ADVANCES IN WAVEFORM AND PHOTON COUNTING LIDAR PROCESSING FOR FOREST VEGETATION APPLICATIONS

A Dissertation

by

TAN ZHOU

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Sorin C. Popescu
Committee Members,	Marian Eriksson
	Anthony Filippi
	Michelle Lawing
Head of Department,	Kathleen Kavanagh

December 2017

Major Subject: Ecosystem Science and Management

ABSTRACT

Full waveform (FW) and photon counting LiDAR (PCL) data have garnered greater attention due to increasing data availability, a wealth of information they contain and promising prospects for large scale vegetation mapping. However, many factors such as complex processing steps and scarce non-proprietary tools preclude extensive and practical uses of these data for vegetation characterization. Therefore, the overall goal of this study is to develop algorithms to process FW and PCL data and to explore their potential in real-world applications.

Study I explored classical waveform decomposition methods such as the Gaussian decomposition, Richardson–Lucy (RL) deconvolution and a newly introduced optimized Gold deconvolution to process FW LiDAR data. Results demonstrated the advantages of the deconvolution and decomposition method, and the three approaches generated satisfactory results, while the best performances varied when different criteria were used.

Built upon Study I, Study II applied the Bayesian non-linear modeling concepts for waveform decomposition and quantified the propagation of error and uncertainty along the processing steps. The performance evaluation and uncertainty analysis at the parameter, derived point cloud and surface model levels showed that the Bayesian decomposition could enhance the credibility of decomposition results in a probabilistic sense to capture the true error of estimates and trace the uncertainty propagation along the processing steps.

In study III, we exploited FW LiDAR data to classify tree species through integrating machine learning methods (the Random forests (RF) and Conditional inference forests (CF)) and Bayesian inference method. Results of classification accuracy highlighted that the Bayesian method was a superior alternative to machine learning methods, and rendered users with more confidence for interpreting and applying classification results to real-world tasks such as forest inventory.

Study IV focused on developing a framework to derive terrain elevation and vegetation canopy height from test-bed sensor data and to pre-validate the capacity of the

upcoming Ice, Cloud and Land Elevation Satellite-2 (ICESat-2) mission. The methodology developed in this study illustrates plausible ways of processing the data that are structurally similar to expected ICESat-2 data and holds the potential to be a benchmark for further method adjustment once genuine ICESat-2 are available.

DEDICATION

This dissertation is dedicated to my parents who raised me, supported me and gave me endless love, and to the mentors who guided me, inspired me and made me become a better man.

ACKNOWLEDGEMENTS

This dissertation would not have been possible in the present thoroughness and quality without the help of many individuals. Foremost, I would like to thank my advisor, Dr. Sorin C. Popescu for his guidance, encouragement and inspiration throughout my doctoral work to make me being a scientific researcher. We spent lots of time together on discussing research and life aspects which made me grow mentally and will continue to be conducive to my further life and career. My sincere gratitude is extended to my committee members, Drs. Marian Eriksson, Anthony Filippi and Michelle Lawing, for their time, efforts and support for the past few years. Their critical reviews and constructive comments helped me improve my dissertation. I am also grateful to Dr. Raghavan Srinivasan who took me to the Texas A&M University and Dr. Thomas W. Boutton for his invaluable guidance and help at the beginning of my Ph.D. I also thank the following professors and researchers for their help in my Ph.D. work: Drs. Keith Krause, Rusty Feagin, Jeffrey D. Hart, Paul-Christian Bürkner and Amy Neuenschwander.

I acknowledge the support by the National Ecological Observatory Network for providing waveform LiDAR data and field data, and NSF Doctoral Dissertation Improvement Grant (DDIG) for partly funding my research.

Many friends have helped me and accompanied me during my Ph.D. period, I would thank all of them: Ryan Sheridan, Tony Ku, Eric Putman, Malambo Lonesome, Lana Narine, Shruthi Srinivasan, Sasa Tapaneeyakul, Tianyi Wang, Yiran Li, Shangguan Liantian, Ruizi Wang, Xiaohan Mei, Minghao Yan, Yinuo Sun, Gang Zhao, Xiangming Sun, Yong Zhou, Yong Wang, Rui Liu and Kunpeng Wang.

Thanks also go to my colleagues and the department faculty and staff for making my time at Texas A&M University a great experience.

Last but not least, thanks to my parents for their unconditional support, patience and endless love for past years.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professors Sorin C. Popescu [advisor], Marian Eriksson, and Michelle Lawing of the Department of Ecosystem Science and Management and Professor Anthony Filippi of the Department of Geography.

The data for Chapter 2, 3 and 4 were provided by the National Ecological Observatory Network (NEON). The data used in Chapter 5 were provided by NASA ICESat-2 Science Definition Team.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by an Excellence Fellowship of Texas A&M University, Dishman Lucas Assistantship teaching assistantship and research assistantship from my advisor and the Department of Ecosystem Science and Management, Texas A&M University and a Doctoral Dissertation Improvement Grant (DDIG) from NSF.

NOMENCLATURE

DR	Discrete-return
FW	Full-waveform
LiDAR	Light Detection and Raning
PCL	Photon counting LiDAR
ICESat-2	Ice, Cloud and Land Elevation Satellite-2
RL	Richardson-Lucy
NEON	National Ecological Observatory Network
RMSE	Root mean square error

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I
INTRODUCTION

1.1 General background

Forest ecosystems are the primary reservoir of carbon and account for 80% of Earth's total plant biomass (Kindermann et al., 2008). The structure of forests is crucial for estimating forest dynamics and determining light use and net primary productivity (Marvin et al., 2014). Acquiring knowledge about vegetation structure and forest biomass is conducive to designing effective plans for sustainable forest management and climate change mitigation (Allouis et al., 2013; Hollaus et al., 2009a).

The development of advanced remote sensing techniques (e.g., Landsat, hyperspectral and radar) over the past decades has contributed significantly to the monitoring of vegetation structure and biomass dynamics. Especially with the emergence of Light Detection and Ranging (LiDAR), a three-dimensional (3D) remote sensing technology, we can now characterize vegetation compositions and structures in both horizontal and vertical dimensions.

Previous studies have demonstrated that discrete-rerun (DR) LiDAR data can accurately estimate many key forest structure variables and predict tree species (Holmgren et al., 2008; Kim et al., 2009), forest biomass (Gleason and Im, 2012; Popescu et al., 2003; Zhao et al., 2009) and carbon stocks (García et al., 2010). For instance, various studies have elaborated how tree height, canopy height, crown width, average basal area, stem volume and leaf area can be retrieved or modeled from DR LiDAR data (Chen et al., 2007; Kraus and Pfeifer, 1998; Lefsky et al., 2002; Lefsky et al., 2005; Zhao and Popescu, 2009). The following section gives a brief review of waveform LiDAR processing, Bayesian inference, tree species identification using FW LiDAR data and photon counting LiDAR (PCL) data processing.

1.1.1 Full-waveform LiDAR processing

The advent of FW LiDAR data demonstrates their remarkable abilities of accurately characterizing forest structures. Unlike DR LiDAR system that only record

several signals along the pulse line, the FW LiDAR system can record the entire echo scattered from illuminated objects with different temporal resolutions (such as 1/2/4 nanosecond(s), ns) through digital sampling. Thus, the resulting datasets of FW are composed of pulses and waves instead of just 3D point clouds. This additional information gives users more control over the data interpretation and application (Chauve et al., 2007). In general, large peaks (high reflected intensity) are interpreted to represent illuminated objects or surfaces along the pulse line.

A complete procedure of a laser pulse propagating along the pulse line consists of three stages: (1) the laser is emitted from a laser sensor at the bottom of aircraft (Outgoing pulse); (2) it interacts with the trees, subshrub and ground; and (3) the receiver records the entire reflected energy along the pulse line.

FW LiDAR data are first introduced as acquired by sensors known as large footprint profilers such as SLICER (Scanning LiDAR Imager of Canopies by Echo Recovery, 10 m footprint), LVIS (Laser Vegetation Imaging Sensor, 25 m footprint) and GLAS (the Geoscience Laser Altimeter System, 70 m footprint). All of them have been successfully applied to estimate various forest parameters such as canopy heights and vegetation studies worldwide (Blair et al., 1999; Drake et al., 2002; Harding and Carabajal, 2005). Recent advances of commercial LiDAR systems have promoted the availability of small footprint waveform from remote sensing industry providers. However, FW LiDAR cannot be directly used without post-processing of these waves through the aid of proprietary software. This limitation hinders the extensive applications of FW LiDAR vegetation characterization and mapping. Therefore, the development of dedicated methods or open source tools for FW LiDAR processing are urgently needed.

Existing methods for FW LiDAR processing can be categorized into two types: the direct decomposition method and the deconvolution and decomposition method. Each method has been successfully applied for echo detection based on its own physical background. For direct decomposition, the Gaussian function is commonly used to decompose waveforms (Wagner et al., 2006) and further obtain waveform components for characterizing the objects along the pulse line. Nevertheless, the echoes reflected back

from the roof materials or surfaces are not always symmetric and that makes the Gaussian function no longer sufficient to fit the waveforms. The process of deconvolution can overcome this problem by reducing unwanted system contributions such as the effect coming from outgoing pulse, and recovering the true distribution of the illuminated surfaces. Various studies have elaborated the application of different deconvolution algorithms such as B-spline, Richardson-Lucy (RL), Non-negative Least Squares (NNLS) and Wiener Filter (WF) (Cawse-Nicholson et al., 2014; McGlinchy et al., 2014; Neuenschwander, 2008; Roncat et al., 2010; Rowe, 2013; Wu et al., 2011) to recover the true cross-sectional profile of an illuminated object.

