ICESat waveform-based land-cover classification using a curve matching approach

Yuhong Zhou\textsuperscript{a}, Fang Qiu\textsuperscript{a,b,}\textsuperscript{*}, Ali A. Al-Dosari\textsuperscript{c}, and Mohammed S. Alfarhan\textsuperscript{d}

\textsuperscript{a}Department of Geospatial Science, University of Texas at Dallas, Richardson, TX 75080-3021, USA; \textsuperscript{b}Centre of Geospatial Information, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China; \textsuperscript{c}Department of Geography, King Saud University, Riyadh 11451, Saudi Arabia; \textsuperscript{d}Space Research Institute, King Abdulaziz City for Sciences and Technology, Riyadh 11442, Saudi Arabia

(Received 12 February 2014; accepted 7 November 2014)

Light detection and ranging waveforms record the entire one-dimensional backscattered signal as a function of time within a footprint, which can potentially reflect the vertical structure information of the above-ground objects. This study aimed to explore the potential of the Geoscience Laser Altimeter System sensor on board the Ice, Cloud, and Land Elevation Satellite to perform land-cover classification by using only the profile curve of the full waveform. For this purpose, a curve matching method based on Kolmogorov–Smirnov (KS) distance was developed to measure the curve similarity between an unknown waveform and a reference waveform. A set of reference waveforms were first extracted from the training data set based on a principal component analysis (PCA). The unknown waveform was then compared with individual reference waveforms derived using KS distance and assigned to the class with the closest similarity. The results demonstrated that the KS distance-based land-cover classification using the waveform curve was able to achieve an overall accuracy of 87.2\% and a kappa coefficient of 0.80. It outperformed the widely adopted rule-based method using Gaussian decomposition parameters by 3.5\%. The research also indicated that the PCA-selected reference waveforms achieved substantially better results than randomly selected reference waveforms.

1. Introduction

Land cover refers to the types of observed biophysical features covering the earth’s surface, such as water, trees, grass, roads, buildings, and bare ground. Land cover is increasingly influenced and altered by human activities, causing it to constantly change over time. These land-cover changes can impose great pressure on the environment leading to widespread problems such as soil nutrient depletion (Abbasi, Zafar, and Khan 2007), landslides (Abbasi, Zafar, and Khan 2007), and global warming (Dallmeyer and Claussen 2011). Consequently, land cover has been monitored and assessed by geographers, ecologists, and other scientists in order to understand the ongoing processes driving its change and identify strategies to achieve sustainable development.

A variety of data sources and methods have been utilized to derive land-cover information (Grey, Luckman, and Holland 2003). The traditional field survey approach is costly in terms of time and labour, inhibited by limited accessibility to many areas of land, and highly dependent on subjective human evaluation (Wu and Xing

\textsuperscript{*}Corresponding author. Email: ffqiu@utdallas.edu

© 2014 Taylor & Francis
Remote-sensing imagery, including multispectral, hyperspectral, and light detection and ranging (lidar) data, provides an effective alternative to a field survey because of its ability to efficiently gather digital data from vast and inaccessible areas on a regular, repetitive basis (Arroyo et al. 2010). Traditionally, remote sensing of land cover primarily relied on multispectral imagery (e.g. Landsat Thematic Mapper and Systeme Probatoire d’Observation de la Terre) (Jawak and Luis 2013). Its use for general land-cover classification has been widely studied (Harvey and Hill 2001; Li, Ustin, and Lay 2005), but poor spectral resolution because of a limited number of bands and wide bandwidths inhibits detailed land-cover classification, especially discriminating different geographical features that have similar spectral responses (Rosso, Ustin, and Hastings 2005). The spectral differences among different types of land cover may occur in very narrow regions of the spectrum that cannot be sensed by multispectral sensors. The advent of hyperspectral data offers a promising alternative for performing land-cover classification at a much more detailed level. The subtle spectral differences among different materials can be sensed by the hundreds of contiguous narrow bands (often 10 nm or less) of the hyperspectral data (Cho et al. 2010) allowing, for example, the discrimination of tree species or differentiation of mineral types. However, both multispectral and hyperspectral data can only provide spectral information of an object’s surface and may fail to separate objects that are spectrally similar but structurally different (Buddenbaum, Seeling, and Hill 2013). For example, roads and buildings are spectrally similar when they are constructed using the same material (such as concrete). Differentiation may be difficult, even with hyperspectral data.

Unlike multispectral and hyperspectral remote sensing, the recently available lidar technology can potentially provide structural information about an object through the three-dimensional positional measurements recorded by its return pulse (Sasaki et al. 2012). Lidar is an active remote-sensing technology based on ranging using visible, near-infrared, or short-wave infrared laser beams. Lidar data can be divided into two categories based on how the signal is recorded: discrete-return lidar and full-waveform lidar (Alexander et al. 2010). For discrete-return lidar, the continuous reflection signal is discretized into one or multiple returns. Conversely, full-waveform lidar records the entire backscattered signal of the objects along the path of the emitted laser using small time intervals (typically 1 ns) which determine the level of vertical detail (Farid et al. 2008; Zaletnyik, Laky, and Toth 2010; Pirotti 2010; Mallet et al. 2011). To perform land-cover classification, discrete-return lidar data are used independently or in conjunction with multispectral or hyperspectral data. For example, Vögtle and Steinele (2003) used lidar elevation data alone to classify urban features into trees, buildings, and terrain objects, achieving an overall accuracy of 91%. Although the discrimination between objects with distinct vertical structures is high, the ability of discrete-return lidar to separate objects with no vertical structure (e.g. bare land and grass) or similar vertical structures is limited. Consequently, numerous subsequent studies have integrated imagery data (either multispectral or hyperspectral) and discrete-return lidar data to improve classification performance (Sugumaran and Voss 2007; Zhou and Troy 2008; Sasaki et al. 2012).

