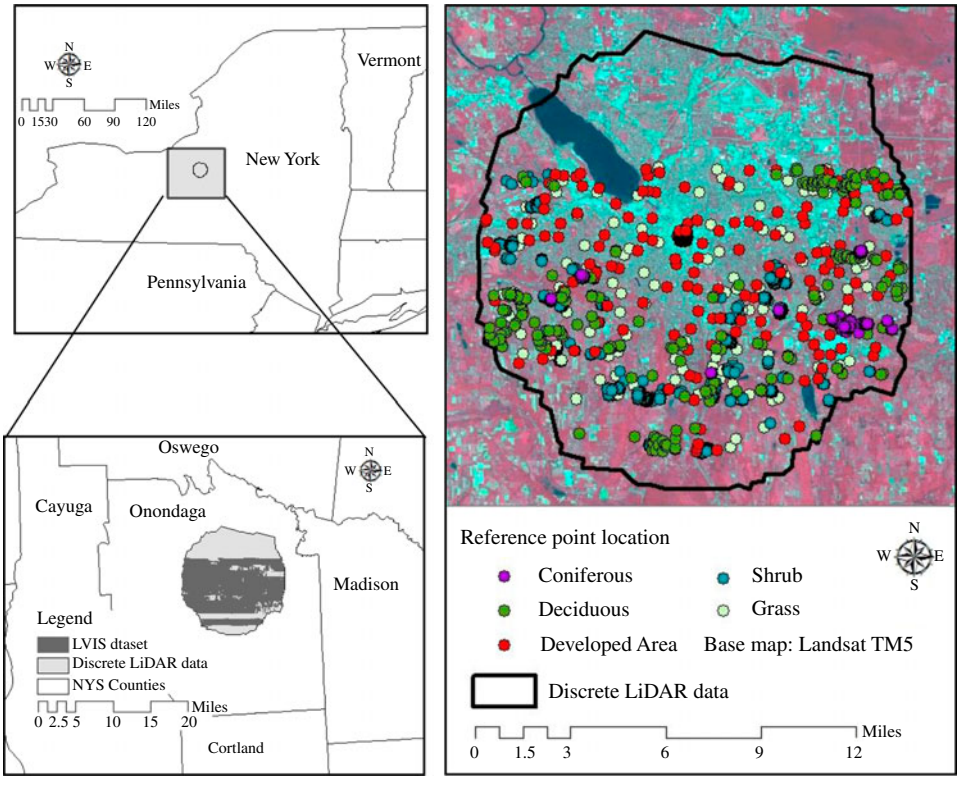

Figure 1. Study area and datasets.

data vendor's accuracy assessment report. The discrete return data were classified into ground and nonground points using the Terrasolid software.

3.3. Aerial photos and Landsat TM images

Digital Orthophotos (DOs) of the area were acquired in April 2009, which was also during the leaf-off season. The orthophoto had 1-ft pixel size and covered the entire study area. Two Landsat 5 TM scenes were collected for this study, acquired on 04/17/2009 and 07/15/2009 representing leaf-off and leaf-on seasons, respectively. Both TM images were cloud-free. Each TM image consisted of six visible and short-wave infrared bands with 30 m pixel size. The thermal band was not included. All Digital Number (DN) values of the TM images were converted to at-satellite radiances (Chander, Markham, and Helder 2009).

4. Reference data for the PCF method

The reference shrub data were identified based on visual interpretation of aerial photos. A total of 1000 shrub plots with 10 m radius (i.e. LVIS footprints) were collected using a random sampling strategy from 5000 identified shrub plots in the classified shrub area, which was delineated manually. For each plot, the overlapping high-density discrete return LiDAR data were identified. As the discrete return LiDAR data were acquired in

the leaf-off season, it was difficult to accurately quantify the coverage of the shrub. The ratio between the number of nonground points and total number of points in a plot varied up to 30% thus providing general picture of vegetation coverage. The mean ground elevation within each plot was then estimated using the elevation of the ground points, which was at least 70% of the total points, labeled in the discrete data. The samples included high variability in slope reaching up to 25°. As the spatial energy distribution of an LVIS pulse was Gaussian-shaped from the plot center, a distance-based weighting strategy was applied when calculating the mean elevation (Chen 2010b), as shown in the equations below:

$$E_M = \sum_{i=1}^{N} \frac{W_i}{\sum_{i=1}^{N} w_i} \times E_i \tag{1}$$

$$W_i = \frac{1}{\sigma\sqrt{2\pi}} \times \exp^{-\frac{\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}}{2\sigma}} \tag{2}$$

where (x_i, y_i) is the coordinate for each discrete LiDAR ground point located within a circular waveform pulse at the center of (x_0, y_0). E_M is the weighted mean elevation calculated from each discrete LiDAR point, and N is the total number of discrete return LiDAR in the waveform circular pulse. W_i is the weight for a specific discrete LiDAR ground point (Blair, Rabine, and Hofton 1999), E_i is the elevation of a discrete LiDAR ground point.

5. Reference data for the MAIC

PCF was specifically designed for ground identification in low-height shrub environments. To ensure that the method was implemented only in areas where its benefits could be realized, a binary MAIC was developed to identify where to apply the PCF method and where to select the typical GD algorithm. For testing the MAIC, sampled plots were initially collected from a variety of land cover types. Beside the 1000 shrub plots described in the previous section, additional sampled plots included deciduous trees (1000), coniferous trees (1000), grass (200), and developed areas (200). The grass and developed area had fewer samples because the vertical structures of the ground objects were simpler and thus do not pose a significant algorithmic challenge. Other land classes, such as water, soil, and agriculture land, were not included in the samples because either no data was collected (water) or the ground identification in these classes was much easier. In order to include only low-height ground objects in MAIC, the aforementioned 3400 sampled plots were further filtered by the RH100 value (i.e. nominal maximum height) provided in NASA's dataset. The threshold for RH100 was set at 10 m and the 3400 plots were reduced to 1456 plots, which included 189 grass plots, 808 shrub plots, 150 deciduous plots, 168 coniferous plots, and 141 urban plots. The average ground elevation within a plot was calculated following Equations 1 and 2. The plots were then divided into training (50%) and testing (50%) datasets using a stratified random sampling. The strata were defined according to the histogram of the difference of the mean absolute error between the PCF method and GD, as shown in Table 1. The number of plots for each stratum was large enough to generate unbiased classification results.

Table 1. Stratification of the peak identification error difference between the PCF method and GD.

Stratum	Threshold (m)	Number of plots
Stratum 1	$\text{Error}_{PCF} - \text{Error}_{GD} -3.3$	107
Stratum 2	$-3.3 \leq \text{Error}_{PCF} - \text{Error}_{GD} -2.7$	155
Stratum 3	$-2.7 \leq \text{Error}_{PCF} - \text{Error}_{GD} -2.1$	230
Stratum 4	$-2.1 \leq \text{Error}_{PCF} - \text{Error}_{GD} -1.5$	194
Stratum 5	$-1.5 \leq \text{Error}_{PCF} - \text{Error}_{GD} -0.9$	107
Stratum 6	$-0.9 \leq \text{Error}_{PCF} - \text{Error}_{GD} -0.3$	96
Stratum 7	$-0.3 \leq \text{Error}_{PCF} - \text{Error}_{GD} < 0.3$	156
Stratum 8	$0.3 \leq \text{Error}_{PCF} - \text{Error}_{GD} < 0.9$	150
Stratum 9	$0.9 \leq \text{Error}_{PCF} - \text{Error}_{GD} < 1.5$	151
Stratum 10	$\text{Error}_{PCF} - \text{Error}_{GD} \geq 1.5$	110

PCF, Partial Curve-Fitting; GD, Gaussian decomposition.