1.1.2 Bayesian statistics and uncertainty

The methods mentioned above are based on the deterministic models that can only generate adequate fitting models, however, the uncertainty of estimations cannot be characterized in a probabilistic sense. In addition, the function for waveform decomposition is non-linear and often suffers from non-uniqueness problems (several models may fit the observations or different combinations of parameters can fit the data with the same model). Mallet et al. (Mallet and Bretar, 2009) have demonstrated that there are four functions including generalized Gaussian, Weibull, Nakagami and Burr functions can be alternatives for the Gaussian function to fit FW LiDAR data. While the focus of my dissertation is on fitting waveforms, future research should investigate how to best incorporate alternative functions.

The Bayesian method enables us to resolve these concerns through obtaining estimates of parameters in terms of probability function in model space. Bayesian inference provides a way of conducting statistical inference by means of Bayes' rule that gives users a rational means of updating prior beliefs in light of the information contained in the data.

Bayes' rules can be expressed as:

$$p(\theta|D) = \frac{p(D,\theta)}{m(D)} = \frac{p(D|\theta)p(\theta)}{\int p(D|\tilde{\theta})p(\tilde{\theta})d\tilde{\theta}} \quad \text{(I-1)}$$

where D is the observed data, $p(\theta)$ is the prior belief about the parameters, $p(D|\theta)$ is the distribution of D conditional on θ being the true value of the unknown parameter vector

or the likelihood function. $\tilde{\theta}$ is the set of all possible values for θ in its parameter space, $p(\theta|D)$ is the posterior distribution.

More generally, Eq. (I-2) can be written as:

$$p(\theta|D) \sim p(D|\theta)p(\theta) \quad \text{(I-2)}$$

Since $\int p(D|\tilde{\theta})p(\tilde{\theta})\, d\tilde{\theta}$ is a constant for a given parameter space, Eq. (I-2) clearly shows us that the posterior distribution of parameters θ is proportional to the prior distribution of parameters times the likelihood function (Hoff, 2009).

To obtain a more comprehensive overview of frequentist statistics and Bayesian statistics, we summarized the main differences in Table I-1.

Table I-1. Comparisons of frequentist statistics and Bayesian statistics

	Frequentist statistics	Bayesian statistics
Data	A repeatable sample	Observed from realized sample (fixed)
Underlying Parameters	Fixed but unknown; no priors	Unknown and described probabilistically; require priors (random variate)
Interpretation of parameters	Coverage of "true" parameter in the long run; parameter is a single value; confidence interval (continuous)	Researchers' belief about given values of parameter are true; parameter is a distribution; credible interval (can be non-continuous)
Uncertainty	How parameter estimates vary from one to the next in repeated sampling from the same population	How much prior opinions about parameters change in light of the observed data

*above comparisons are summarized from https://en.wikipedia.org/wiki/Frequentist_inference and https://www.quora.com/What-is-the-difference-between-Bayesian-and-frequentist-statisticians

The Bayesian method has been successfully applied in numerous domains such as electroencephalogram separation (Roonizi and Sassi, 2016), geophysical inversion (Oh and Kwon, 2001; Sen and Stoffa, 1996) and seismic waves inversion (Gouveia and Scales, 1998; Hong and Sen, 2009). However, few studies have applied the Bayesian method to processing FW LiDAR data. Furthermore, the uncertainty of processing steps also cannot be captured using the deterministic methods such as the direct decomposition, the Gold

and RL deconvolution and decomposition methods. Therefore, we propose to apply the Bayesian concept to decompose FW LiDAR data in this study. Details of this method can be found in Chapter 3.

MCMC introduction

Markov chain Monte Carlo (MCMC) (Gelfand and Smith, 1990) is a crucial technique for the rapid expansion of the Bayesian method. There are cases that some parameters' posterior distributions are difficult or impossible to sample when the non-conjugate prior (the posterior distributions $p(\theta|x)$ are not in the same family as the prior probability distribution $p(\theta)$) is used or the integration of parameters is conducted over a high dimensional parameter space (Hoff, 2009). In such conditions, the MCMC method can be helpful by approximating the true posterior distribution using the joint distribution $p(D|\theta)p(\theta)$ instead of directly sampling from the integration of posterior distribution for parameters of interest $p(\theta|D)$. There are several commonly used MCMC methods such as Gibbs sampling, Metropolis algorithm, and Metropolis-Hastings (MH) algorithm.

Gibbs sampling

Among the common MCMC methods, the Gibbs sampling is the simplest version of the MCMC method, while it requires the full conditional distribution of each parameter that can be sampled exactly. Suppose that $f(\theta_1, \theta_2, \ldots, \theta_p)$ is the joint density of all parameters, the full conditionals for each parameter are

$$f(\theta_1|\theta_2, \ldots, \theta_p), f(\theta_2|\theta_1, \ldots, \theta_p), \ldots, f(\theta_p|\theta_1, \ldots, \theta_{p-1}).$$

Denote the i^{th} sample by $\theta^{(i)} = (\theta_{i1}, \theta_{i2}, \ldots, \theta_{ip})$. Assuming we have arrived at $\theta^{(s)}$, the Gibbs sampling for generating $\theta^{(s+1)}$ from f is as follows:

(1) 1^{st} sample, draw a value $\theta_{(s+1)1}$ from $f(\theta_{(s+1)1} | \theta_{s2}, \ldots, \theta_{sp})$.

(2) 2^{nd} sample, draw a value $\theta_{(s+1)2}$ from $f(\theta_{(s+1)2} | \theta_{(s+1)1}, \theta_{s3}, \ldots, \theta_{sp})$. …

(3) j^{th} sample, draw a vale $\theta_{(s+1)j}$ from $f(\theta_{(s+1)j} | \theta_{(s+1)1}, \ldots, \theta_{(s+1)(j-1)}, \theta_{s(j+1)}, \ldots, \theta_{sp})$.

(4) Repeat above step until we reach p^{th} sample. And our $\theta^{(s+1)} = (\theta_{(s+1)1}, \theta_{(s+1)2}, \ldots, \theta_{(s+1)p})$.

Metropolis algorithm and MH algorithm

The Metropolis algorithm differs from the Gibbs sampling in that it introduces a proposal distribution J that can reject some proposed moves. The MH algorithm is a more general form of the Metropolis algorithm. Both Gibbs sampling and Metropolis algorithm are special cases of the MH algorithm. It proceeds in following ways:

(1) Generate θ^* from $J(.|\theta^{(s)})$

(2) Compute the acceptance ratio $r = p(\theta^*|y) / p(\theta^{(s)}|y) = \frac{f(\theta^*)}{f(\theta^s)}$ for the Metropolis algorithm with J is symmetric in the sense that J(x|y) = J(y|x). When it comes to the MH algorithm,

$$r = \frac{f(\theta^*)}{f(\theta^s)} \times \frac{J(\theta^{(s)}|\theta^*)}{J(\theta^*|\theta^{(s)})} \quad \text{(I-3)}$$

(3) Generate a value u from the U (0,1) distribution. If $u < r$, set $\theta^{(s+1)} = \theta^*$. Otherwise set $\theta^{(s+1)} = \theta^{(s)}$

The intuition behind the Metropolis algorithm is drawing a proposed value θ^* nearby the current value $\theta^{(s)}$ using proposal distribution J (Hoff, 2009). We want to know if θ^* should be selected as a sample by calculating $r = p(\theta^*|y) / p(\theta^{(s)}|y)$. If $r > 1$ which indicate that θ^* is more probable than $\theta^{(s)}$, and then we should include θ^* in the samples.

MCMC diagnostic

The mixing of parameters is an important criterion to assess the performance of MCMC method that could characterize the speed of chain moves around the parameter space (Hoff, 2009). Good mixing is indicated by the Markov chain dispersing around the population median that makes samples more likely to be independent. The faster the dependence decays in successive iterations, the quicker the MCMC method converges. Poor mixing can happen in the above three algorithms, but the generated values can stay stuck at the same value for many consecutive iterations only using the Gibbs sampling.

The above MCMC methods are based on the random walk process. Recently, a non-random walk MCMC, Hamiltonian Monte Carlo (HMC), is introduced by adopting

physical system dynamics instead of a probability distribution to propose further states in the Markov chain (Neal, 2011).

1.1.3 Full-waveform LiDAR data for tree species classification

Over the past decade, the utility of LiDAR data for identifying the tree species has become a popular topic with increasing availability of LiDAR data and development of processing methods. DR LiDAR data alone (Holmgren and Persson, 2004; Vaughn et al., 2012) or in conjunction with ancillary data such as multispectral imagery (Heinzel et al., 2008; Holmgren et al., 2008; Ke et al., 2010; Leckie et al., 2003), and hyperspectral imagery (Jones et al., 2010; Zhang and Qiu, 2012) have proved their potential for tree species identification by using variables such as percentile heights and spectral features extracted from these data.