Full-waveform lidar, which has become operational only recently, provides a promising new source to monitor land-cover conditions. The shape of a waveform is the result of interaction between the transmitted laser pulse and intercepted object surfaces within the lidar footprint (Alexander et al. 2010; Pirotti 2010). For example, a waveform over a building is primarily determined by the joint effects of the building’s
height, vertical structure, and the percentage of area covered by ground versus building within the footprint. A waveform over a tree is often determined by tree density, structure, and vegetation phase. Due to the finer vertical resolution, a lidar waveform offers an enhanced capability to reflect the vertical structures of geographical objects compared to the traditional discrete-return lidar that typically consists of six or fewer measurements separated by several metres in distance along its vertical profile (Zaletnyik, Laky, and Toth 2010; Zhang et al. 2011; Ussyshkin and Theriault 2011).

The large-footprint waveform data collected by the Geoscience Laser Altimeter System (GLAS) on board the Ice, Cloud, and Land Elevation Satellite (ICESat) is the most commonly used due to its availability free-of-charge, global coverage, and repeatable measurements (Ranson et al. 2004; Duong, Pfeifer, and Lindenbergh 2006; Pirotti 2010; Wu and Xing 2010; Zhang et al. 2011; Molijn, Lindenbergh, and Gunter 2011). Although originally designed for cryosphere research, ICESat waveforms have been widely used for various other studies, such as estimation of above-ground biomass (Lefsky et al. 2005; Ballhorn, Jubanski, and Siegert 2011; Boudreau et al. 2008; Duncanson, Niemann, and Wulder 2010; Nelson 2010), canopy vertical structure (Harding and Carabajal 2005; Lee et al. 2011; Sun et al. 2008; Duncanson, Niemann, and Wulder 2010; Lefsky et al. 2007; Pang et al. 2011; Rosette, North, and Suárez 2008; Wang et al. 2012, 2014), timber volume (Nelson et al. 2009; Ranson et al. 2007), and building height (Gong et al. 2011).

ICESat waveforms have also been used for land-cover classification by a number of researchers (Ranson et al. 2004; Duong, Pfeifer, and Lindenbergh 2006; Zhang et al. 2011; Wu and Xing 2010; Molijn, Lindenbergh, and Gunter 2011; Wang et al. 2012; Liu et al. 2014), since a waveform shape is largely determined by the vertical structures of the objects within the footprint. In general, the waveform shape is composed of one or many echoes, where one echo is a peak in the waveform that may correspond to an individual target encountered or a cluster of objects too close to be separated, while multiple echoes often represent multiple targets or one target with multiple components that are vertically separable within the footprint (Alexander et al. 2010). Consequently, the shapes for the waveforms over multiple objects can be very complex. To utilize the waveforms for land-cover classification, existing studies have focused on simplifying the complex full-waveform data. Shape-related summary metrics derived from the full waveform, such as the number of echoes, peak amplitude, echo width, skewness, and kurtosis, are used to represent the important characteristics of the waveform shapes (Ranson et al. 2004; Duong, Pfeifer, and Lindenbergh 2006; Zhang et al. 2011; Narayanan, Kim, and Sohn 2009; Zaletnyik, Laky, and Toth 2010; Wu and Xing 2010; Molijn, Lindenbergh, and Gunter 2011; Mallet et al. 2011; Wang et al. 2012). These metrics have had varied success in differentiating land-cover types. Because they aim to reduce the complexity of the waveform data, these metrics are not taking advantage of the entirety of information that the full waveform offers. The full-waveform curve provides a more comprehensive picture of the vertical structure of the intercepted surfaces than these summary metrics. This vertical structure information is potentially more valuable for land-cover classification by using a curve-matching-based approach, but has been scarcely investigated.

To complement the research based on summary metrics of the full waveforms, this study focuses on analysing the complete curve distribution. It investigates a novel approach derived from the Kolmogorov–Smirnov (KS) statistic (Burt and Barber 1996) for analysing the full ICESat waveform to distinguish land-cover types. The specific objectives of the study are (1) to quantify the similarity between waveform curves using a KS-based approach; (2) to select a set of waveforms as references
through principal component analysis (PCA); and (3) to determine whether analysing the complete curve distribution using the KS-based approach can achieve superior results for land-cover classification over a rule-based approach using Gaussian decomposed parameters.

In this study, the emphasis is on exploring the potential of full waveforms alone to differentiate objects having different vertical structures using a curve-matching-based approach. Consequently, spectral information was deliberately excluded, and a relatively level study area was used to minimize the complications of a varied terrain. Furthermore, only three general categories of land cover that have distinct vertical structures are used: buildings, trees, and open spaces. All land-cover types with a flat open surface such as bare land, water, roads, and parking lots are assigned to the open space category. They are not identified as separate categories since they have generally similar vertical waveform curves and it is very difficult to classify them using the waveform lidar alone. Similarly, there is no differentiation of building subtypes (such as residential or commercial) or tree subtypes (such as deciduous or coniferous). These very general land-cover categories were a necessity following from the use of ICESat waveform data, which do not allow a more detailed level of classification because of the large size of the ICESat footprint (Hilbert and Schmullius 2012). Albeit not optimal for land-cover classification since it was designed for other purposes, the free availability of its waveforms makes it an especially useful test bed for assessing curve-matching-based approaches. Furthermore, global coverage is available since it is a satellite-based system. If the effectiveness of using full-waveform curves alone for high-level land-cover classification can be achieved, future studies can then be conducted by incorporating spectral information of high spatial resolution (HSR) imagery with small-footprint waveform curves from commercial airborne lidar systems at the object level to achieve a more detailed land-cover classification or fusing hyperspectral imagery with pseudo-waveforms derived from discrete-return lidar to identify vegetation species using the same curve matching approach.

2. Study area and data

2.1. Study area

The study area is located in the Dallas, Texas metropolitan area, bounded approximately by 97.031° to 96.518° W and 32.553° to 32.996° N, where a large number of ICESat waveform data are available (Figure 1). The area is characterized by a mixture of mature residential neighbourhoods and commercial and industrial buildings.

ICESat waveforms are sensitive to surface terrain due to their large footprints (Hilbert and Schmullius 2012). For example, a waveform over open space with a flat surface (Figure 2(a)) usually presents a single echo with a narrow width because a high concentration of energy is reflected back at nearly the same time, while a waveform over open space with a sloped surface (Figure 2(b), not in our study area but was selected from a hilly area of San Diego) is stretched with multiple echoes due to rough terrain topography. The waveform widths can be noticeably wider compared to those of flat surface if terrain slopes are larger than 10° (Hilbert and Schmullius 2012). Since the majority of the study area is composed of flat surfaces (slope \( \leq 10° \)), the terrain is not expected to exert significant influence on the waveform shapes (Gong et al. 2011; Hilbert and Schmullius 2012; Duong et al. 2009) and therefore the classification.
Figure 1. Study area with waveform footprints.