In the binary MAIC process, the two reference classes were defined based only on the comparison between the error of the ground peak identified by the PCF method and error of the ground elevation from GD. Considering the vertical resolution of the LVIS waveform, which is 0.3 m, we defined two classes shown in the equations below:

$$\begin{aligned} C1 &= \text{Avoid PCF application}, \quad \textit{if } \text{Error}_{PCF} - \text{Error}_{GD} \geq -0.3 \text{ m} \\ C2 &= \text{Suggest PCF application}, \ \textit{if } \text{Error}_{PCF} - \text{Error}_{GD} < -0.3 \text{ m} \end{aligned} \tag{3}$$

where the errors were from different methods of the ground peak identification.

6. Methods

In this study, a new method based on PCF was developed for identifying ground peaks in the LiDAR waveform acquired in shrub areas. As the method was specifically designed for shrub waveforms, applicability of the method was further investigated using a MAIC to identify areas that PCF should be applied. This classifier incorporated metrics from both TM images and waveform datasets.

6.1. Extracting ground peak using PCF and comparison with existing methods

Figure 2 describes the different processing steps for the proposed PCF method. The hypothesis of the PCF method is that in order to separate the shrub and ground returns first, the shrub signal should be extracted and ground peak detection should take place on the residual signal. The shrub signal is extracted as follows: First, the point with maximum intensity is identified. It is assumed that in a dense shrub plot, the reflection of shrub canopy is generally stronger than the ground reflection because of the near-infrared (NIR) bandwidth of the laser pulse. Therefore, it is appropriate to assume that the maximum intensity in a waveform would represent the shrub peak. Second, a Gaussian function is interpolated to describe the shrub peak. The amplitude and the center location of the Gaussian function were set as the intensity and the location of maximum intensity, respectively. The only parameter that is investigated to define the Gaussian function is the width.

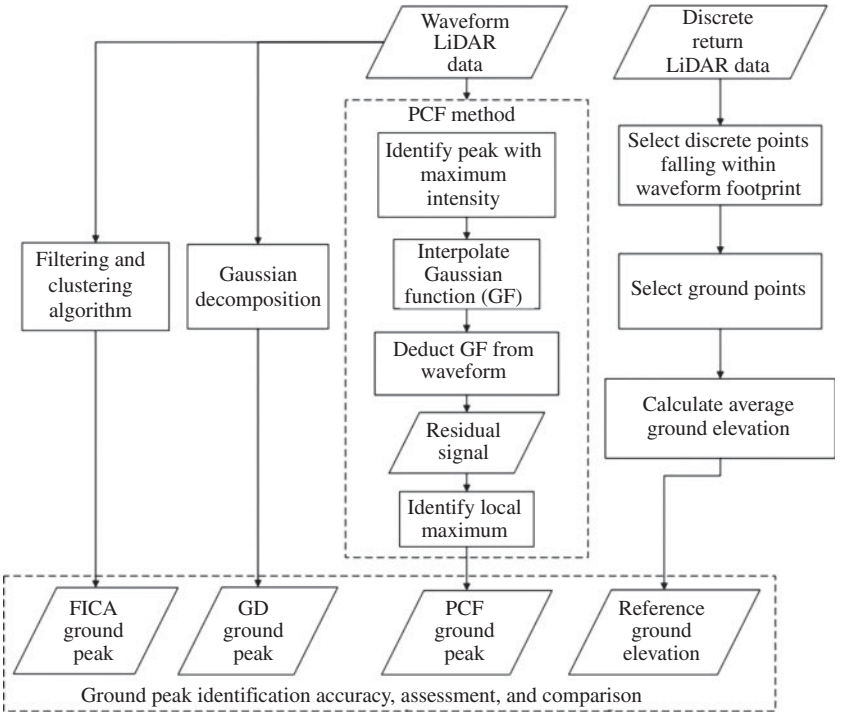

Figure 2. Flowchart of the PCF ground peak identification method.

The widths of the candidates were determined by Equation 4, which is simply an inverted solution for the width assuming two points are known:

$$\sigma_i = \frac{x_i - \mu}{\sqrt{2 * \log A/y_i}}, i = 1, 2, \ldots, 2N-1, 2N \tag{4}$$

values for A, μ, and N were held constant. A, μ varied for each waveform, while N was the same for all waveforms. A is the maximum intensity of the waveform under examination, μ is the (time) location of the maximum intensity for that particular waveform, and N is the maximum number of bins (time samples) defining the maximum neighborhood (in time, x-axis in Figure 3).The neighborhood size was assigned to 85% of the width of the LVIS-transmitted pulse, and it was calculated once for all waveforms by a Gaussian fit on the transmitted pulse. The rationale behind this decision was that a simulated Gaussian function cannot be separated if the distance between the centroids of two overlapping Gaussian functions was less than 85% of the width of the transmitted pulse (Jutzi and Stilla 2006).In Equation 4, the width was calculated for every solution involving two points, the center of the Gaussian described by point (μ, A) and another point to the left or the right of the center described as (x_i, y_i). In the two parameters tested x_i expressed the location of a waveform bin (time bin on the x-axis of Figure 3), and y_i expressed the intensity (y-axis in Figure 3). For each pair of (x_i, y_i), a width value σ_i was obtained. The minimum of σ_i values was selected as the σ_G width for the Gaussian

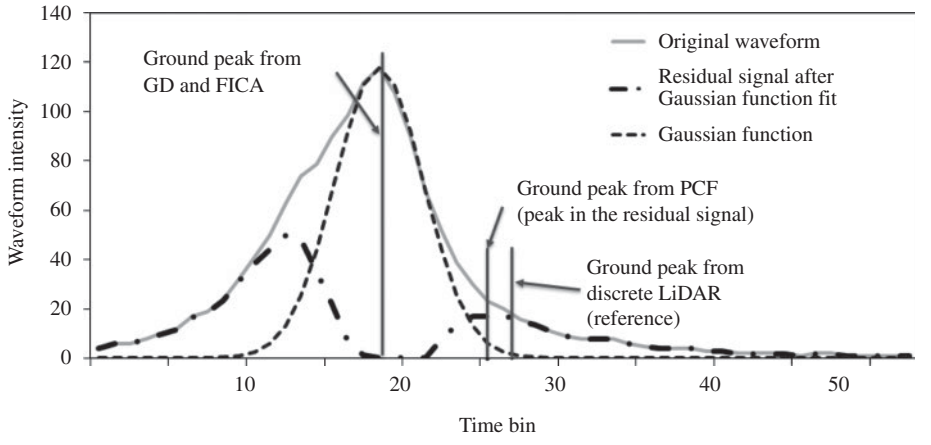

Figure 3. Example with ground peak identification from GD, FICA, and PCF in a waveform of dense shrub plot.

function representing the vegetation canopy:

$$\sigma_G = \min(\sigma_i), i = 1 \ \ldots \ 2N \tag{5}$$

The selection of the minimum width value was to generate a narrow Gaussian function so that the residual signal would include a strong representation of the ground return.