The automatic individual tree segmentation is the general preprocessing step of tree species discrimination. The accuracy of segmentation is crucial for the subsequent tree species classification and other tree structural attributes such as crown width, tree height, basal area (Chen et al., 2007; Koch et al., 2006; Li et al., 2012; Popescu and Wynne, 2004). Most of tree segmentation studies are based on the LiDAR-derived canopy height model (CHM) using different algorithms such as the local maximum filtering algorithm (Popescu et al., 2002), watershed algorithms (Chen et al., 2006) and pouring algorithm with empirical geometric shape of trees (Koch et al., 2006), and some segment tree crowns directly from the point cloud (Li et al., 2012; Morsdorf et al., 2003; Wang et al., 2008). The accuracy of tree species identification is influenced by many factors, such as forest types and point density (Vauhkonen et al., 2011), while the extraction method has been shown to be the main factor for precise individual tree crown delineation (Kaartinen et al., 2012).

Likewise, the individual tree segmentation step is also pivotal to the tree crown delineation using FW LiDAR data. Currently, most studies using FW LiDAR data (Reitberger et al., 2008; Reitberger et al., 2009; Yao et al., 2012) also use the CHM derived from different algorithms as mentioned above to automatically segment trees. Specific examples of emerging FW LiDAR data for tree species identification mainly use

variables extracted from the waveform decomposition, such as echo width, amplitude, backscatter cross section or intensity (Hollaus et al., 2009a; Reitberger et al., 2008; Vaughn et al., 2012; Yao et al., 2012; Yu et al., 2014). The main advantage of the decomposition method is that the general view of shape for each waveform can be obtained by deriving the echo width, peak location and amplitude. Compared to DR LiDAR data, this extra information can be applied for tree species classification (Hollaus et al., 2009a; Reitberger et al., 2008). However, the rich information contained in the waveforms may automatically decrease during this step.

Several studies have explored the variables directly from the raw FW LiDAR data, named FW metrics, to investigate their potential for tree species identification using linear discriminant analysis (Heinzel and Koch, 2011), neural network, support vector machine (SVM) (Vaughn et al., 2012), maximum-likelihood method (Yao et al., 2012) or Random Forests (Cao et al., 2016; Yu et al., 2014). What these methods share is that they are mostly based on the deterministic model and uncertainty of these input variables are not taken into account. Therefore, we propose to integrate Bayesian inference with FW metrics to check whether a stochastic model can improve the accuracy of tree species identification and quantify the uncertainty of the estimation. A detailed description of the basic Bayesian concept is provided in Section 1.1.2.

1.1.4 Photon counting LiDAR

Over the past decades, PCL data have been successfully applied for ranging in various domains such as measuring surface elevation, roughness and slope of an ice sheet or sea ice in the cryosphere (Dabney et al., 2010). In contrast to analog LiDAR, the PCL is unique in that it requires low energy expenditure to transmit a power laser pulse and uses a highly sensitive detector which can capture any returned photon from the reflected signal or an event triggered by solar background within the detector. These advantages enable PCL systems suitable to generate dense along-track sampling (Zhang and Kerekes, 2014) and ultimately result in large spatial coverage (Wulder et al., 2012).

Unlike the Ice, Cloud and land Elevation Satellite (ICESat) using an analog, full waveform system, the upcoming Ice, Cloud and Land Elevation Satellite-2 (ICESat-2)

mission, with launch date in September 2018, will employ a PCL sensor to detect spatial variability of ice surface, monitoring ice dynamics (Herzfeld et al., 2014) and measuring vegetation canopy height over large areas. Compared to the ICESat, the ICESat-2 will generate overlapping footprints on the Earth surface with a diameter of 14 m, spaced at 0.7 m along the track, which is denser than the illuminated spots (footprints) of 70 m in diameter, spaced at 170 m intervals of the Geoscience Laser Altimeter System (GLAS) aboard the ICESat.

To clearly articulate the measurement process of different LiDAR systems, Figure I-1 demonstrates an example of three different laser detector modalities including FW, DR, and PCL for measuring vegetation. The main difference of full waveform and DR LiDAR systems lies in the recording logic, for example, full waveform systems can digitize and store the waveform samples along the pulse line while DR systems only record samples with their corresponding amplitude of reflected signal energy exceed a certain threshold, and thus it is insensitive to noise impact (Neuenschwander and Magruder, 2016). In contrast, PCL systems can capture any detected event including signal and noise within the footprint along the vertical distribution. Through accumulating a number of shots (>100) of a PCL sensor, a probability distribution of vertical locations of shots similar to waveform profile is generated as shown in Figure I-1. This histogram also highlights where the photons are actually reflected from and implies uncertainty and error of height measurement with PCL systems due to that the vertical sampling of one shot can occur in any high probability regions with one to three photons. However, this overview of the PCL's measurement process provides solid support for measuring height with dense sampling as expected from the ICESat-2 mission. The shapes or recorded observations of these LiDAR systems are influenced by the transmitted pulse, scattering surfaces, the interaction of pulse with atmosphere and the characteristics of the signal path in the receiver (Hovi, 2015).

A critical task for the ecosystem community is to identify the ground and canopy surface from these photons to meet the science objective of determining global canopy

heights hinges upon the ability to detect both the canopy surface and the underlying topography (Neuenschwander and Magruder, 2016).

Generally, there are two major steps to derive terrain and canopy height from PCL data: 1) noise filtering of raw signals, and 2) canopy and terrain classification of possible signals. The performance of noise filtering, on which canopy and terrain classifications depend on, may be of greater significance. Regarding the noise filtering, a detailed discussion of these methods is provided in Chapter V.

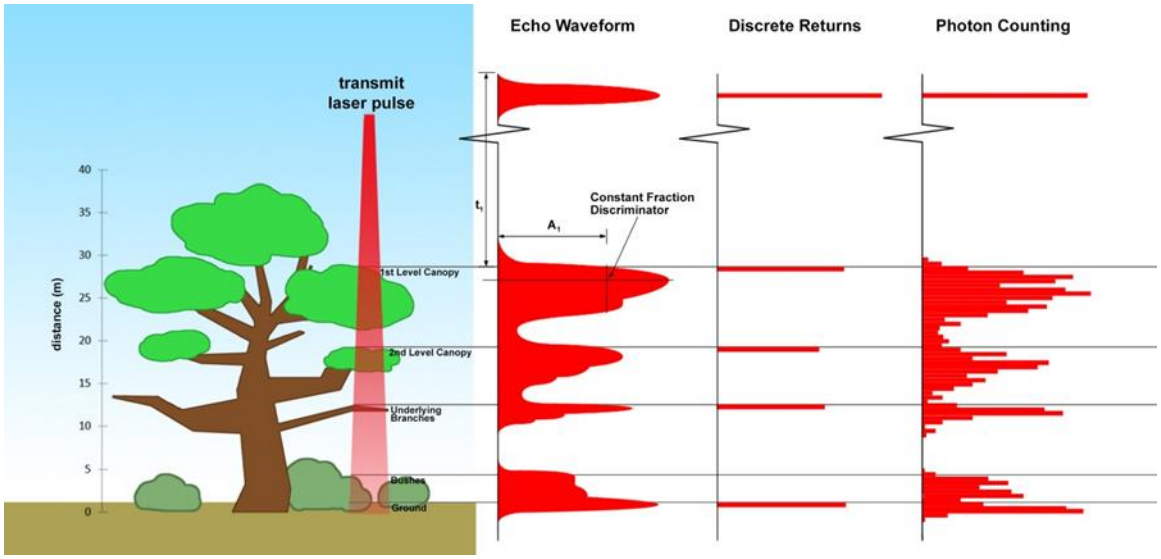


Figure I-1.Illustrations of different measurement processes of various LiDAR systems for vegetation detection (from (Neuenschwander and Magruder, 2016)).

1.2 Rationales

Largely available small footprint FW LiDAR data have gained the attention of researchers in forestry recently since a more transparent view of measurement process can be achieved. Owing to the fact that FW LiDAR data typically cannot be directly used like DR LiDAR data and scarce FW LiDAR data processing tools or algorithms are available for researchers to apply them to real-world tasks, developing open-source tools or novel algorithms is urgently needed. In addition, there is no justification for using the

Gaussian function to waveform decomposition from the model perspective and the uncertainty of waveform decomposition process also cannot be characterized. The Bayesian method can resolve these concerns and give us a thorough understanding of FW LiDAR processing. To validate the effectiveness of FW LiDAR data, we proposed a new way to discriminate tree species by integrating waveform metrics with machine learning methods such as the Random forest (RF) and Conditional inference forests (CF), and the Bayesian method. The upcoming ICESat-2 mission will provide a significant benefit to society through a variety of real-world applications such as the monitoring of the dynamics of sea ice, forest structural mapping, biomass estimation and improved estimation of Global Digital Terrain Models (GDTM). To pre-validate scientific objectives of the ICESat-2 mission for ecosystem community, we explored two test-bed sensor data, both of which are similar to the data expected from genuine ICESat-2 data to identify possible challenges of data processing for measuring canopy height and generate a basic methodological framework for processing ICESat-2 data.

1.3 Research objectives

The overall goal of this work is to develop algorithms for processing small footprint FW LiDAR and PCL data and investigate the potential application of FW LiDAR data by integrating FW metrics with advanced statistical methods for classifying tree species.