Figure 2. (a) A waveform over open space with flat surface and (b) a waveform over open space with sloped surface.
2.2. ICESat/GLAS full waveform

Up to now, ICESat is the only spaceborne laser satellite (Wu and Xing 2010), although airborne lidar systems are increasingly available with the capability of collecting medium- or small-footprint waveforms. One of the objectives for ICESat is to measure surface elevation and vegetation cover globally. It was launched by NASA in January 2003, came into operation about one month later, and operated at an orbit altitude of 600 km until its termination in October 2009. There were 18 operational periods in total during this 7-year mission, collecting about two billion waveforms (Nelson 2010). During most operational years, the ICESat was programmed with a 91-day repeat orbit (Sun et al. 2008). The GLAS on board ICESat uses the near-infrared (~1064 nm) channel to acquire vertical profiles of returned energy as a waveform with a frequency of 40 Hz (Pirotti 2010). For the beginning missions in 2003, the received waveforms over land have 544 bins that correspond to an 81.6 m vertical extent with a 1 ns temporal interval. Starting from 2004, the vertical extent of the waveform was increased to 150 m in order to obtain returns from taller objects (>81.6 m). The returns lower than 58.8 m (lower 392 bins) have a 1 ns interval and the returns higher than 58.8 m (upper 152 bins) have a 4 ns interval (Harding and Carabajal 2005). This results in a total of 1000 ns to record reflected intensity of ICESat waveforms. The footprints are elliptical in shape and their sizes vary a great deal between different missions, with a mean major axis ranging from about 50 to 150 m (Attributes for ICESat Laser Operations Periods, 2009). The average spacing between adjacent footprints along the track is about 175 m.

The ICESat programme distributes 15 Level-1 and Level-2 data products (labelled GLA01 to GLA15), which are readily available on the National Snow & Ice Data Center website. For this study, we employed the Level-1A global altimetry data GLA01, which includes the raw waveform, Level-1B global waveform-based range correction data GLA05, which contains parameters for the shape of footprints, and the Level-2 global land surface altimetry product GLA14, which contains the centroid coordinates of the waveform footprints. All GLA01, GLA05, and GLA14 contain the unique record index number and shot number, which together can be used to link a waveform to its corresponding footprint. For the study area, the waveform data set was obtained for the entirety of 2005 (including 17 February 2005, 23 May 2005, and 24 October 2005) and 2006 (25 February 2006, 27 May 2006, and 27 October 2006), including 2183 waveforms.

2.3. Ground reference data

The National Land Cover Database 2006 (NLCD2006) was used to obtain ground reference data for each waveform. NLCD2006 is a 16-class land-cover classification data set that is consistently available across the conterminous United States with a spatial resolution of 30 m (Fry et al. 2011). In this study, the pixels of interest were those predominantly occupied by buildings, trees, or open spaces. Google Earth images were further employed to refine the ground reference data for each selected ICESat footprint. Google Earth has been used as a reliable source of reference data for remote-sensing-based land-cover classification in recent years (Helmer, Lefsky, and Roberts 2009) as it provides imagery with HSR. The availability of historical imagery in Google Earth since 2001 allowed comparison with the 2005 and 2006 ICESat data. In order to minimize the time difference between the ICESat waveform data and the land-cover ground reference data, Google Earth reference images closest in time to the waveform data were used.
The approximate areas of the waveform footprints were computed using the ellipses defined by their short and long axes, together with the orientations of the long axis, derived from the parameters provided in the GLA05 product (Gong et al. 2011). The waveform footprint polygons thus obtained were then converted to keyhole markup language and imported to Google Earth. When these ICESat footprints are overlaid on top of Google Earth images, the land-cover conditions within the selected footprints can then be determined through visual interpretation.

Since the waveform shapes were primarily determined by the vertical structures of the dominant land-cover type within the footprints, only waveforms having a dominant land-cover type were identified for each class to investigate whether the vertical structure information provided by lidar waveform can be better extracted by using a curve-matching-based method when compared to a rule-based method using Gaussian decomposed parameters. This was relatively easy to achieve for the open space category due to common presence of large open spaces in Dallas. However, due to the likely presence of bare land, grass with a mix of buildings, and trees inside the large ICESat footprints, only waveforms where buildings or trees covered more than 50% of the area of their footprints were considered as buildings or trees, respectively. The waveform footprints that did not have a dominant land-cover type but had a mixture of several dominant types were excluded from the training and testing processes because they cannot be easily labelled with a single class in a hard classification such as the ones used in this study. Like a mixed pixel, the classification of mixed footprints should be investigated using a fuzzy classification algorithm in the future, which is not the focus of this study. If the potential of waveform curves using a curve-matching-based approach can be demonstrated, we will incorporate waveform curves with HSR imagery at the object level for fine-scale land-cover classification. Since the footprints of these objects are derived from HSR imagery through image segmentation, the mixed footprint problem will not be an issue any more because each image segment usually has a dominant land-cover type as a result of segmentation.

3. Methodology

As discussed earlier, the land cover of the study area was classified into three categories: buildings, trees, and open spaces. The methodological steps required for this classification are summarized in the flow chart in Figure 3. The waveform samples are initially preprocessed by transforming coordinates, excluding unreliable waveforms, and removing system noise. The whole data set is then equally split into training and testing parts by random sampling within each category. A set of reference waveforms is selected from the training data set by using a method based on a PCA. For the testing part, each waveform is assigned to the class of a reference waveform with the closest similarity. KS distance is used as a similarity measurement to match the unknown waveform with a reference waveform of known land-cover type. To assess the performance of this KS distance-based land-cover classification approach, a widely used rule-based approach based on Gaussian decomposed parameters is also implemented for comparison purposes. Finally, the accuracies of the classification results from the KS distance and rule-based approaches are evaluated. These steps are described in more detail later.