After the Gaussian function representing shrub canopy reflection was identified, it was subtracted from the original waveform. This residual signal was further examined to detect the ground peak. Our assumption was that as the ground return was normally a peak in a waveform (Drake et al. 2002a), which may be fitted by a Gaussian function. Therefore, a local maximum representing the mean elevation of the plot was expected to be found within the residual signal. Further analysis concentrated on waveform points returning at a time later than the center of the shrub waveform detected on the previous step. In other words, examined points of the residual signal were the ones on the right side of the Gaussian center identified in the previous step. On these residual points a moving mask was applied, where the center point was compared to the immediate left and right neighbors (i.e. a derivative function). Points with higher values than both of the immediate neighbors were considered to be local maxima. After all local maxima were extracted, the one whose intensity at the residual (not the original) signal was highest was selected as the ground peak unless that maximum intensity was below a noise threshold. The noise threshold was identified as the maximum intensity of the first 150 bins in the waveform. In the case the selected local maximum was lower than the noise threshold, the ground was not selected from this process; instead, the ground peak was identified as the largest peak in the waveform. In other words, this verified that the previously extracted shrub was actually the ground since there was no other strong peak at a later return.

An example demonstrating graphically the implementation of the method is shown in Figure 3. The original signal is depicted in light gray color. The maximum intensity point is identified (time bin = 18) and the Gaussian function that fits best the waveform signal is selected (dashed line). By examining the residual signal (dashed-dotted line), the ground peak is assigned as the maximum value of the lower height.

The height accuracy of the identified ground peak was evaluated using reference data from discrete return LiDAR data. A further comparison took place between the ground peaks identified by the PCF method, and those identified by existing methods such as GD and FICA. The parameter settings for GD and FICA were identified after testing multiple options. The optimal settings for GD were a smoothing filter width set to 0 m and maximum iteration number set to 60, and for FICA the smoothing filter width was assigned the value of 0.3 m, the threshold for peak identification was set to 0.5 and the number of clusters was 4.

6.2. Landsat/LVIS fusion classifier

The PCF method was designed for dense shrub plots. It may not obtain better results than GD or FICA in other land cover types, such as grass, short trees, or very sparse shrub. In other words, the strongest peak may be the ground and not the vegetation peak. In that case, the height of the PCF-detected ground peak would be lower than the reference data as the Gaussian function representing the correct ground reflection would already be eliminated. Therefore, it was necessary to investigate the applicability of the PCF method and restrict its application in plots that could benefit from it.

A MAIC was built using metrics from both waveform and the corresponding TM image pixels. The classifier targeted plots where the PCF method outperformed the GD method in the ground peak identification with the aforementioned metrics. Due to the original design of the PCF method, we only examined the plots with RH100 less than 10 m. The process flow of the MAIC process is shown in Figure 4. Input metrics of the classifier were extracted from both waveform data and the multispectral dataset. The reference binary classes were determined by comparing the performance of the PCF method and GD. Depending on TM data availability, four different classification schemes were used in order to evaluate the data contribution in the classification. In Scheme 1, only the LVIS waveform metrics were used for the classification. Metrics from leaf-off and leaf-on season TM images were integrated with LVIS waveform metrics in Schemes 2 and 3, respectively. Finally, all metrics from TM leaf-on and leaf-off season and LVIS waveform were examined in Scheme 4. The input datasets in each scheme were then split into training and testing datasets. MAIC was built based on Random Forest (RF; Breiman 2001). The accuracy assessment and metrics importance were reported.

6.2.1. Metric extraction from LVIS waveform and Landsat TM images

Metrics from waveform and Landsat TM images were extracted as inputs for the MAIC process (Table 2). Waveform metrics were extracted from both the residual signal and the original waveform (Figure 2). By removing the canopy reflection from the original waveform, the residual signal between the largest peak in the original waveform and the end of the waveform was expected to capture the reflection from the ground. In addition, metrics from the original waveform captured the canopy structure, which indicated the intensity of the occlusion effect from the canopy. For example, stronger canopy reflection will lead to a weaker ground return.

Metrics were also extracted from TM pixels corresponding to the center locations of the sampled pulses. In addition to the spectral bands, different vegetation indices were calculated for each scene, such as Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Modified Soil-Adjusted Vegetation Index (MSAVI), which reduced the soil impact in areas with large amounts of exposed soil.

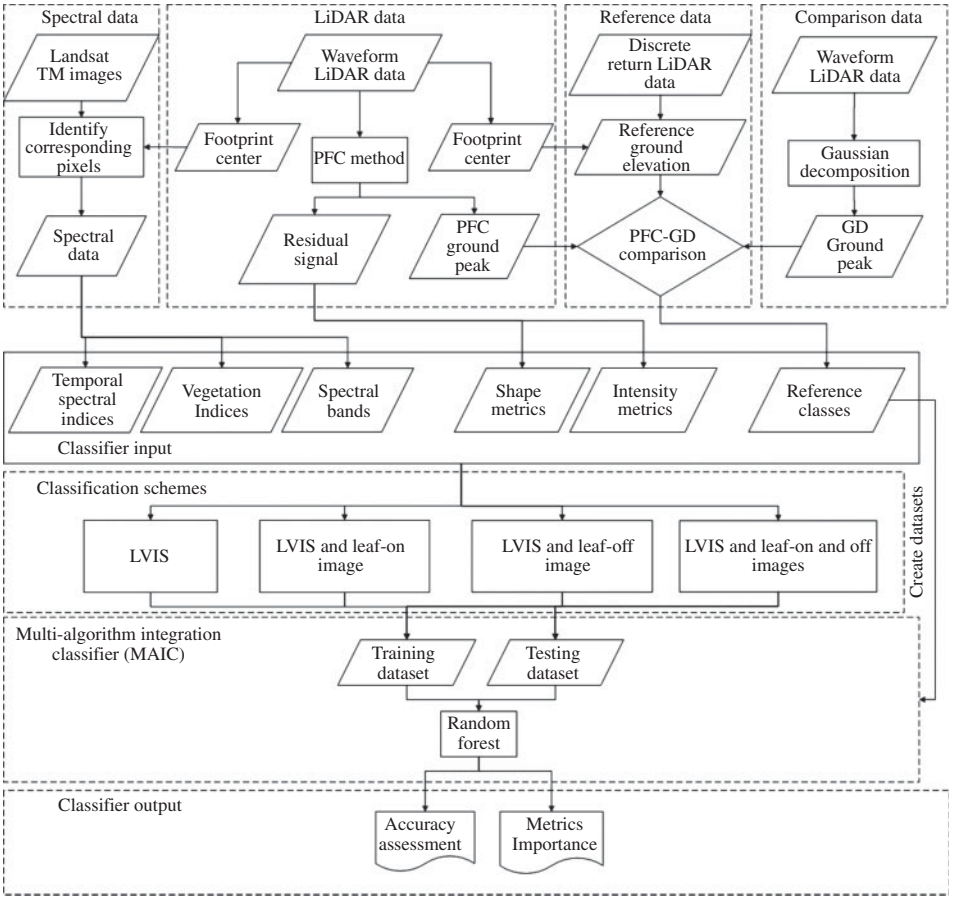

Figure 4. Flowchart of the MAIC.

The vegetation indices have been shown to correlate with canopy coverage (Qi et al. 1994; Carlson and Ripley 1997; Purevdorj et al. 1998; Gitelson et al. 2002; Jiang et al. 2006). In addition, the difference of reflectance and the difference of vegetation indexes between the two TM scenes were also calculated in order to amplify the reflectance difference in shrubs. Table 3 shows the metrics extracted from the two TM images.