Specific research objectives are formulated as follows:

(1) exploring classical waveform decomposition methods and introducing optimized Gold deconvolution to process FW LiDAR data in various topography and vegetation conditions (Chapter II).

(2) applying the Bayesian concept for waveform decomposition and quantifying the propagation of error and uncertainty along the processing steps (Chapter III).

(3) integrating waveform metrics with machine learning (the Random Forests and Conditional inference Forests) and Bayesian methods to discriminate tree species using FW LiDAR data alone (Chapter IV).

(4) developing a methodological framework to derive terrain elevation and vegetation canopy height from test-bed sensor data which share similar characteristics of

the expected data from the ICESat-2 mission and further pre-validate the capacity of the mission to meet its science objectives for the ecosystem community (Chapter V).

CHAPTER II

GOLD – A NOVEL DECONVOLUTION ALGORITHM WITH OPTIMIZATION FOR WAVEFORM LIDAR PROCESSING*

Waveform Light Detection and Ranging (LiDAR) data have advantages over discrete-return LiDAR data in accurately characterizing vegetation structure. However, we lack a comprehensive understanding of waveform data processing approaches under different topography and vegetation conditions. The objective of this paper is to highlight a novel deconvolution algorithm, the Gold algorithm, for processing waveform LiDAR data with optimal deconvolution parameters. Further, we present a comparative study of waveform processing methods to provide insight into selecting an approach for a given combination of vegetation and terrain characteristics. We employed two waveform processing methods: 1) direct decomposition, 2) deconvolution and decomposition. In method two, we utilized two deconvolution algorithms - the Richardson-Lucy (RL) algorithm and the Gold algorithm. The comprehensive and quantitative comparisons were conducted in terms of the number of detected echoes, position accuracy, the bias of the end products (such as digital terrain model (DTM) and canopy height model (CHM)) from the corresponding reference data, along with parameter uncertainty for these end products obtained from different methods. This study was conducted at three study sites that include diverse ecological regions, vegetation and elevation gradients. Results demonstrate that two deconvolution algorithms are sensitive to the pre-processing steps of input data. The deconvolution and decomposition method is more capable of detecting hidden echoes with a lower false echo detection rate, especially for the Gold algorithm.

Compared to the reference data, all approaches generate satisfactory accuracy assessment results with small mean spatial difference (<1.22 m for DTMs, < 0.77 m for CHMs) and root mean square error (RMSE) (<1.26 m for DTMs, <1.93 m for CHMs).

*Reprinted with permission from "Gold–A novel deconvolution algorithm with optimization for waveform LiDAR processing." by Zhou, Tan, Sorin C. Popescu, Keith Krause, Ryan D. Sheridan, and Eric Putman. ISPRS Journal of Photogrammetry and Remote Sensing 129 (2017): 131-150. Copyright 2017 by Elsevier.

More specifically, the Gold algorithm is superior to others with smaller root mean square error (RMSE) (<1.01 m), while the direct decomposition approach works better in terms of the percentage of spatial difference within 0.5 and 1 m. The parameter uncertainty analysis demonstrates that the Gold algorithm outperforms other approaches in dense vegetation areas, with the smallest RMSE, and the RL algorithm performs better in sparse vegetation areas in terms of RMSE. Additionally, the high level of uncertainty occurs more on areas with high slope and high vegetation. This study provides an alternative and innovative approach for waveform processing that will benefit high fidelity processing of waveform LiDAR data to characterize vegetation structures.

Key words: Waveform LiDAR, Deconvolution, Gold, Richardson–Lucy (RL), Decomposition, Parameter uncertainty

2.1 Introduction

Full-waveform airborne laser scanners (ALS) are increasingly available to remote sensing data providers. The data acquired by such systems is widely applicable to vegetated ecosystem assessment and monitoring (Gwenzi and Lefsky, 2014; Hollaus et al., 2009b; Lefsky, 2010; McGlinchy et al., 2014). Waveform airborne laser scanning is an active form of remote sensing technique that could provide additional geometric and physical information of the scattering substance along the path. It also gives users more control of data interpretation (Chauve et al., 2007) compared with the conventional discrete-return Light Detection And Ranging (LiDAR) technique. The physical principle of the waveform LiDAR system is similar to the conventional LiDAR system, but waveform LiDAR system can record the entire echo scattered from illuminated objects with different temporal resolutions (such as 1/2/4 nanosecond(s) (ns)) through digital sampling.

The waveforms not only include responses from ground level, but also comprise multiple-scattering responses of illuminated surfaces along the laser line which give more information of objects through waveform shapes, widths, intensities and skewness. These characteristics of waveform LiDAR make its processing more difficult than discrete-

return LiDAR processing which only needs to combine the time between the emitted signal and received signal, the speed of light and geolocation information (like GPS, platform altitude of scanner). Through extracting and decoding these characteristics, the physical attributes of vegetation like canopy height (Gao et al., 2015; Gwenzi and Lefsky, 2014), target cross section (Roncat et al., 2011), stem volume (Reitberger et al., 2009), and above ground biomass of forest (Boudreau et al., 2008) can be modeled. Thus, gaining knowledge of the forest from waveform LiDAR data is a pivotal step toward efficiently and comprehensively understanding forest roles such as biomass change and carbon cycle under climate change.

Generally, the waveform processing method can be categorized into two types: one is the direct decomposition method and another is the deconvolution and decomposition method. Each method has been successfully applied for echo detection based on its own physical background. For direct decomposition, the emitted pulse is generally assumed to be Gaussian shape, as well as the scatterers' differential backscatter cross section (Wagner et al., 2006). The return waveform is obtained through the convolution of the Gaussian shape emitted pulse and the Gaussian substance scattering function (Wagner et al., 2006). Therefore, the Gaussian decomposition is the most frequently used approach to process the received signals (Mallet and Bretar, 2009). Many studies have been carried out on Gaussian decomposition to perform processing and analysis of different types of waveform LiDAR data such as LVIS (Laser Vegetation Imaging Sensor) data (Zhuang and Mountrakis, 2014), ICESat (Ice Cloud and land Elevation Satellite) data (Gwenzi and Lefsky, 2014; Harding, 2005; Keller, 2007; Lefsky et al., 2005), airborne data (Chauve et al., 2007; Hancock et al., 2017; Wagner et al., 2006; Wu et al., 2011). These researches have demonstrated that the Gaussian model is sufficient for processing waveform LiDAR data to characterize the vegetation structure, no matter whether its footprint is large or small. Fitting the waveform with the Gaussian function can provide peak amplitude, the width of each Gaussian component and peak location information. The peak amplitude can be used as a criterion to filter the points from below ground (Rowe, 2013) and it also provides us information about surface of objects along the pulse. The echo width has been

employed as tool to characterize the crown depth, crown variability, and topographic relief (Harding, 2005; Keller, 2007). The range or elevation of a specific reflecting surface can be calculated using the peak location provided through Gaussian decomposition (Hofton et al., 2000). Furthermore, these pieces of information combined can be utilized to estimate woody cover and biomass, classify tree species (Reitberger et al., 2008) and map land-cover etc. (Wang and Glennie, 2015).

The commonly used approaches to fit a sum of Gaussian functions are the Non-linear Least Square (NLS) method with Levenberg-Marquardt (LM) optimization algorithm (Hofton et al., 2000), the maximum likelihood methods (Persson et al., 2005) with the Expectation Maximization (EM) algorithm, and the Progressive Waveform Decomposition (PWD) method (Zhu et al., 2011). The limitations of the first two methods are the lack of prior knowledge about waveforms and the difficulties of identifying initialization of waveform parameters, such as the number of peaks, peak amplitude and width. Mallet et.al (2009) employed a stochastic method to reconstruct the waveform LiDAR by decomposing each echo with suitable functions like the generalized Gaussian function, Weibull function, Nakagami function and Burr function. These methods are robust and have shown good potential for applications in waveform processes and analysis (Chauve et al., 2007; Fieber et al., 2015; Reitberger et al., 2008). Whereas the pulse detection method such as the Average Square Difference Function (ASDF) (Zhu et al., 2011) may omit the important waveform parameters, the PWD method may lead to false echo detection due to the ringing effect.

Based on the standard LiDAR equations and real world constraints (Carlsson et al., 2001), the power of the received pulse can also be expressed as the sum of echoes from N targets with system and environment contributions (Mallet and Bretar, 2009). The direct decomposition method does not take into account the detector and system's contributions to the waveform, which results in the loss of illuminated surface information. To reduce unwanted system contribution and recover the true distribution of the illuminated surface, the second approach, the deconvolution and decomposition, is proposed. Several published studies have successfully applied different deconvolution

algorithms such as B-spline, Richardson-Lucy (RL), Non-negative Least Squares (NNLS), Wiener Filter (WF) (Cawse-Nicholson et al., 2014; McGlinchy et al., 2014; Neuenschwander, 2008; Roncat et al., 2011; Wu et al., 2011), sparsity-constrained regularization approach (Azadbakht et al., 2016) and Bayesian inference method (Jalobeanu and Gonçalves, 2014) to recover the true cross-sectional profile of an illuminated object.