3.1. Preprocessing

Preprocessing was performed to make waveform data suitable for further analysis. First, both GLA01 and GLA14 were converted from distributed binary format (*.dat)
to ASCII format and the unit of GLA01 was converted from count (0–255) to voltage. Then, the ellipsoid of GLA14 was transformed from the default TOPEX/Poseidon to the WGS84 coordinate system, since other geographical data used the WGS system. Second, unreliable waveforms, such as those with a reflectivity (the i_reflectUncorr field in GLA14) lower than 0.1 (Molijn, Lindenbergh, and Gunter 2011), a value of i_satCorrFlg in GLA14 larger than 0 (Duong et al. 2009), or a signal-to-noise ratio lower than 10 (Hilbert and Schmullius 2012; Hayashi et al. 2013), were excluded because they were primarily caused by heavy cloud. As a result, 844 waveforms were retained with 391, 270, and 183 waveforms for open space, tree, and building categories, respectively. Lastly, a noise removal procedure was applied to eliminate system noise at the beginning and end of each waveform signal. Any waveform signal
value below the threshold was regarded as noise and assigned a value of 0. The threshold was determined using $i\_4nsBgMean + 4.5 \times i\_4nsBgSDEV$ (Lefsky et al. 2007), where $i\_4nsBgMean$ is the mean value of the background noise for the 4 ns filter and $i\_4nsBgSDEV$ is the standard deviation of the background noise for the 4 ns filter. Both were the variables in GLA01. The starting location of a waveform was defined by finding the furthest left location of the waveform where the signal was larger than the threshold value, and the ending location was identified by the right-most location where the signal was larger than the threshold value (Figure 4). Among all the reliable waveforms, the earliest starting location was the 651th ns (corresponding to the top of the highest object), and the last ending location was close to the 1000th ns value (corresponding to the lowest ground level). Therefore, only signals over the last 350 ns (from 651th to 1000th ns) were analysed in this study.

3.2. Kolmogorov–Smirnov distance based approach

3.2.1. Reference waveform selection

After preprocessing, the data set was randomly split into training and testing parts (Table 1). The number of training waveforms for buildings, trees, and open space was 91, 135, and 195, respectively. In traditional land-cover classification based on statistical methods such as maximum likelihood, all training data are used to estimate the statistical parameters of the classification model, and subsequent classification is performed based on the trained model without further reference to the original training

<table>
<thead>
<tr>
<th></th>
<th>No. of training waveforms</th>
<th>No. of testing waveforms</th>
<th>No. of selected waveforms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building</td>
<td>91</td>
<td>92</td>
<td>183</td>
</tr>
<tr>
<td>Tree</td>
<td>135</td>
<td>135</td>
<td>270</td>
</tr>
<tr>
<td>Open space</td>
<td>195</td>
<td>196</td>
<td>391</td>
</tr>
<tr>
<td>Total</td>
<td>421</td>
<td>423</td>
<td>844</td>
</tr>
</tbody>
</table>
samples. Since the KS distance is a curve matching approach that is based on measuring the similarity between an unknown sample and individual references, a method is required to select a subset of references from the training samples. On the one hand, many waveforms for a given land-cover category have a similar profile in the training data set. For example, most of the waveforms for open space have similar single-echo shapes, with minor differences in the starting and ending location, the width, or the peak amplitude of the echoes. Consequently, to avoid data redundancy, it is not appropriate to use all training samples in a land-cover category. On the other hand, there also exist substantial differences between the waveform curves of the same category, a phenomenon often referred to within-class variability. Due to their different height levels and vertical structures, the waveform curves of the same land-cover category in the training data set can be very different from each other, especially for buildings and trees. Therefore, it is also important that the reference waveforms be able to adequately capture the within-class variability for each category. To resolve the dilemma of avoiding duplication while also reflecting variation, a multiple reference waveform selection approach was developed, beginning with a PCA. PCA has the capability to decorrelate the input data and remove duplication by maximizing the variability on the principal components and selecting subsequent orthogonal components in descending order of variance (Sarrazin et al. 2012). Thus, the first principal component should be able to reflect the main trend of waveform distributions and is used to select the first reference waveform in each land-cover category. Intuitively, we should use the subsequent components to select other references. However, the curves of the subsequent principal components are drastically different in shape from all the training waveforms in each category and so an alternative approach was used to capture within-class variability.

For each land-cover category, each training waveform was assumed to be an input variable to the PCA, and each time bin (1 ns) of a waveform was assumed to be an observation. For example, the open space category had 195 waveforms and each waveform had 350 bins of interest; thus, there were 195 input variables each with 350 observations in the open space category. Although the original waveform curves could be used, this study used their cumulative distribution functions (CDFs) as input variables to the PCA. This provides consistency with the KS approach which is based on CDFs. Additionally, correlation coefficients between the CDFs are significantly higher compared to those for the original waveforms due to the monotonic growth of CDFs. For example, the correlation coefficient between the original waveforms for building in Figure 5(e) and building in Figure 5(g) is about 0.27, although they present similar patterns. The coefficient is increased to 0.92 when the CDF curves are used.

The procedure for the PCA-based reference waveform selection was implemented as follows.

Step 1: For each category, apply PCA to all the training waveforms and then find the training waveform with the best similarity to the first PCA component using the smallest KS distance (which is discussed later). Store it as the first reference waveform in the reference set since it is closest to the first component that explains the largest majority of variance in the data.

Step 2: Compute the KS distance between all other training waveforms and the first reference waveform. The training waveform with the smallest (least) similarity (that is, the largest KS distance) to the first reference is selected as the second
reference waveform. This reference waveform is most different from the first reference waveform within the same land-cover category.

Step 3: The next reference waveform to be selected is based on the smallest total similarity (i.e. the largest total KS distance) between a training waveform and all currently selected reference waveforms. The reference set $R$ now has two reference waveforms, i.e. $R = \{R_1, R_2\}$, and assume that the training set $Q$ has $n$ remaining samples, i.e. $Q = \{Q_1, Q_2, \ldots, Q_i, \ldots, Q_n\}$. Calculate the KS distance between $Q_i$ and $R_1$ and $R_2$ as $KS_{i1}$ and $KS_{i2}$, respectively. Sum $K_{i1}$, $KS_{i1}$, and $KS_{i2}$ to obtain the total KS distance $TKS_i$. Sort all $TKS_i$ and the one with the largest total KS distance (i.e. the smallest similarity) is selected as the next reference waveform.