6.2.2. RF in the MAIC process

RF, a method that assembles decision tree classifiers, has been shown to be a powerful classification method of remotely sensed and ecological datasets (Cutler et al. 2007; Chan and Paelinckx 2008; Timm and McGarigal 2012). As a nonparametric method, RF does not have any data distribution assumptions. The large number of trees used in RF minimizes the overfitting problem that exists in a single decision tree (Breiman 2001; Pal 2005). During the process of the classification in RF, a large number of trees (normally between 100 and 1000) are built. In each tree, a random subset of training data is selected using the bagging strategy, and a random subset of the predicting features is used for each tree split. RF determines the classification results by a voting procedure from the all

Table 2. Metrics calculation from the possible ground reflection in a waveform.

Name	Calculation	Description
Metrics from residual waveform		
RISum	$\sum$ Residual intensity	The intensity sum of a residual waveform
RIMax	Max(Residual intensity)	The maximum intensity of a residual waveform
RIMaxRatio	$\frac{\text{Max(Residual intensity)}}{\text{Max(Waveform Intensity)}}$	The ratio between the maximum intensity of a residual waveform and the maximum intensity of total waveform
RISumRatio	$\frac{\sum \text{Residual intensity}}{\sum \text{Total intensity}}$	The ratio between the intensity summation residual waveform and a total waveform
RAngle	$\frac{\text{Max intensity} - \text{Residual peak}}{\text{L(Residual peak)} - \text{L(max intensity)}}$	The relationship between largest peak in a waveform and peak in residual signal
Metrics from original waveform		
Leading edge	L(treetop)−L(first half max intensity)	Distance between location of treetop signal and location of first half max intensity
Trailing edge	L(last half max intensity)−L(end)	Distance between location of last half max intensity and last signal larger than background noise
Intensity variance	Var(Intensity)	Variance of waveform intensity

decision trees. Each tree in RF only uses approximately two-thirds of the data for training. The out-of-bag classification result is calculated by the other one-third of the data (Liaw and Wiener 2002). Moreover, RF also produces importance metrics (i.e. decrease of the mean square error) that shows the contribution of each metric to the classification. The number of trees in our RF implementation was set to 1000. In this study, the 10 most important metrics were used for the classification. The importance metric in the RF was the decrease of mean square error (Liaw and Wiener 2002).

6.2.3. Accuracy assessment and validation of the MAIC

The bias, which is the summation of the errors, Mean Absolute Error (MAE) and RMSE were reported. The classification accuracy (both overall accuracy and confusion matrix) of the MAIC process was reported for each classification scheme to allow comparisons.

7. Results and discussion

7.1. Ground peak identification results using PCF method and benchmark comparison

The PCF method was applied in peak identification for shrub waveforms. GD and FICA were also used for peak identification for comparison purposes. The reference dataset included the 1000 shrub plots. The ground identification results from the three methods

Table 3. Spectral and phonological metrics derived from Landsat TM images.

Metrics name	Description
S1b1, S1b2, … S1b7 S2b1, S2b2, … S2b7	Radiance of six spectral bands at different Scenes, S1 is the scene on April 17, S2 is for July 15 (thermal and panchromatic bands were excluded)
NDVIS1, NDVIS2	Normalized Difference Vegetation Index at different scenes: $\frac{\text{NIR}-\text{Red}}{\text{NIR}+\text{Red}}$
EVIS1, EVIS2	Enhanced Vegetation Index for different scenes: $G\frac{\text{NIR}-\text{Red}}{\text{NIR}+C_1\text{Red}-C_2\text{Blue}+L}$ $L = 1$, $C1 = 6$, $C2 = 7.5$, and $G = 2.5$
MSAVIS1, MSAVIS2	Modified Soil-adjusted Vegetation Index at different scenes: $\frac{2\text{NIR}+1-\sqrt{2(\text{NIR}+1)^2-8(\text{NIR}-\text{Red})}}{2}$ (Qi et al. 1994)
B1SpecDiff, … B6SpecDiff	Bands at S1 – Bands at S2
NDVIDiff	NDVI at S1 – NDVI at S2
EVIDiff	EVI at S1 – NDVI at S2
MSAVIDiff	MSAVI at S1 – MSAVI at S2

are shown in Table 4. The PCF method offers significant improvements over the other two methods in all accuracy metrics. The GD had a significantly higher bias, which meant that, in general, GD tended to overestimate the ground elevation. This overestimation is probably attributed to GD misinterpreting the shrub peak as the ground peak.

The FICA method performed better since it adopts a peak identification strategy that can detect small ground peaks even when they overlap with strong vegetation reflection. Therefore, FICA obtained better results (RMSE = 2.81 m) than GD (RMSE = 3.36 m). Still the PCF method had smaller MAE, which was 0.88 m less than that of FICA (RMSE of PCF was 0.89 m smaller). Moreover, the MAE of the PCF was less than 50% of the MAE in GD. The RMSE of the PCF method (1.92 m) was also much smaller than that of GD (3.36 m). The PCF method reduced the impact of the canopy reflection by deducting the Gaussian function, which represented the vegetation canopy. Therefore, even though the ground peak was hidden into the vegetation reflection, it was more accurately detected.

A detailed error distribution of the three ground peak identification methods is shown in Figure 5. The errors of all methods followed a normal distribution according to Kolmogorov–Smirnov test with a significance level of 0.05. In GD, more than 70% of the

Table 4. Ground identification results from PCF, GD, and FICA methods.

	PCF (m)	FICA (m)	GD (m)
Bias	0.50	1.07	2.84
MAE	1.41	2.29	2.84
RMSE	1.92	2.81	3.36

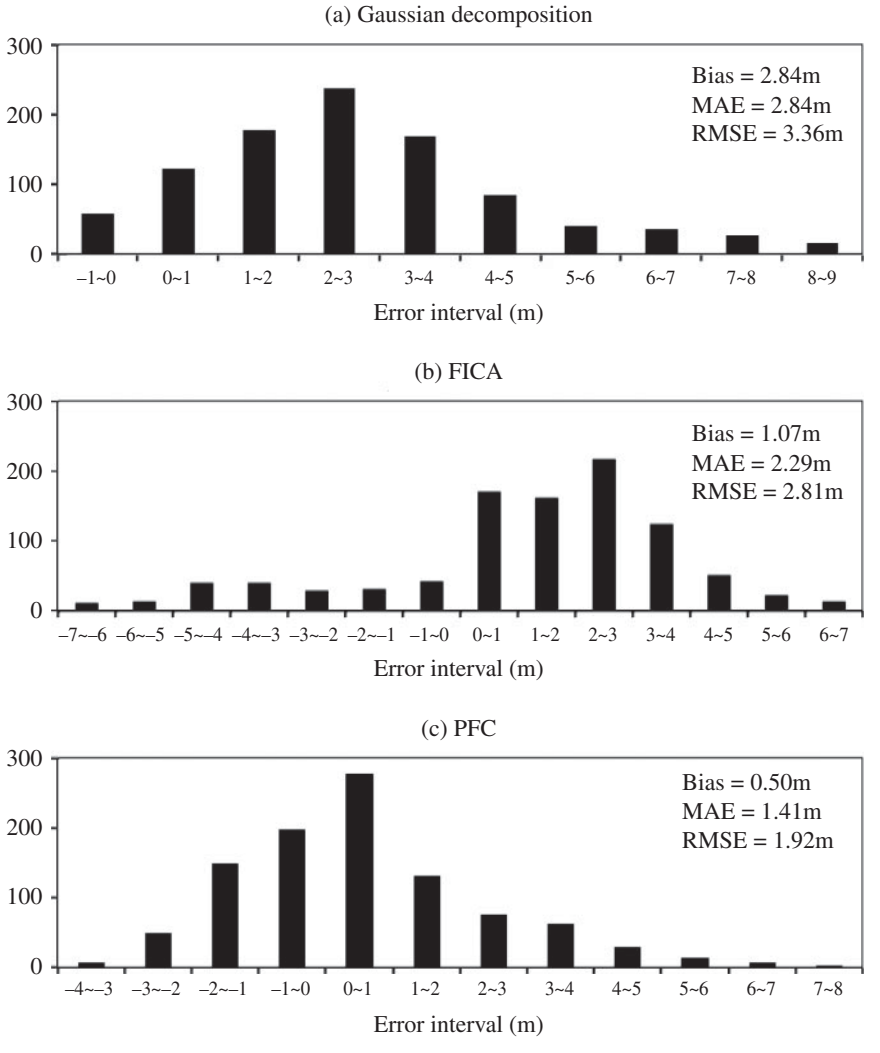

Figure 5. Error distribution of GD, FICA, and PCF methods.