Though the widely used deconvolution algorithms in the waveform LiDAR are RL (Lucy, 1974), NNLS and WF (Jutzi and Stilla, 2006), the studies of Nordin (2006) and Wu et al. (2011) have demonstrated that the RL algorithm is superior to other algorithms for the estimation of tree biomass and detection of unobservable peaks. The detailed information of the above three algorithms can be found in Wu's study (2011). However, each algorithm has its own advantages and limitations when applied to the deconvolution. For example, the RL and NNLS can lead to more accurate results at the expense of taking a longer time to complete the iterative process; WF may require less implementation time but results in less accurate solution. Additionally, developing open source tools for the waveform processing is also a pressing need for the extensive applications of waveform LiDAR data with different format. Hancock et al. (2008) proposed the Gold's method to process the large-footprint waveform GLAS (Geoscience Laser Altimeter System), however, the information contained in the large-footprint and small-footprint data is different (Mallet and Bretar, 2009). Large footprint data, e.g., up to 65m, have the waveform returned from multiple tree crowns and is affected by topography, especially on high slope terrain, therefore such data have a significantly different shape compared to small footprint waveform data that samples a small portion of tree crowns intersected by the laser beam, possibly not reaching the ground under dense vegetation. Thus, the Gold's method has not been proven in prior studies that it is suitable for small-footprint waveform LiDAR processing. Additionally, at the time of developing our study, we could not identify any publication that has conducted the parameter optimization of waveform deconvolution for vegetation applications. Results of waveform deconvolution depend significantly on the choice of parameters used with deconvolution functions.

To enrich the existing waveform processing methods and enhance the performance of the deconvolution, the optimized Gold algorithm described in Section 2.3.1 is proposed to reconstruct the differential backscatter cross section from the waveform LiDAR collected by the National Ecological Observatory Network (NEON). Meanwhile, there is a lack of quantitative and comprehensive comparisons of waveform LiDAR processing methods. The overall goal of this research is to propose a novel deconvolution approach to process waveform data and contribute to a better understanding of advantages and limitations of different small-footprint waveform LiDAR processing approaches. Specific objectives are to: (1) introduce a novel deconvolution algorithm, the Gold algorithm, which is a non-negative iterative solution toward generating more accurate and representative ground elevation and canopy height; (2) develop an optimization methodology for finding appropriate deconvolution parameters; (3) explore advantages and limitations of various waveform processing techniques to derive topography and canopy height information; (4) perform comprehensive comparisons of results with different approaches and assess each approach's accuracy and parameter uncertainty.

Our hypothesis is that better results are expected with the new algorithm in terms of the echo detection, accuracy assessment and parameters uncertainty analysis. The innovative aspects of this study consist of: (1) extending the current small-footprint waveform processing methods by adapting the Gold algorithm to process the SF waveform LiDAR data and then investigating their performance in different topography and vegetation conditions, (2) introducing an optimization process of determining the deconvolution parameters and (3) implementing processing steps within an open source software or tools such as R (2013) and LAStools (Isenburg, 2012).

2.2 Methodology

2.2.1 Study sites

This study was conducted using the data from three NEON terrestrial sites: (1) the Harvard Forest (HF), north-central Massachusetts; (2) the San Joaquin Experimental Range (SJER), north of Fresno, California and; (3) the Ordway-Swisher Biological Station (OSBS), near Melrose of Putnam County, Florida (Figure II-1). We hypothesize

that the performance of approaches will be affected by factors such as topography and elevation gradients. These study sites were selected to test the robustness of different approaches for processing waveform LiDAR data. They were extended over diverse ecological regions, climate and elevation gradients with different number of flight lines.

The HF is a core wild-land site and statistically represents unmanaged wildlife conditions across the NEON's 30-year history (Kampe, 2010). One flight line of the cropped waveform sample area is chosen. The data covers about 60m × 60m with the center located at 731156.6 Easting, 4712671.4 Northing, and UTM Zone 18N (Figure II-1A). This site primarily consists of dense mixed hardwood trees with dominant species being white pine (*Pinus strobus*) and red oak (*Quercus rubra*) in the center 20m × 20m area. The landscape is characterized by flat terrain with an elevation difference of approximately 5m.

The second site, the SJER, is located in the foothills of Sierra Nevada Mountains, about 32km north of Fresno, California. The cropped waveform sample is about 6.25 ha (250m × 250m) with the center at 256840.0 Easting, 4110820.0 Northing, and UTM Zone 11N (Figure II-1B). The elevation ranges from 380 to 425m dominated by sparse blue oak (*Quercus douglasii*), interior live oaks (*Quercus wislizeni*) and digger pine (*Pinus sabiniana*). The topography is complex with coarse, large hills and valleys.

The OSBS is located near Melrose of Putnam County, Florida with an elevation range from 21 to 48m. The cropped area is covered by four flight lines and it is about 60 ha (1172m × 534m) with the center at 402507.6 Easting, 3282045.1 Northing, and UTM Zone 17N (Figure II-1C). It is composed of homogenous forest dominated by Longleaf Pines (*Pinus palustris Mill*.) and Loblolly (*Pinus taeda*), areas of mixed patches of vegetation structure and heterogonous land cover types, including water body, wetland, open ground and road. These make the OSBS site well suited for comparing and testing performance of different processing waveform LiDAR processing over a range of simple to complex vegetation communities.

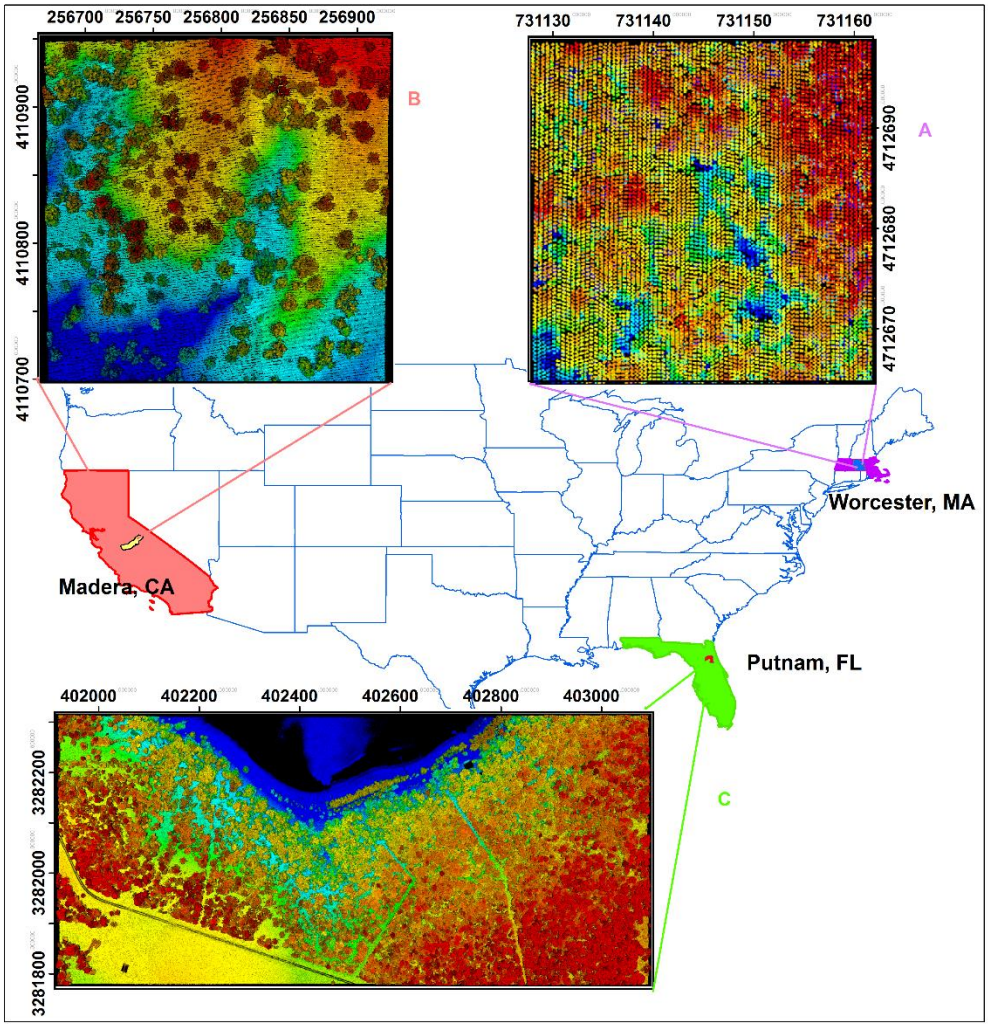


Figure II-1. Locations of three study sites in Massachusetts (one flight line), California (two flight lines) and Florida (four flight lines) with discrete-return LiDAR points.

2.2.2 LiDAR data

2.2.2.1 Waveform LiDAR Data

The three waveform LiDAR datasets were acquired with an Optech Gemini instrument at a nominal range of 1000 m (the aircraft flew at 1000 m above ground level). It achieved a nominal density of 3.82 points per square meter with a 0.8m diameter spot and a spot spacing of about 0.524 m in the across-track direction and 0.5m in the along-track direction. Both datasets were collected during the leaf-on condition on August 8,

2012 for the HF site, June 13, 2013 for the SJER site, and May 7 and May 19, 2014 for the OSBS site. All data were distributed by the NEON data center (http://www.neonscience.org/content/airborne-data). The detailed technical data specifications are shown in Table II-1.