Step 4: Repeat step 3 until the smallest total similarity between the training waveform and existing reference waveforms is smaller than a given threshold. Multiple threshold values were tested, and threshold values of 0.20, 0.25, and 0.20 for tree, building, and open space categories, respectively, were empirically found to produce the most accurate classification results for the training data set. Theoretically, the extracted reference waveforms, although being a small subset of the original training samples, should represent most of the typical samples for each category in the training data set. These selected reference waveforms are then used to classify waveforms in the testing data set.

3.2.2. Kolmogorov–Smirnov distance

KS distance is often used to measure the similarity between two empirical frequency distributions in the non-parametric KS statistical test (Burt and Barber 1996). The
motivation to use KS distance to classify waveform data is inspired by our recently
developed fuzzy KS classifier for land-cover classification using HSR data (Sridharan
and Qiu 2013). In this classifier, the empirical cumulative distribution of spectral values
within each object is used to compare the object-to-object similarity. According to
Zhang et al. (2011) and Pirotti (2010), the ICESat waveforms of GLA01 were originally
stored as the counts (0–255) of returning impulse along the time axis. Therefore, a
waveform can be treated as a time-varying frequency distribution function (histogram)
(Zaletnyik, Laky, and Toth 2010). Muss, Mladenoff, and Townsend (2011) employed a
height frequency distribution of discrete-return lidar points within a forest stand or a plot
to simulate the corresponding pseudo-waveform using a cubic spline function. It was
demonstrated that the resultant pseudo-waveform is similar in its characteristics to the
traditional waveform, since the height essentially corresponds to time multiplied by the
speed of light. For these reasons, the KS distance can be calculated based on either the
true or pseudo-waveform to measure the similarity of their corresponding empirical
frequency distributions.

Figure 5 shows that the CDFs of waveforms corresponding to different land-
cover types are noticeably distinct from each other. Waveforms for open space often
have a single, narrow echo resulting in a steeply increasing edge in the correspond-
ing CDFs, which quickly reaches one. A typical building waveform usually has two
narrow echoes, one for the building roof and the other for the ground, with a gap
between them corresponding to the building height. This results in the steep steps
observed in the CDF. As for trees, the waveforms usually have two major echoes
with varied width: a wider top echo corresponding to the tree crown and a narrower
second echo corresponding to the ground. Accordingly, the CDF gradually increases
initially with gentle slopes but the end reaches one abruptly when the signal hits the
ground.

Figure 5 also suggests that, compared with the original waveforms, the CDFs may
be able to reduce within-class variability. For example, due to the differences in the
local ups and downs and in the global patterns such as the echo widths and peak
amplitudes, the original waveforms for tree in Figure 5(i) and tree in Figure 5(k) have
quite different shapes, but their corresponding CDFs are more similar to each other in
shape because the local and global differences are minimized during the aggregation and
normalization process inherent in creating the CDFs. Therefore, the use of CDFs may
help reduce the omission error of the classification. On the other hand, the CDF curves
may also increase the between-class similarity, because the similarity between the CDFs
of different categories may become larger than that between their original waveforms for
the same reason, which may potentially increase the omission error. For example, the
CDFs of open space in Figure 5(b) and building in Figure 5(f) are now a little more
similar to each other compared to their original waveforms, but still with considerable
distinction, especially in the slope of the steps. The KS distance is not based on a point-
by-point comparison of the two CDFs at every bin. It is the maximum difference
between two CDF curves, which is actually the cumulative difference between the
two. Therefore, KS distance can capture the major difference in the patterns of the
CDFs, such as the slope of the steps.

Based on the two-sample KS distance, the following steps were employed to evaluate
whether the CDF of an unknown waveform is similar to that of a reference waveform.

Step 1: For all input waveforms, the CDF is first computed using Equation (1). This
also serves as a normalization process to make waveforms acquired in different epochs or
with different atmosphere conditions comparable with each other.
\[ S(i) = \frac{\sum_{j=\min V_j}^{i=\min V_j} V_j}{\sum_{j=\max V_j}^{i=\max V_j} V_j}, \]  

(1)

where \( V_j \) is the received energy at the time \( j \)th ns. In this study, \( \min \) equals 651 and \( \max \) equals 1000, because only the last 350 ns (from 651 ns to 1000 ns) are of interest.

Step 2: To assess similarity, KS distance is measured as the maximum absolute distance between the CDF of an unknown waveform and that of each reference waveform using Equation (2),

\[ D_k = \max_{\min i}^{\max i} \left| S_{i}^{k}(i) - S_{i}(i) \right|, \quad k = 1, 2, \cdots, K, \]  

(2)

where \( S_{i}^{k} \) is the CDF of the \( k \)th reference waveform and \( S_{i} \) is the CDF of an unknown waveform.

Step 3: For classification, the unknown waveform is assigned to the reference class with the smallest KS distance (Equation (3)).

\[ \hat{k} = \arg \min_{1 \leq k \leq K} D_k. \]  

(3)

As an example, it is apparent in Figure 6 that the CDF of an unknown waveform closely resembles that of the open space reference waveform, rather than the CDFs of the building and tree reference waveforms. The maximum absolute distance, measured vertically, between the CDF of this unknown waveform and that of the open space reference waveform is therefore the minimum among all the \( D_k \). As a result, the unknown waveform is classified into the open space category.

3.3. Rule-based approach

Rule-based methods based on parameters extracted from preprocessed and Gaussian decomposed waveforms have been widely adopted for land-cover classification (Duong, Pfeifer, and Lindenbergh 2006; Alexander et al. 2010). In this study, a

![Figure 6. CDFs of three reference waveforms and an unknown waveform. The vertical black lines are the maximum absolute distances between the CDF of the unknown waveform and that of the individual reference waveform.](image-url)
rule-based classification approach using ICESat waveforms was also implemented for comparison with the KS-based approach. The first step of the rule-based approach involved preprocessing as described in Section 3.1 to extract the values of the starting and ending locations. The next step, Gaussian decomposition, assumes that waveform $w(t)$ is a mixture of $k$ Gaussian distribution components (echoes) $W_j(t)$. Thus, a waveform can be expressed as Equation (4),

$$w(t) = \sum_{j=1}^{k} W_j(t), \text{ with } W_j(t) = A_j e^{-\frac{(t-\mu_j)^2}{2\sigma_j^2}}.$$  

where $w(t)$ is the received energy of the waveform at time $t$, $k$ is the number of Gaussian components, $W_j(t)$ is the contribution of the $j$th component at time $t$, $A_j$ is the amplitude of component $j$, $\mu_j$ is the mean of component $j$, and $\sigma_j$ is the standard deviation of component $j$. An expectation-maximization (EM) algorithm (Oliver, Baxter, and Wallance 1996) is applied to decompose the waveform into a series of Gaussian components by estimating the values for $k$, $A_j$, $\mu_j$, and $\sigma_j$. Figure 7 presents an original waveform with multiple peaks, which is decomposed into three Gaussian components by using the EM algorithm.