plots had an error between 0–4 m. The error of FICA also underestimated ground peak location and had negative errors, which may be due to false identification of the background noise as the ground peak. In the PCF errors, more than 75% of the plots were between −2 m and 2 m. The error distribution showed that the ground peak in shrub plots was not overestimated in general. The relatively symmetrical distribution of the errors around zero indicated that PCF not only had better accuracy but better precision than the other two methods.

7.2. MAIC results in four classification schemes

Features were extracted from both LVIS and Landsat leaf-on and leaf-off images to perform a binary classification, indicating whether the PCF method would outperform the GD method, which is a popular method for ground identification in the previous large

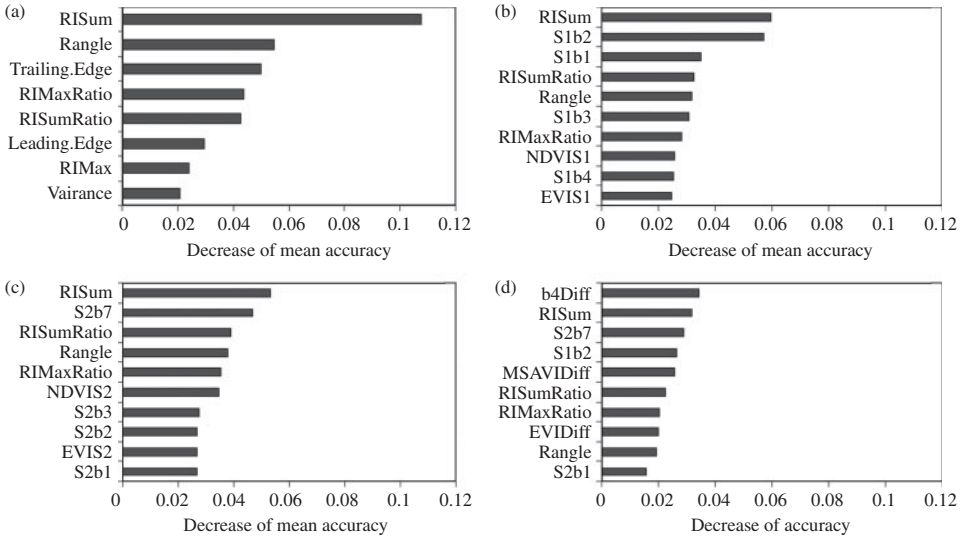

Figure 6. Important metrics in the MAIC process: (a) Scheme 1, (b) Scheme 2; (c) Scheme 3, (d) Scheme 4. For Schemes 2, 3, and 4 only the 10 most important metrics are shown.

footprint waveform LiDAR applications (Drake et al. 2002b; Anderson et al. 2006; Sun et al. 2008). First, the metrics importance was calculated, which then provided guidance for the feature selection.

7.2.1. Feature importance in MAIC

One of the benefits of RF classifiers is their ability to evaluate the importance of each input. Figure 6a–d show the selected most important metrics that were used in each of the four classification schemes. In Schemes 1, 2, and 3, *RISum*, which was the summation of the intensity in the residual signal, obtained the highest importance, as shown in Figure 6a, 6b, and 6c. In a waveform, the residual signal after the deduction of a Gaussian function would contain the ground peak, which would have a larger intensity sum than that of the false peaks created from background noise. However, if no ground peak existed in the residual signal, the sum of the intensity would be inevitably small. Similarly, RISum ratio was the relative measurement as RISum was also important in the classification. Spectral information also showed high importance in Schemes 2, 3, and 4. For example, in Scheme 4, the radiance value difference between the two different TM images in Near Infrared (NIR) (i.e. b4Diff) became the most important metric. Other spectral bands, such as S1b2 and S2b7, also showed high importance when combining LVIS data and both TM images together, as shown in Figure 6d. Those highly important metrics providing key information in discriminating the plots, where the ground peak location can be improved by the PCF method, may also be useful in the other studies.

7.2.2. Classification results

The results of the accuracy assessment of the four binary classification procedures are shown in Figure 7. It can be seen from the figure that the overall accuracy of all four schemes is high, suggesting that a site selection process for PCF application can be

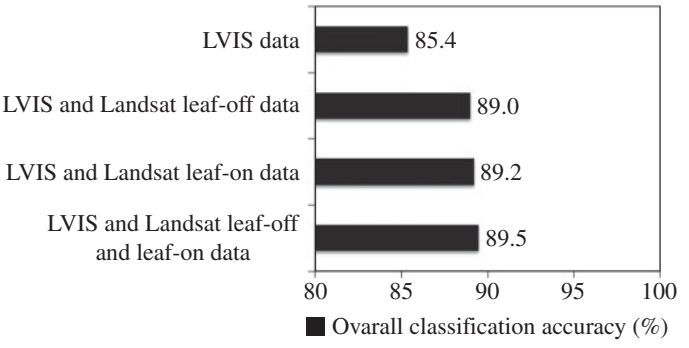

Figure 7. Classification accuracy for each scheme.

successfully implemented. The accuracy of Scheme 1 that uses only LVIS data was sufficiently high (85.4%) to allow the implementation of this method even in the absence of TM imagery. Also, the different seasonality of the TM image or the integration of images from multiple seasons did not improve classification accuracy over using a single Landsat image.

To investigate further the propagation of the binary misclassification errors to ground detection errors (i.e. errors of the MAIC process in the ground peak identification) Figure 8 was created. For each of the strata discussed in Section 5 an accuracy assessment took place on the MAIC classification prediction. It was encouraging to see that strata with smaller binary classification accuracy, in other words, where the MAIC process may have selected PCF instead of GD or vice versa, were present in bins that the actual ground detection performance difference between PCF and GD was small (strata S6 and S7). In bins that the difference between using one method over the other was substantial (strata S1, S2, S3, S9,