There was one flight line with 13,902 waveforms included in the HF sample area. For the SJER site, two flight lines were available with 167,019 waveforms for flight line 03 and 91,648 waveforms for flight line 12. For the OSBS site, four flight lines were available for our study region as shown in Figure II-1C. The number of waveforms for four flight lines is 660,995, 859,919, 597,455 and 1,371,186, respectively. Each waveform was segmented into 500 time bins with 1 ns temporal spacing.

The waveforms stored the digital number (DN) of return pulses, which can be assumed to be the amplitude of the waveform. Simultaneously, NEON provided geolocation information and corresponding outgoing pulses that consist of 100 time bins with a temporal resolution of 1 ns. The geolocation data comprised Easting, Northing, height, dx (m), dy (m), dz (m), and first return bin location. And dx, dy, and dz were the pulse direction vector that can be used to calculate the accurate geolocation of any other time bins in a given waveform without registration and rectification. All data were zero padded. NEON also provided us with a prototype system impulse which was a return pulse of single laser shot from a hard ground target with a mirror angle close to nadir and corresponding outgoing waveform (Figure II-2). This can help us remove the outgoing pulse and system response effect and perform a deconvolution on the waveform.

2.2.2.2 Reference data

To validate the performance of methods and end products of the waveform LiDAR, the discrete-return LiDAR data and the Digital Terrain Models (DTMs) and Canopy Height Models (CHMs) provided by NEON were used as the reference data. According to the NEON's discrete-return LiDAR Algorithm Theoretical Basis Document (ATBD), the maximum horizontal accuracy of discrete-return LiDAR is about 0.4m, with maximum LiDAR vertical accuracy 0.36m, respectively (Keith and Tristan, 2015). Discrete-return LiDAR data will be used as reference in the Number of echoes section

test whether more points were extracted from the waveform LiDAR data. Additionally, we compared the waveform-based end products such as DTMs and CHMs with corresponding reference data provided by the NEON to conduct the accuracy assessment of our approaches.

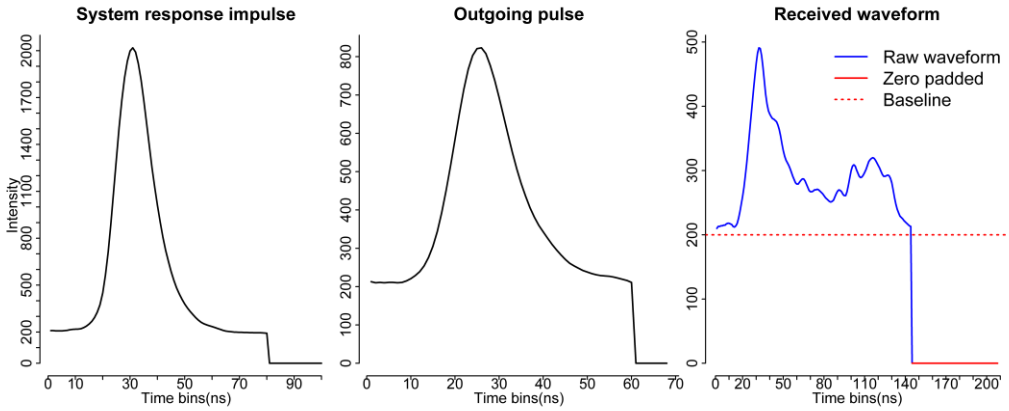


Figure II-2. A subset of the system response impulse, a sample of outgoing pulse and corresponding raw waveform recorded by NEON's full waveform LiDAR system.

Table II-1. Main technical specifications of the NEON waveform data.

Study Sites	Technical specifications
Operating altitude	~1000 m
Wavelength	1064 nm
Pulse repetition frequency	100 kHz
Scan frequency	50 Hz
Beam divergence	0.8 mrad
Scan angle range	±18.5°
Spot spacing	0.524 m(across-track), 0.5 m(along-track)
Footprint size	3.83 points/m2 (0.8m)
Digitizer	1 ns (12 bit A/D , baseline signal is 200)
Outgoing pulse width	~ 14 ns
Flying direction	HF: East, heading 90 degree
	SJER & OSBS: Northing to South or South to North, heading 180 or 0 degree

2.2.3 Waveform processing

Waveform processing involves a series of steps, including noise detection, smoothing, radiometric calibration (Briese et al., 2008), deconvolution and decomposition, etc. Many studies have been conducted to interpret the waveform data and used it to estimate vegetation structure and function, such as canopy height and above-ground biomass (Chauve et al., 2007; Gwenzi and Lefsky, 2014; Roncat et al., 2011; Wagner et al., 2006). The major steps in the waveform processing used by these studies are signal deconvolution and decomposition. The deconvolution is an algorithm-based process that is used to reverse the effect of convolution on the recorded signals, and the decomposition is a process which can provide estimates of the location and properties of objects along the pulse (Wagner et al., 2006). In this study, we employed two distinct methods to process the waveform data. The first method was direct decomposition (I), which only applied the Gaussian decomposition to the waveform data. The second method was the deconvolution and decomposition method. For the second method, we utilized both new and classical deconvolution algorithms, the Gold and RL algorithms, to explore the advantages and limitations of deconvolution algorithms. The results from the above deconvolution were then subjected to Gaussian decomposition and we refer to them as the Gold approach (II) and RL approach (III) in this study. The RL algorithm was selected as a reference deconvolution algorithm because it is a superior widely used deconvolution algorithm (Harsdorf and Reuter, 2000; Nordin, 2006; Wu et al., 2011). Before performing analysis, we converted all delivered waveform data into ASCII format and then pre-processed them with steps, such as de-noising and mean filtering.

2.2.3.1 Parameter optimization for deconvolution

The returning pulses of the waveform were the product of interaction among outgoing pulses, atmospheric scattering, system noise and reflecting surfaces (Briese et al., 2008). For the NEON waveform LiDAR data, the backscatter responses can be expressed as the convolution of the outgoing pulse, impulse response (atmospheric

scattering, system noise, etc.) and effective target cross section (Eq.(II-13)). The deconvolution approach can remove the effect of the outgoing pulse and system impulse response and then improve the separability (Neuenschwander, 2008) of close peaks and reveal the true distribution of the scattering substances ($\delta_i(t)$) along the optical path (Cawse-Nicholson et al., 2014; Wagner et al., 2006; Wu et al., 2011).

$$P_r(t) = \sum_{i=1}^{n} \frac{D^2}{4\pi\gamma^2 {R_i}^4} P_t(t) * \tau(t) * \delta_i(t) \quad \text{(II-1)}$$

where $P_r(t)$ is the received laser power, $P_t(t)$ is the outgoing pulse, $\tau(t)$ is the receiver impulse function, D is the aperture diameter of the receiver optics, λ is the wavelength, R is the range from the LiDAR system to the target, $\delta_i(t)$ is effective target cross section and n is the number of targets detected along the pulse line.

The deconvolution is sensitive to the pre-processing of input data and the choice of parameters for deconvolution. For our optimization of processing parameters, we randomly selected 2000 waveforms of each dataset and processed this subset in two different ways with all other parameters as constant: one method is to keep all input data (the outgoing pulses, impulse response and return pulses) as raw values as supplied by the data provider; another method is to normalize the input data by subtracting the minimum non-zero value of each dataset. Through comparing the peaks' location of deconvolved waveforms with raw waveforms, we determined which method was employed in the subsequent analysis.

The parameters of the deconvolution function used in this study include boost, iterations and repetitions. The experiment demonstrated that boost was not as sensitive as the other two parameters and its recommended range was 1 to 2. Detailed information of these parameters can refer to the R package Peaks (Morhac, 2012). In our case, 1.5 was selected as a constant in the subsequent analysis. The number of iterations and repetitions in estimated impulse response and deconvolution algorithm are critical to the performance of the deconvolution and decomposition method. As such, our second step was to optimize these parameters for deconvolution. This step could be highly subjective and case-dependent for different datasets. To avoid subjectivity and a trial-and-error approach, we developed a general rule to find the reasonable range of these parameters.

The total number of echoes generated from these 2000 waveforms and the percentage of matched waveforms were selected as criteria to narrow down the parameter ranges. The matched waveforms were defined as the deconvolved waveform with the similar peak locations (the difference within 3 ns) compared to the peak locations of raw waveforms after using the mean filter.

The preliminary experiment was conducted on about 2300 combinations of these parameters, and then we selected the potential parameter combinations based on the criteria above. In our case, we eventually narrowed down the estimate impulse response's iteration and repetition to 15-30 and 2-4, respectively. For the iteration and repetition in the deconvolution algorithms, the range for them was 30-55 and 3-5, respectively.

In this study, the Gold algorithm (Morhac et al., 1997) and RL algorithm were employed to deconvolve the raw waveform data, and then we compared these two algorithms in terms of detection of peaks, false detection rate, accuracy assessment, and parameter uncertainty analysis. The following two sections provide the principles of these two algorithms.

Richardson–Lucy (RL) algorithm

The RL algorithm was developed from the Bayesian's theorem which could reconstruct noisy images by taking into account statistical fluctuations in the signal (Fish et al., 1995). It was originally developed for recovering the image by searching iteratively for solutions to deconvolution problems. The basic idea is to calculate the most likely value $f_t(x)$ given the observed $d(x)$ and the known point-spread function $g(x)$. One waveform LiDAR profile can be seen as an image with the dimension $1 \times N$ and t^{th} iteration solution written as follows in terms of convolution (Fish et al., 1995):

$$f_{t+1}(x) = f_t(x) \cdot \left(\frac{d(x)}{f_t(x) * g(x)} * g(-x)\right) \quad \text{(II-2)}$$

where * is the convolution operation, d(x) is the observed value at location x, $f_t(x)$ is the most likely value at location x and g(x) is the known point spread function, f(x) can be solved by iterating Eq. (II-2) until convergence.