Preprocessing and Gaussian decomposition result in the following parameters for each waveform: $s.loc$ (starting location) and $e.loc$ (ending location), width (the distance between starting location and ending location of the noise-removed waveform), $k$ (the number of Gaussian components), and $A_m$ (amplitude), $\mu$ (mean), and $\sigma$ (standard deviation) for each Gaussian component. These parameters are used to classify the waveforms into different land-cover types. $k$ is usually the most important parameter. It can be used, for example, to effectively differentiate open space waveforms from building and tree waveforms. A majority of open space waveforms are single-echo with a narrow width due to the flat surface; only a small number have two or three echoes, caused by the existence of rough terrain or other objects within the footprints. By comparison, all buildings and trees have multiple echoes, thus $k$ is greater than 1. The $s.loc$ is another useful parameter that can be used to separate high trees from buildings. In our study area, a large number of trees are higher than buildings, which results in an earlier $s.loc$ for high trees than for buildings. The width of a waveform, measured from $s.loc$ to $e.loc$, can also be a useful indicator. However, it basically
follows the same trend as s.loc. An earlier s.loc usually leads to a larger width because all e.loc values, which correspond to ground, are generally similar. As a result, in this study, width is not used for classification to avoid redundancy.

The other three parameters, Am, μ, and σ, which characterize the decomposed Gaussian components, have not been widely used for land-cover classification in the literature. However, the sum of the standard deviation (total.sd) of the Gaussian components may be used to differentiate between trees and buildings. Echoes over trees are usually wider (thus total.sd is larger) because of the cone shape of the tree crown, whereas the echoes over flat-roofed buildings are narrower (total.sd is smaller) due to the flat surface. Slope-roofed buildings may have one or two roof echoes that are wider than those of flat roofs, but usually their total.sd is still smaller than that for most trees.

Consequently, only k, s.loc, and total.sd are used to derive the ‘if–then’ rules for land-cover classification. The threshold values of these parameters are determined by analysing the frequency distributions of the individual parameters from the training waveforms. For example, if the frequency distribution of k shows that the majority of waveforms over open space have a value of 1 for k, then the threshold should be set to 1 in order to differentiate open space from the other two categories.

3.4. Accuracy assessment

The classification results were assessed using producer’s accuracy, user’s accuracy, overall accuracy, and kappa statistics. The producer’s accuracy is the probability of unknown waveforms being correctly classified as the corresponding ground reference data; user’s accuracy is the probability of a classified waveform being in that category. The overall accuracy is calculated by the total number of correctly classified waveforms divided by the total count of waveforms. The kappa coefficient (ranging from 0 to 1) is used to evaluate the agreement between ground reference data and classification results (Congalton 1991). Values of kappa larger than 0.5 indicate a good agreement, and values greater than 0.7 indicate a strong agreement (Qiu 2008; Khanna et al. 2011).

4. Results and discussion

This section compares the classification performance of the KS-based approach using the PCA selected reference waveforms with both the rule-based approach and the KS-based approach using randomly selected reference waveforms.

4.1. Results for KS-based classification using PCA selected reference waveforms

Based on the PCA analysis, a total of 54 reference waveforms were selected from the 421 training samples, comprising 19 reference waveforms for buildings, 24 for trees, and 11 for open space. These selected reference waveforms were used to classify the testing waveforms, using KS distance to assess similarity. Table 2 is a contingency table of the classification results: the columns show ground reference, and the rows show the classified land cover. Producer’s accuracy, user’s accuracy, overall accuracy, and kappa statistics are computed to assess the classification results. For visualization purposes, the classification map is shown in Figure 8.

In general, most waveforms have been correctly classified with an overall accuracy of 87.2%, a kappa coefficient of 0.80, and a producer’s accuracy of 74%, 86%, and 94% for
Table 2. Confusion matrix of the KS-based approach.

<table>
<thead>
<tr>
<th>Classified data</th>
<th>Ground reference data</th>
<th>Producer’s accuracy (%)</th>
<th>User’s accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building</td>
<td>68</td>
<td>74</td>
<td>69</td>
</tr>
<tr>
<td>Tree</td>
<td>22</td>
<td>86</td>
<td>84</td>
</tr>
<tr>
<td>Open Space</td>
<td>2</td>
<td>94</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>Building</td>
<td>Tree</td>
<td>Open space</td>
</tr>
</tbody>
</table>

Notes: Overall accuracy: 87.2%. Kappa coefficient: 0.80.

Figure 8. The classification result of ICESat waveforms. (a) Map of classification for all footprints in the study area. (b) The zoom-in map of the area within the yellow rectangle in (a) with the latitude, longitude of its upper right corner being 32.915° N, -96.885° E.
building, tree, and open space categories, respectively, indicating a strong agreement between true ground references and classification results.