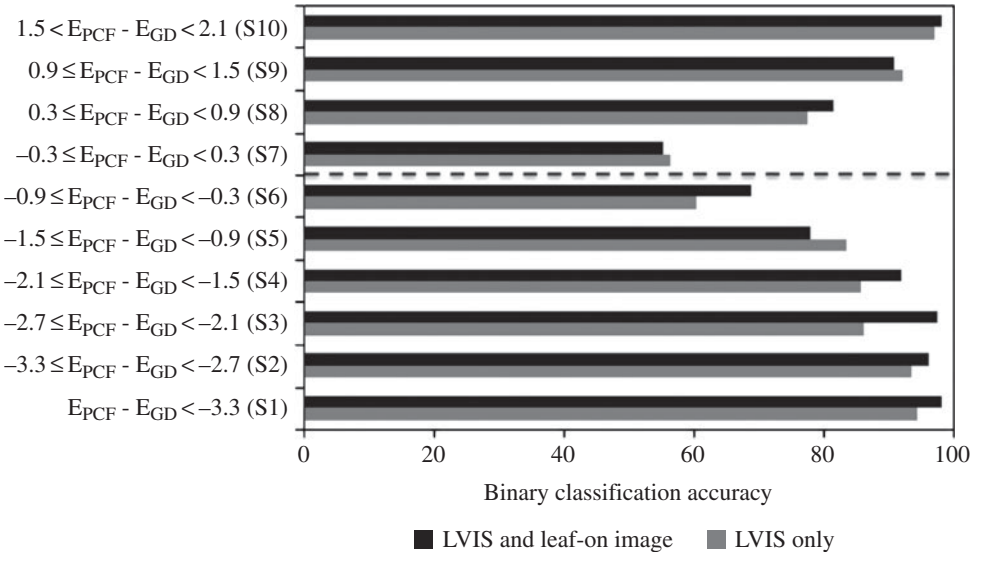

Figure 8. Accuracy of the MAIC process in different strata in Schemes 1 and 3. Strata below dash line are in class C1; other strata are in class C2, following Equation 3.

Table 5. Overall Accuracy (OA) of the MAIC for each class.

	Shrub	Grass	Deciduous	Coniferous	Urban
OA	85%	97%	88%	86%	100%

and S10), the binary classification accuracy was high thus allowing the selection of the more accurate PCF or GD method. Figure 8 also contrasts performance differences between Schemes 1 and 3. The similarity of both patterns suggests that the above conclusions are applicable to either LVIS-only classifiers or ones that fuse LVIS data with leaf-on Landsat imagery.

7.3. Integration of PCF and GD methods using the MAIC

In order to assess the practical benefits of the MAIC process, the 1456 reference points from multiple classes were revisited. Two objectives guided this process: (1) assess improvements in the shrub class that the PCF algorithm was optimized for, and (2) ensure that if the PCF method is applied to nonshrub classes it is done in a selective manner that does not cause overall degradation in the results. Three scenarios were examined, the exclusive use of GD or PCF and their fusion guided by the MAIC process. The MAIC process involved in this experiment assumed availability of Landsat leaf-on and leaf-off images along with the LVIS dataset (Scheme 4). In the fused scenario, the PCF method was only applied on plots where the MAIC process indicated that PCF would be more accurate for ground peak location. The GD method was applied on all other plots. The binary classification results of MAIC, which indicated whether PCF or GD method obtained better ground peak identification accuracy, are shown in Table 5 for each class.

After applying the MAIC results to the ground identification procedure, the RMSE of the three scenarios for the five land cover classes are recalculated, as shown in Table 6. Both of our objectives were met as the MAIC-based fusion: (1) improved results in the shrub class, and (2) produced at least as accurate results as the PCF or the GD approaches in all other classes. More specifically, Table 7 shows the results only for the shrub plots, using PCF, GD, and the fused algorithm. It is important to note here that the MAIC-based process is different than a general land cover/use classifier. The latter would select a single algorithm (PCF or GD) to be applied on all plots in a given class (e.g. all grass plots), while the MAIC fusion allows on demand integration of the two methods (PCF or GD) for each individual plot rather than an entire class.

Table 6. RMSE (m) of PCF, GD, and their fusion algorithm in ground identification.

	Shrub	Grass	Deciduous	Coniferous	Urban
PCF	1.92	0.67	5.46	7.62	1.16
GD	3.36	1.54	6.85	5.40	0.45
PCF + GD	1.79	0.60	5.34	5.40	0.42

Table 7. Errors for the shrub plots using PCF, GD, and the fused algorithm.

	PCF	GD	PCF + GD
Bias	0.5	2.84	0.39
MAE	1.41	2.84	1.28
RMSE	1.92	3.36	1.79

8. Conclusion

In this study, a novel PCF method peak was introduced for identifying ground peaks in LiDAR waveform datasets acquired in dense shrub areas. The proposed method improved detection accuracy of the ground peak when compared to existing methods such as the GD and theFICA. This improvement is substantial, especially compared to the currently prevailing GD method (reduced error almost in half). We are confident that our findings could generalize to other sites as a wide range of slopes was tested. This is particularly important for biodiversity, forest succession, and carbon sequestration studies, where accurate shrub height detection is essential. Ground identification is the most challenging task associated with height estimation, especially in dense shrubland. As the LVIS sensor tested in this study is expected to become space borne in the near future, the increase in data availability will multiply the benefits of our method.

In addition, for the first time, a MAIC is proposed, establishing the framework for dynamic integration of multiple ground peak detection algorithms. A proof of concept is discussed in this paper demonstrating potential benefits, especially for different vegetation coverage. The PCF method does not require any parameter optimization procedure, which is usually associated with time-consuming iterative computations. Thus, the algorithm is computationally efficient. In future implementations, algorithmic execution speed could also benefit from this integration, especially considering the high data volume. For example, in this paper, the PCF method is substantially faster to execute than the GD method, so in cases where results are equally accurate preference could be given to PCF.

Acknowledgments

We would like to thank Dr Blair and Dr Hofton for providing the LVIS dataset as part of that grant and Dr Beier with Mr Wiley for helpful discussions.

Funding

This work was supported through NASA's Biodiversity Program [grant number NNX09AK16G].

References

Anderson, J., M. E. Martin, M.-L. Smith, R. O. Dubayah, M. A. Hofton, P. Hyde, B. E. Peterson, J. B. Blair, and R. G. Knox. 2006. "The Use of Waveform Lidar to Measure Northern Temperate Mixed Conifer and Deciduous Forest Structure in New Hampshire." *Remote Sensing of Environment* 105 (3): 248–261. doi:10.1016/j.rse.2006.07.001.

Anderson, J. E., L. C. Plourde, M. E. Martin, B. H. Braswell, M. L. Smith, R. O. Dubayah, M. A. Hofton, and J. B. Blair. 2008. "Integrating Waveform Lidar with Hyperspectral Imagery for Inventory of a Northern Temperate Forest." *Remote Sensing of Environment* 112 (4): 1856–1870. doi:10.1016/j.rse.2007.09.009.

Blair, J. B., M. A. Hofton, and D. L. Rabine. 2006. *Processing of NASA LVIS Elevation and Canopy (LGE, LCE and LGW) Data Products, Version 1.01*. http://lvis.gsfc.nasa.gov.

Blair, J. B., D. L. Rabine, and M. A. Hofton. 1999. "The Laser Vegetation Imaging Sensor: A Medium-altitude, Digitisation-only, Airborne Laser Altimeter for Mapping Vegetation and Topography." *ISPRS Journal of Photogrammetry and Remote Sensing* 54 (2–3): 115–122. doi:10.1016/S0924-2716(99)00002-7.

Boudreau, J., R. F. Nelson, H. A. Margolis, A. Beaudoin, L. Guindon, and D. S. Kimes. 2008. Regional Aboveground Forest Biomass Using Airborne and Spaceborne LiDAR in Québec." *Remote Sensing of Environment* 112 (10): 3876–3890. doi:10.1016/j.rse.2008.06.003.

Breiman, L., 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. doi:10.1023/A:1010933404324.

Carlson, T. N., and D. A. Ripley. 1997. "On the Relation between NDVI, Fractional Vegetation Cover, and Leaf Area Index." *Remote Sensing of Environment* 62 (3): 241–252. doi:10.1016/S0034-4257(97)00104-1.