Gold algorithm

The Gold algorithm is a non-oscillating and stable deconvolution method that can give us non-negative solutions (Morháč et al., 2003). This vital property is suitable for the waveform processing, since it is unreasonable and senseless if negative solutions appear. The Gold algorithm has been successfully applied to deconvolve the γ-ray spectra (Morhac et al., 1997) and nuclear data (Morháč et al., 2003). For discrete values, it searched iteratively to solve the deconvolution problem using Eq. (II-3):

$$y(i) = \sum_{k=0}^{n-1} h(i-k)x(k) \qquad i = 0, 1, 2, \dots, m \qquad \text{(II-3)}$$

where x, y are input and output vector, h(i) is the impulse response or outgoing impulse function, n is the number of samples of vector h, i is i^{th} sample point and x(k) represent the k^{th} waveform's differential backscatter cross-section. Here, the convolution system is one dimension. After a matrix transformation, the Gold algorithm can be expressed as Eq. (II-4) (Morhac et al., 1997):

$$x^{(k)}(i) = \frac{x^{(k-1)}(i)}{\sum_{j=1}^{n} h(i-j)x^{(k-1)}(j)} x^{(k-1)}(i) \qquad \text{(II-4)}$$

For both algorithm applications, the impulse response and outgoing pulses must be known first. The NEON waveform LiDAR data provided each waveform with the corresponding outgoing pulse. In this study, three major steps were utilized to obtain the effective target cross section ($\delta_i(t)$): (1) the system response was used to deconvolve the corresponding outgoing waveform for each dataset to derive the estimated impulse response. (2) using the return waveform to deconvolve the outgoing pulse to get the immediate waveform. (3) the immediate waveform was employed to deconvolve the estimated impulse response to reveal the effective target cross section. After deconvolution, the results from the above were also decomposed using a mixture of Gaussian function. The major steps were the same as described in section 2.3.2.

2.2.3.2 Gaussian decomposition

Since the outgoing laser pulse of the NEON data is nearly Gaussian in shape, as shown in Figure II-2, the returned waveform can be fitted by a mixture of Gaussian function (Wagner et al., 2006). The Gaussian components characterize the different targets when the laser beam interacts with objects along the path like vegetation and

ground (Harding, 2005). The analytical expression of the Gaussian function (f(x)) can be written as:

$$f(x) = \sum_{i=1}^{n} A_i \exp\left(-\frac{(x-u_i)^2}{2{\delta_i}^2}\right) \quad \text{(II-5)}$$

where n is number of Gaussian components, A_i is the amplitude of peak at ith waveform component, δ_i is the standard deviation of ith waveform component, and u_i is the time location of peak at ith waveform component.

In this study, a mean filter was first performed on each individual waveform to remove the noise and then we normalized these waveforms through subtracting the minimum value of each waveform. The return pulses were fitted with a mixture of Gaussian functions using a NLS method and optimized using the LM algorithm, which was implemented in the R package minpack.lm (Elzhov et al., 2013). Detailed steps can refer to the study of Chauve et al. (2007). The difficult part of using NLS to fit the Gaussian function was to determine the initial values of parameters. For the direct decomposition, we estimated the number of Gaussian components (n) by finding the number of peaks of raw waveform data and then used amplitude threshold to delete the "fake" peak(s). The initial amplitude (A_i) was assumed to be 2/3 of the corresponding peak value, initial echo width (δ_i) of each component was estimated as the half-widths of consecutive peaks and initial peak locations (u_i) were derived from corresponding raw waveforms' peak locations. For the deconvolution and decomposition, the initial n, u_i and A_i were derived from corresponding deconvolved waveform with the same step as the direct decomposition. However, the initial δ_i for each deconvolved waveform was estimated as half of the difference between peak location and the closest time bin with negative lagged difference.

2.2.3.3 Geolocation Extraction

Through deconvolving the waveforms and fitting echoes to a mixture of Gaussian function with the LM algorithm, the time of peak positions, intensity and width were obtained. The 3D point cloud was generated based on the time of peak positions, the location of the first time bin of the return pulse and position change of pulse per nanosecond (dx, dy, dz).

For the direct decomposition method, the leading edge position of each waveform was used to calculate the geolocation of any time bin in a return waveform by incorporating the full width at half maximum (FWHM). The waveform is fitted with Gaussian function, so FWHM can be obtained through the standard deviation (σ). Therefore, the leading edge position can be calculated by the Eq. (II-6) and (II-7).

$$FWHM = 2\sqrt{2ln2}\,\sigma \tag{II-6}$$

$$t_l = t_p - 0.5 * \text{FWHM} \tag{II-7}$$

where t_l is the leading edge position for each echo, t_p is the time of peak position for each echo and σ is the standard deviation of individually fitted function. t_p and σ can be obtained from Gaussian decomposition. Then the new geolocation of any time bin for given waveform can be calculated by the Eq. (II-8).

$$\left.\begin{aligned} X &= (t_l - t_r) * dx + X_r \\ Y &= (t_l - t_r) * dy + Y_r \\ Z &= (t_l - t_r) * dz + Z_r \end{aligned}\right\} \tag{II-8}$$

where X, Y, Z is the new geolocation of peak, t_r is the first return reference bin location, dx, dy, dz are the position change for every ns, X_r, Y_r, Z_r are the Easting, Northing and height of the first return. t_r, dx, dy, dz, X_r, Y_r, Z_r are provided by the NEON geolocation dataset.

The new geolocation was determined by using the time of peak location (t_p) for the deconvolution and decomposition method, since the deconvolution can reveal the real geometry of objects. The NEON datasets provided us another geolocation dataset for deconvolution and decomposition method. The new geolocation is computed as:

$$\left.\begin{aligned} X &= [(t_p - t_r) - (t_{op} - t_{or})] * dx + X_r \\ Y &= [(t_p - t_r) - (t_{op} - t_{or})] * dy + Y_r \\ Z &= [(t_p - t_r) - (t_{op} - t_{or})] * dz + Z_r \end{aligned}\right\} \tag{II-9}$$

where t_{op} is the time of peak location for the each outgoing pulse and t_{or} is the time of corresponding reference bin location for the outgoing pulse. Both can be found in the NEON geolocation datasets.

2.2.3.4 Digital models extraction

The original point cloud derived from the decomposition step had some noisy points, since some raw waveforms were not exactly Gaussian shape. We filtered these

points based on the intensity and height, which was achieved through the LAStools(Isenburg, 2012) .

Digital terrain models (DTMs) and canopy height models (CHMs) were generated from these filtered 3D point cloud for each method using the LAStools (Isenburg, 2012). DTMs and CHMs were chosen mainly because sets of vegetation metrics were derived from CHMs, and any error of DTMs would propagate to affect the accuracy of CHMs. To derive a DTM from waveform LiDAR data, the point cloud has to be classified into ground and non-ground points. Generally, the intensity and width of the last echoes can be used as criteria to remove the non-ground laser points. The intensity of the echoes can provide additional information about the reflectance properties of an object, such as judging whether the echoes came from below ground response (Chhatkuli et al., 2012). The echo width not only gives information on the range distribution of individual object that produce a single echo, but it also can assist to decide whether a pulse was reflected from solid ground or vegetation (Doneus et al., 2008; Ioannides et al., 2006). In this study, we employed the intensity and width of echoes to filter the non-ground points or noise. After exploring different thresholds of width and intensity of echoes, a width threshold of 20 was applied to remove noise or wrongly fitted echoes, and 1/10 of the corresponding maximum intensity of each waveform was selected as a threshold to remove the below ground response. These thresholds may not be universally valid, as they may vary by regions and types of data. Finally, the filtered points were imported into the LAStools to generate a refined DTM. To further evaluate the performance of the methods, the CHM was generated from the non-ground points based on the steps described by Khosravipour (2014).

2.2.4 Algorithms' performance comparisons

2.2.4.1 Accuracy assessment

The discrete-return LiDAR data and its derived end products can be regarded as the ground truth data due to that they can potentially achieve better accuracy than direct field measurement (Chen, 2007). In this study, the discrete-return LiDAR data were adopted as reference data to verify whether more points can be extracted from the waveform

LiDAR data. Additionally, we employed the reference DEMs and CHMs provided by the NEON to conduct accuracy assessment of our results generated from waveform LiDAR data from both visual and statistical perspectives. The SJER site was selected to show the visual comparison results for its complex topography. The results of the statistical comparisons were measured by mean difference, standard deviation (SD), root mean square error (RMSE) and percentage of spatial difference within 0.5, 1, 2 and >2m of each study site.