For buildings, 68 out of 92 waveforms were correctly classified. In the study area, most of the buildings were flat-roofed. Waveforms over the flat-roofed buildings usually show two modes, one for the roof and the other for the ground. Both modes have a narrow echo width caused by their flat surfaces because, when the laser hits a flat roof or ground, a majority of the incoming pulses are reflected over a short time period. This leads to a narrow echo width with high amplitude of reflection in the waveform, which in turn results in steep slopes in their associated CDF. This made it possible to separate waveforms of flat-roofed buildings from those of trees. The buildings misclassified as trees were mainly slope-roofed, because their waveform curves were similar to those of trees (Alexander et al. 2010). For trees, 19 out of 135 samples were misclassified as buildings, most of which were low trees surrounding residential buildings, which were confused with the slope-roofed buildings. The open space had the highest accuracy (94%) among the three categories, although based on the smallest number of reference waveforms. Waveforms for open space have the least variation in their curves, usually with a single mode and a narrow echo width. The small differences in their waveforms lie primarily in the starting location and the amplitude of the wave caused by variation in open space elevation and surface conditions, respectively. As a result, they could be easily differentiated from buildings and trees, both of which have a wider echo width, an earlier starting location and are multimodal. Only 11 open space waveforms were misclassified as buildings, most of which were multiecho due to the existence of rough terrain or other objects within their footprints making their profiles similar to those of low buildings.

4.2. Results for rule-based classification

For the rule-based approach, three parameters $k$ (the number of Gaussian components), total.sd (the sum of the standard deviation of individual Gaussian components), and s.loc (the starting location of the noise-removed waveform) were used to derive the ‘if–then’ classification rules, as discussed earlier.

Figure 9 presents the frequency distributions of $k$, total.sd, and s.loc of the samples in the training data set. For the open space category, all three parameters are quite different from those for the building and tree categories. This is especially true for $k$, which is larger than 1 for all buildings and trees, whereas the vast majority of waveforms for open space are single-echo with $k$ equal to 1. Therefore, the first classification rule was that, if $k$ equals 1, the corresponding waveform is open space; otherwise, the waveform remains unclassified. The parameter total.sd can be used to further separate the open space samples with more than one echo ($k > 1$) from the unclassified samples. A few waveforms for open space were affected by slope or rough terrain, resulting in multiple, narrow-width echoes. However, the total.sd for most of these is smaller than 15 ns in the study area due to their small number of echoes and narrow echo width, which was clearly different from buildings and trees. After applying a second classification rule that, if total.sd < 15, the corresponding waveform is also open space, the remaining unclassified samples were mostly trees and buildings.

Compared with open space, waveforms for buildings and trees are more complex, evidenced by the wide range and overlap of their three parameters. Consequently, a combination of parameters, rather than a single parameter, was required to differentiate between them. Since the trees in the study area were usually higher than buildings, and the
cone shapes of tree crowns causes the tree waveforms to have wider echoes, most trees in
the study area have an earlier s.loc (<870 ns) and a larger total.sd (>30) than buildings.
These two parameters were thus used to distinguish trees from buildings. The complete set
of rules derived for land-cover classification is shown as a decision tree in Figure 10.

Table 3 presents the classification results using these rules. The rule-based approach
performed reasonably well with an overall accuracy of 83.7% and a kappa coefficient of
0.75. It produced acceptable producer’s accuracies of 72% for buildings and 79% for
trees, with 16 out of 92 buildings misclassified as trees and 28 out of 135 trees
misclassified as buildings. This was primarily due to the overlaps in total.sd and s.loc
between buildings and trees, since these two categories have some waveforms with similar
heights or vertical structures. This was particularly true for waveforms over slope-roofed
residential buildings surrounded by similar height or lower height trees, which results in
similar s.loc and total.sd values with trees of a similar height. These buildings have a high
probability of being misclassified as trees. As a consequence, the combined use of s.loc
and total.sd to discriminate buildings from trees met only with limited success. As for
open space, their waveforms usually have a single-echo due to the flat surface, with a few

![Figure 9](image_url)

**Figure 9.** Frequency distributions of $k$ (i.e. the number of Gaussian components), total.sd (i.e. the
sum of standard deviation of the individual Gaussian components), and s.loc (starting location) of
training samples.
exceptions that have slope or rough terrain within the footprint. Consequently, the combination of $k$ and total.sd was capable of differentiating open space, achieving a producer’s accuracy of 92% and a user’s accuracy of 95%.

4.3. Comparison of KS-based classification and rule-based classification

In the rule-based approach, extracted parameters are used to differentiate the waveforms between land-cover types. In general, each parameter corresponds to the physical characteristics of geographical objects. For example, s.loc corresponds to the height of an object, $k$ assesses the number of reflected objects, and total.sd relates to the object’s vertical structure (flat or sloped). The simple thresholds set for these parameters were able to achieve acceptable discrimination results. However, when two different objects have similar heights (such as a tree and a slope-roofed building) and similar structure profiles (such as a sloped top), these parameters may not be able to differentiate them. In comparison, the KS-based approach classified waveforms based on matching their entire CDF. Since the CDF is derived from the original waveform, it contains not only the characteristics captured by the extracted parameters but also the information that could not

Table 3. Confusion matrix of the rule-based approach.

<table>
<thead>
<tr>
<th>Classified data</th>
<th>Ground reference data</th>
<th>Producer’s accuracy (%)</th>
<th>User’s accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building</td>
<td>66 28 15</td>
<td>72 61</td>
<td></td>
</tr>
<tr>
<td>Tree</td>
<td>16 107 0</td>
<td>79 87</td>
<td></td>
</tr>
<tr>
<td>Open space</td>
<td>10 0 181</td>
<td>92 95</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Overall accuracy: 83.7%.
Kappa coefficient: 0.75.
be represented by these parameters, such as the small local peaks, and the local skewness and kurtosis that cannot be completely characterized by a few Gaussian components. Basically, the entirety of the information provided by the waveform is made available for classification via the CDF. As a result, the KS-based approach was able to achieve better performance with a 3.5 percentage point improvement in overall accuracy, and 2, 7, and 2 percentage point improvements in producer’s accuracy for building, tree, and open space, respectively, over the rule-based classification.

These numbers, although not large, do validate the superiority of the KS distance approach. However, an even more significant advantage of the KS approach is the absence of any need for interpretation by analysts, which helps ensure consistency in classification results. The rule-based approach requires the analysts to determine rules based on an inspection of parameter values, and there is no guarantee that two analysts would derive identical rules. The use of PCA, coupled with KS distance measures, for selection of the reference waveforms relies on a purely quantitative approach which will always be duplicable given the same input data. Furthermore, the derivation of rules can be very specific to the characteristics of the study area. This is evidenced by the discussion of rule selection earlier, which references particular characteristics of our study area, such as the large height of trees relative to buildings reflecting the maturity of these residential neighbourhoods. If the study area included newer neighbourhoods with recently planted trees, these rules would likely require modification by the analysts. Using the PCA approach, these differences are automatically incorporated.