Chan, J. C., and D. Paelinckx. 2008. Evaluation of Random Forest and Adaboost Tree-based Ensemble Classification and Spectral Band Selection for Ecotope Mapping Using Airborne Hyperspectral Imagery." *Remote Sensing of Environment* 112 (6): 2999–3011. doi:10.1016/j.rse.2008.02.011.

Chander, G., B. L. Markham, and D. L. Helder. 2009. "Summary of Current Radiometric Calibration Coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors." *Remote Sensing of Environment* 113 (5): 893–903. doi:10.1016/j.rse.2009.01.007.

Chen, Q., 2010a. "Assessment of Terrain Elevation Derived from Satellite Laser Altimetry Over Mountainous Forest Areas Using Airborne Lidar Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 65 (1): 111–122. doi:10.1016/j.isprsjprs.2009.09.004.

Chen, Q., 2010b. "Retrieving Vegetation Height of Forests and Woodlands Over Mountainous Areas in the Pacific Coast Region Using Satellite Laser Altimetry." *Remote Sensing of Environment* 114 (7): 1610–1627. doi:10.1016/j.rse.2010.02.016.

Clark, M. L., D. B. Clark, and D. A. Roberts. 2004. "Small-footprint Lidar Estimation of Sub-canopy Elevation and Tree Height in a Tropical Rain Forest Landscape." *Remote Sensing of Environment* 91 (1): 68–89. doi:10.1016/j.rse.2004.02.008.

Cutler, D. R., T. C. Edwards Jr, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. "Random Forests for Classification in Ecology." *Ecology* 88 (11): 2783–2792. doi:10.1890/07-0539.1.

Drake, J. B., R. O. Dubayah, D. B. Clark, R. G. Knox, J. B. Blair, M. A. Hofton, R. L. Chazdon, J. F. Weishampel, and S. Prince. 2002a. "Estimation of Tropical Forest Structural Characteristics Using Large-footprint Lidar." *Remote Sensing of Environment* 79 (2–3): 305–319. doi:10.1016/S0034-4257(01)00281-4.

Drake, J. B., R. O. Dubayah, R. G. Knox, D. B. Clark, and J. B. Blair. 2002b. "Sensitivity of Large-footprint Lidar to Canopy Structure and Biomass in a Neotropical Rainforest." *Remote Sensing of Environment* 81 (2–3): 378–392. doi:10.1016/S0034-4257(02)00013-5.

Estornell, J., L. A. Ruiz, B. Velázquez-Martí, and T. Hermosilla. 2012. "Estimation of Biomass and Volume of Shrub Vegetation Using Lidar and Spectral Data in a Mediterranean Environment." *Biomass and Bioenergy* 46: 710–721. doi:10.1016/j.biombioe.2012.06.023.

Gitelson, A. A., Y. J. Kaufman, R. Stark, and D. Rundquist. 2002. "Novel Algorithms for Remote Estimation of Vegetation Fraction." *Remote Sensing of Environment* 80 (1): 76–87. doi:10.1016/S0034-4257(01)00289-9.

Glenn, N. F., L. P. Spaete, T. T. Sankey, D. R. Derryberry, S. P. Hardegree, and J. J. Mitchell. 2011. "Errors in Lidar-derived Shrub Height and Crown Area on Sloped Terrain." *Journal of Arid Environments* 75 (4): 377–382. doi:10.1016/j.jaridenv.2010.11.005.

Goetz, S., D. Steinberg, R. Dubayah, and B. Blair. 2007. "Laser Remote Sensing of Canopy Habitat Heterogeneity as a Predictor of Bird Species Richness in an Eastern Temperate Forest, USA." *Remote Sensing of Environment* 108 (3): 254–263. doi:10.1016/j.rse.2006.11.016.

Gould, S. B., N. F. Glenn, T. T. Sankey, and J. P. Mcnamara. 2013. "Influence of a Dense, Low-height Shrub Species on the Accuracy of a Lidar-derived Dem." *Photogrammetric Engineering and Remote Sensing* 79 (5): 421–431. doi:10.14358/PERS.79.5.421.

Hodgson, M. E., and P. Bresnahan. 2004. "Accuracy of Airborne Lidar-derived Elevation: Empirical Assessment and Error Budget." *Photogrammetric Engineering and Remote Sensing* 70 (3): 331–339. doi:10.14358/PERS.70.3.331.

Hofton, M. A., J. B. Minster, and J. B. Blair. 2000. "Decomposition of Laser Altimeter Waveforms." *IEEE Transactions on Geoscience and Remote Sensing* 38 (4 II): 1989–1996. doi:10.1109/36.851780.

Hyde, P., R. Dubayah, B. Peterson, J. B. Blair, M. Hofton, C. Hunsaker, R. Knox, and W. Walker. 2005. "Mapping Forest Structure for Wildlife Habitat Analysis Using Waveform Lidar: Validation of Montane Ecosystems." *Remote Sensing of Environment* 96 (3–4): 427–437. doi:10.1016/j.rse.2005.03.005.

Jakubowksi, M. K., Q. Guo, B. Collins, S. Stephens, and M. Kelly. 2013. "Predicting Surface Fuel Models and Fuel Metrics Using Lidar and Cir Imagery in a Dense, Mountainous Forest." *Photogrammetric Engineering and Remote Sensing* 79 (1): 37–49. doi:10.14358/PERS.79.1.37.

Jiang, Z., A. R. Huete, J. Chen, Y. Chen, J. Li, G. Yan, and X. Zhang. 2006. "Analysis of NDVI and Scaled Difference Vegetation Index Retrievals of Vegetation Fraction." *Remote Sensing of Environment* 101 (3): 366–378. doi:10.1016/j.rse.2006.01.003.

Jutzi, B., and U. Stilla. 2006. "Range Determination with Waveform Recording Laser Systems Using a Wiener Filter." *ISPRS Journal of Photogrammetry and Remote Sensing* 61 (2): 95–107. doi:10.1016/j.isprsjprs.2006.09.001.

Lee, S., W. Ni-Meister, W. Yang, and Q. Chen. 2011. "Physically based Vertical Vegetation Structure Retrieval from ICESAT Data: Validation using LVIS in White Mountain National Forest, New Hampshire, USA." *Remote Sensing of Environment* 115 (11): 2776–2785. doi:10.1016/j.rse.2010.08.026.

Liaw, A., and M. Wiener. 2002. "Classification and Regression by Random Forest." *R news* 2: 18–22.

Mallet, C., and F. Bretar. 2009. "Full-waveform Topographic Lidar: State-of-the-art." *ISPRS Journal of Photogrammetry and Remote Sensing* 64 (1): 1–16. doi:10.1016/j.isprsjprs.2008.09.007.

Martinuzzi, S., L. A. Vierling, W. A. Gould, M. J. Falkowski, J. S. Evans, A. T. Hudak, and K. T. Vierling. 2009. "Mapping Snags and Understory Shrubs for a Lidar-based Assessment of Wildlife Habitat Suitability." *Remote Sensing of Environment* 113 (12): 2533–2546. doi:10.1016/j.rse.2009.07.002.

Means, J. E., S. A. Acker, D. J. Harding, J. B. Blair, M. A. Lefsky, W. B. Cohen, M. E. Harmon, and W. A. Mckee. 1999. "Use of Large-footprint Scanning Airborne Lidar to Estimate Forest Stand Characteristics in the Western Cascades of Oregon." *Remote Sensing of Environment* 67 (3): 298–308. doi:10.1016/S0034-4257(98)00091-1.

Mitchell, J. J., N. F. Glenn, T. T. Sankey, D. R. Derryberry, M. O. Anderson, and R. C. Hruska, 2011. Small-footprint Lidar Estimations of Sagebrush Canopy Characteristics." *Photogrammetric Engineering and Remote Sensing* 77 (5): 521–530. doi:10.14358/PERS.77.5.521.

Nyland, R. D., W. C. Zipperer, and D. B. Hill. 1986. "The Development of Forest Islands in Exurban Central New York State." *Landscape and Urban Planning* 13 (C): 111–123. doi:10.1016/0169-2046(86)90016-2.

Pal, M., 2005. "Random Forest Classifier for Remote Sensing Classification." *International Journal of Remote Sensing* 26 (1): 217–222. doi:10.1080/01431160412331269698.

Pinotti, B. T., C. P. Pagotto, and R. Pardini. 2012. "Habitat Structure and Food Resources for Wildlife Across Successional Stages in a Tropical Forest." *Forest Ecology and Management* 283: 119–127. doi:10.1016/j.foreco.2012.07.020.

Popescu, S. C., and K. Zhao. 2008. "A Voxel-based Lidar Method for Estimating Crown Base Height for Deciduous and Pine Trees." *Remote Sensing of Environment* 112 (3): 767–781. doi:10.1016/j.rse.2007.06.011.

Purevdorj, T. S., R. Tateishi, T. Ishiyama, and Y. Honda. 1998. "Relationships between Percent Vegetation Cover and Vegetation Indices." *International Journal of Remote Sensing* 19 (18): 3519–3535. doi:10.1080/014311698213795.

Qi, J., A. Chehbouni, A. R. Huete, Y. H. Kerr, and S. Sorooshian. 1994. "A Modified Soil Adjusted Vegetation Index." *Remote Sensing of Environment* 48 (2): 119–126. doi:10.1016/0034-4257(94)90134-1.

Ramankutty, N., and J. A. Foley. 1999. "Estimating Historical Changes in Global Land Cover: Croplands from 1700 to 1992." *Global Biogeochemical Cycles* 13 (4): 997–1027. doi:10.1071/WF08173.

Rango, A., M. Chopping, J. Ritchie, K. Havstad, W. Kustas, and T. Schmugge. 2000. "Morphological Characteristics of Shrub Coppice Dunes in Desert Grasslands of Southern New Mexico Derived from Scanning Lidar." *Remote Sensing of Environment* 74 (1): 26–44. doi:10.1016/S0034-4257(00)00084-5.

Riaño, D., E. Chuvieco, S. L. Ustin, J. Salas, J. R. Rodríguez-Pérez, L. M. Ribeiro, D. X. Viegas, J. M. Moreno, and H. Fernández. 2007. "Estimation of Shrub Height for Fuel-type Mapping Combining Airborne Lidar and Simultaneous Color Infrared Ortho Imaging." *International Journal of Wildland Fire* 16 (3): 341–348.

Saatchi, S., K. Halligan, D. G. Despain, and R. L. Crabtree. 2007. "Estimation of Forest Fuel Load from Radar Remote Sensing." *IEEE Transactions on Geoscience and Remote Sensing* 45 (6): 1726–1740. doi:10.1109/TGRS.2006.887002.

Skowronski, N., K. Clark, R. Nelson, J. Hom, and M. Patterson. 2007. "Remotely Sensed Measurements of Forest Structure and Fuel Loads in the Pinelands of New Jersey." *Remote Sensing of Environment* 108 (2): 123–129. doi:10.1016/j.rse.2006.09.032.

Smart, L. S., J. J. Swenson, N. L. Christensen, and J. O. Sexton. 2012. "Three-Dimensional Characterization of Pine Forest Type and Red-cockaded Woodpecker Habitat by Small-footprint, Discrete-return Lidar." *Forest Ecology and Management* 281: 100–110. doi:10.1016/j.foreco.2012.06.020.

Smith, B. E., P. Marks, and S. Gardescu. 1993. "Two Hundred Years of Forest Cover Changes in Tompkins County, New York." *Bulletin of the Torrey Botanical Club* 120 (3): 229–247. doi:10.2307/2996988.

Spaete, L. P., N. F. Glenn, D. R. Derryberry, T. T. Sankey, J. J. Mitchell, and S. P. Hardegree. 2011. "Vegetation and Slope Effects on Accuracy of a Lidar-derived Dem in the Sagebrush Steppe." *Remote Sensing Letters* 2 (4): 317–326. doi:10.1080/01431161.2010.515267.

Streutker, D. R., and N. F. Glenn. 2006. "Lidar Measurement of Sagebrush Steppe Vegetation Heights." *Remote Sensing of Environment* 102 (1–2): 135–145. doi:10.1016/j.rse.2006.02.011.

Su, J. G., and E. W. Bork. 2007. "Characterization of Diverse Plant Communities in Aspen Parkland Rangeland Using Lidar Data." *Applied Vegetation Science* 10 (3): 407–416. doi:10.1111/j.1654-109X.2007.tb00440.x.

Sun, G., K. J. Ranson, Z. Guo, Z. Zhang, P. Montesano, and D. Kimes. 2011. "Forest Biomass Mapping from Lidar and Radar Synergies." *Remote Sensing of Environment* 115 (11): 2906–2916. doi:10.1016/j.rse.2011.03.021.

Sun, G., K. J. Ranson, V. I. Kharuk, and K. Kovacs. 2003. "Validation of Surface Height from Shuttle Radar Topography Mission Using Shuttle Laser Altimeter." *Remote Sensing of Environment* 88 (4): 401–411. doi:10.1016/j.rse.2003.09.001.

Sun, G., K. J. Ranson, D. S. Kimes, J. B. Blair, and K. Kovacs. 2008. "Forest Vertical Structure from GLAS: An Evaluation Using LVIS and SRTM Data." *Remote Sensing of Environment* 112 (1): 107–117. doi:10.1016/j.rse.2006.09.036.

Swanson, M. E., J. F. Franklin, R. L. Beschta, C. M. Crisafulli, D. A. Dellasala, R. L. Hutto, D. B. Lindenmayer, and F. J. Swanson. 2011. "The Forgotten Stage of Forest Succession: Early-successional Ecosystems on Forest Sites." *Frontiers in Ecology and the Environment* 9 (2): 117–125. doi:10.1890/090157.

Timm, B. C., and K. Mcgarigal. 2012. "Fine-scale Remotely-sensed Cover Mapping of Coastal Dune and Salt Marsh Ecosystems at Cape Cod National Seashore Using Random Forests." *Remote Sensing of Environment* 127: 106–117. doi:10.1016/j.rse.2012.08.033.

Wang, C., and N. F. Glenn. 2009. "Estimation of Fire Severity Using Pre- and Post-fire Lidar Data in Sagebrush Steppe Rangelands." *International Journal of Wildland Fire* 18 (7): 848–856. doi:10.1071/WF08173.

Wang, K., S. E. Franklin, X. Guo, and M. Cattet. 2010. "Remote Sensing of Ecology, Biodiversity and Conservation: A Review from the Perspective of Remote Sensing Specialists." *Sensors* 10 (11): 9647–9667. doi:10.3390/s101109647.

Wasser, L., R. Day, L. Chasmer, and A. Taylor. 2013. "Influence of Vegetation Structure on Lidar-derived Canopy Height and Fractional Cover in Forested Riparian Buffers during Leaf-off and Leaf-on Conditions." *PLoS ONE* 8 (1): e54776. doi:10.1371/journal.pone.0054776.