4.4. Results for KS-based classification using randomly selected reference waveforms

We also compared KS-based classification results using the PCA selected reference waveforms with classification results using 10 trials of randomly selected reference waveforms. The number of references for each category was kept the same as that for PCA selected references across all 10 trials, that is, 19 reference waveforms for buildings, 24 for trees, and 11 for open space. The classification results (Table 4) indicate that the PCA selected reference waveforms consistently outperformed randomly selected

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>Building (%)</th>
<th>Tree (%)</th>
<th>Open space (%)</th>
<th>Overall (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomly selected reference waveforms</td>
<td>1th</td>
<td>54</td>
<td>79</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>66</td>
<td>82</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>61</td>
<td>81</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>4th</td>
<td>63</td>
<td>75</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>5th</td>
<td>66</td>
<td>84</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>6th</td>
<td>72</td>
<td>82</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>7th</td>
<td>64</td>
<td>80</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>8th</td>
<td>73</td>
<td>81</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>9th</td>
<td>68</td>
<td>81</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>10th</td>
<td>67</td>
<td>79</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>std</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>66</td>
<td>80</td>
<td>90</td>
</tr>
<tr>
<td>PCA selected reference waveforms</td>
<td>74</td>
<td>86</td>
<td>94</td>
<td>87.2</td>
</tr>
</tbody>
</table>
references. The overall accuracy of 87.2% for the PCA approach was higher than the overall accuracy of any of the random selection trials. Nevertheless, the randomly selected references were able to produce acceptable results with an average overall accuracy of 81.5%. However, accuracy varied substantially between the land-cover classes. The performance for buildings was consistently the poorest among all categories in the 10 trials, which resulted in the lowest average producer’s accuracy of 66% compared with 74% for the PCA approach. Among all the building training samples, only a small number were slope-roofed residential buildings surrounded by trees of similar height. Random reference selection missed picking up a reference for these buildings, causing many to be misclassified as trees. For trees, the randomly selected references achieved consistent and relatively accurate results for all 10 trials, with an average producer’s accuracy of 80%, compared with 86% for the PCA approach, and a small standard deviation of 2%. As for open space, an average accuracy of 90% was achieved (with a minimum of 81% and a maximum of 94%) compared with 94% for the PCA approach. There are only a few open space training samples having multiechoes due to the existence of rough terrain or other objects within the footprint. Missing references for these samples by the random selection method is the likely cause of poorer classification results in some of the trials.

In general, the high variance of the producer’s accuracies (with standard deviations of 5% and 4% for building and open space, respectively) suggests that the quality of the selected reference waveforms has an important impact on the classification results. It is critical that the selected references for each category be able to adequately capture the within-class variation in the training data set. This is the main advantage of the PCA-based reference selection approach. The randomly selected references, on the contrary, may include many similar waveforms, but miss the ones with unique patterns due to a limited number of examples in the training data set. Compared to the randomly selected references, the PCA selected references are capable of capturing the within-class variability, which is critically important for methods based on curve matching.

5. Conclusions

In this study, we investigated the potential of using the ICESat waveform as a one-dimensional signal for land-cover classification over level terrain. For this purpose, the KS distance is employed to measure the similarity between test waveforms and individual reference waveforms. The following conclusions can be drawn.

1. Objects with distinct vertical structures (e.g. building, tree, and open space) within the ICESat footprint exhibit unique waveform curves that can be used to perform land-cover classification.

2. The use of the full-waveform curve to discriminate buildings, trees, and open space over level terrain provides an alternative approach to existing metrics-based methods using a limited number of parameters (e.g. amplitude, width, number of echoes, etc.) derived from the waveforms. By taking the full-waveform curve into consideration, the KS-based approach achieves a superior performance over the widely adopted rule-based method, with the additional advantages of simplicity of mathematical form, ease of implementation, no requirement for computationally complex Gaussian decomposition, and consistency of results stemming from the lack of need for analyst input.

3. The references selected based on PCA play an important role in achieving better classification results compared with randomly selected references. Generally, the
better the reference samples are able to capture within-class variability, the higher
the classification accuracies achieved by a curve matching approach.

(4) The utilization of full-waveform curve information is sufficient to separate objects
with distinct vertical structures. However, its ability to differentiate objects with
similar vertical structures is limited. Future studies will focus on fusing full-
waveform data and spectral data to perform fine-level land-cover classification
by taking advantage of both vertical structure information and surface spectral
reflectance information.

(5) ICESat waveforms are sensitive to underlying terrain variation within the large
footprints (Hilbert and Schmullius 2012), which makes it challenging to apply
both KS-based and rule-based classification using ICESat waveform to sloped
areas (slope > 15°). Future studies may focus on removing the terrain effects from
the ICESat waveforms in advance and then evaluate the performance of curve-
based approaches to waveform-based classification over slopes.

The use of ICESat waveforms for land-cover mapping is not likely to be widespread due
to the large spacing between its footprints along and between the tracts, and the presence
of mixed land covers within its large footprints. It was used here primarily to test a novel
approach to land-cover classification using full-waveform data and KS-distance-based
curve matching. This approach will support the laser altimeter data expected from ICESat-
2, the 2nd-generation ICESat mission scheduled for launch in early 2016. The approach is
equally appropriate for feature extraction from small-footprint waveforms. More and more
airborne lidar systems now provide full waveforms with small footprints. These are likely
to have pure land-cover types in each footprint and minimal spacing between them. These
will be the focus of our future studies.

Acknowledgements
The authors would like to thank the National Snow and Ice Data Center for providing ICESat data.

References
Changes of Selected Soil Properties in the Mountain Region of Rawalakot Azad Jammu and

Attribute for the Classification of Full-Waveform Airborne Laser Scanning Data in Urban
isprsjprs.2010.05.002.

Imagery for Mapping Riparian Biophysical Parameters and Land Cover Types in Australian
foreco.2009.11.018.

Peatland Topography and Peat Swamp Forest Biomass in Kalimantan, Indonesia.” *Remote

“Regional Aboveground Forest Biomass using Airborne and Spaceborne LiDAR in Québec.”