2.2.4.2 Parameter uncertainty

The predictive parameter error estimates not only enable us to objectively quantify the expected quality of the results from available data but also allow us to estimate the rigorous error propagation through to the end products. Through the approaches we employed in this study, the standard error (se) of peak location for each estimated echo was obtained. Parameter uncertainty was represented by the 95% confidence interval (95% CI) of peak location. For each echo, the estimated peak location's confidence bounds were calculated using Eq. (II-10) and Eq. (II-11):

$$t_{ll} = t_l - 1.96 * se \quad \text{(II-10)}$$

$$t_{lu} = t_l + 1.96 * se \quad \text{(II-11)}$$

The above equations were for the direct decomposition approach, while t_l became t_p for the Gold approach and RL approach. The above biased peak locations (t_{ll}, t_{lu}) would form two datasets for each approach: the 95% lower confidence level dataset (Lower dataset) and 95% upper confidence level dataset (Upper dataset). Once six DTMs and six CHMs for these three approaches were generated using these uncertainty datasets, we compared them with the DTMs and CHMs derived from the original decomposition to get the biases and quantitatively assess the approaches' robustness. The parameter uncertainty for each method was calculated as follows:

$$LU = EGE_L - EGE \quad \text{(II-12)}$$

$$UU = EGE_U - EGE \quad \text{(II-13)}$$

where EGE was the estimated ground elevation generated from dataset of peak locations, EGE_L was the estimated ground elevation generated from Lower dataset of peaks locations, EGE_U was the estimated ground elevation generated from Upper dataset of peak

locations, LU was the lower uncertainty and UU was the upper uncertainty. The uncertainty of maximum CHM was calculated in the same way. The visual comparisons of parameter uncertainty were also conducted using the site which had the largest uncertainty based on the statistical results.

In order to quantify the effect of factors such as slope and vegetation height on the parameter uncertainty, the uncertainty was divided into three levels: high (> 2.00m and < -2.00m), medium (-2.00 to -0.51m and 0.51 to 2.00m) and low (-0.5 to 0.5m). The slope and vegetation height for each corresponding level were also extracted. The Upper dataset and Lower dataset of SJER site were combined for each approach. The Analysis of variance (ANOVA) was used to analyze the effect of factors like slope and vegetation on uncertainty levels, and identify which region may be more likely to have higher uncertainty. Box plots were chosen to visualize the uncertainty's statistical results. An overview of the whole waveform processing and comparisons procedure is given in Figure II-3.

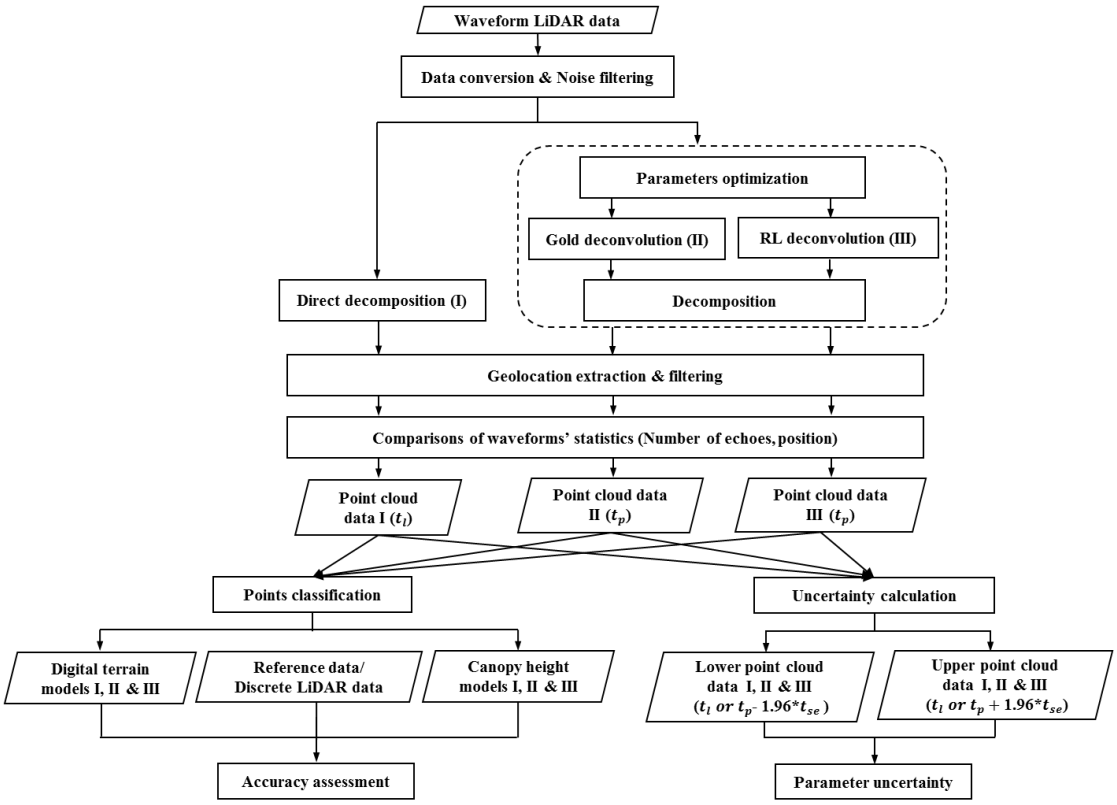


Figure II-3. Flowchart for waveform LiDAR data processing and comparisons

2.3 Results and discussion

2.3.1 Deconvolution parameters optimization

Since deconvolution was sensitive to the input data, we first explored different pre-processing steps of the input data. The three sample results of deconvolution by the RL and Gold algorithms were plotted in Figure II-4, with different pre-processing steps against corresponding original waveforms after noise deletion.

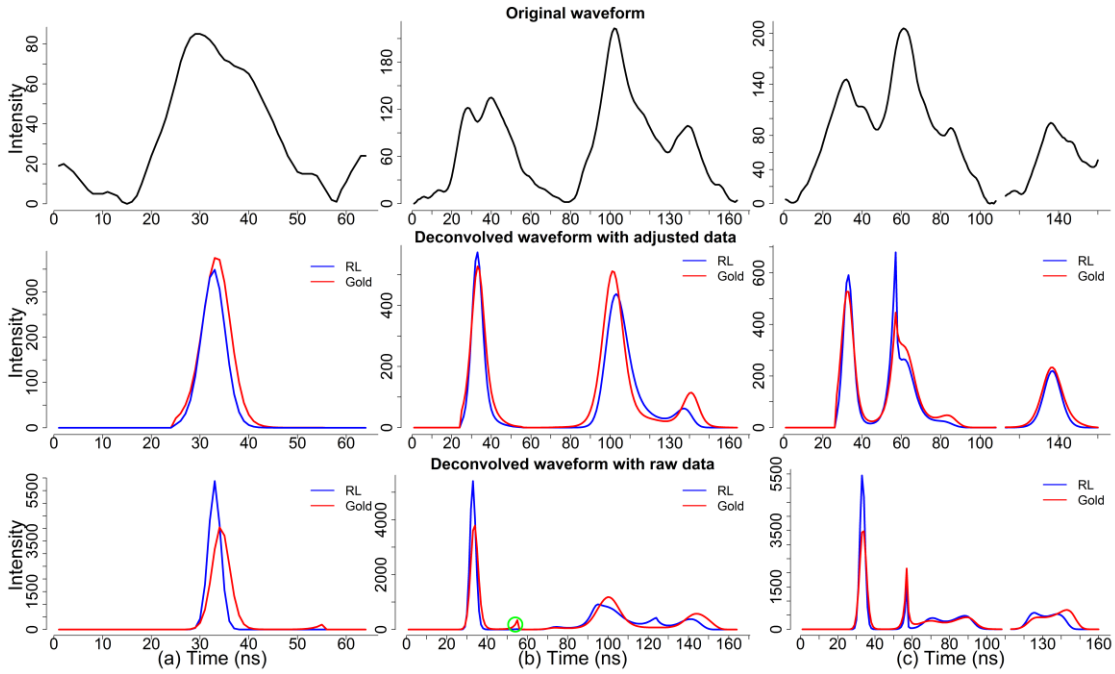


Figure II-4. Three samples of original waveforms (a, b and c) versus deconvolved waveforms by the RL algorithm and Gold algorithm with different pre-processing steps: *Deconvolved waveform with adjusted data* where the input data (the outgoing pulse, impulse response and return pulse) of deconvolution were normalized by subtracting minimum non-zero values; *Deconvolved waveform with raw data* where the input data (the outgoing pulse, impulse response and return pulse) of deconvolution were raw pulse values, excluding the zero padded values. The blank section in the original waveform of Figure II-4(c) resulted from unrecorded values in the raw data.

Figure II-4 shows that the peaks of the waveforms with deconvolution were much easier identified and that the shape of waveform components was closer to the Gaussian distribution. This verified our assumption that the waveform can be simulated by a mixture of the Gaussian components after removing the system impulse's contribution. Also, the intensity of the waveforms was much higher and the width of the waveform was narrower than the raw waveform, which could be conducive to precisely revealing the geometry of objects along the pulse.

It was evident that input data with different preprocessing steps could impact the deconvolution results. Compared with the original waveform, the results using the raw data were more likely to detect a wrong peak (green circle) as shown in Figure II-4(b).

Therefore, we adopted adjusted data to do the subsequent steps in this study, which was consistent with the same pre-processing steps described by Wu (2011). However, when using the adjusted data, we observed a downward shift in the time bin axis for both deconvolution algorithms. This was also found in the study of Cawse-Nicholson et al. (2014). The main reason may be that the system impulse was not acquired under the ideal condition, which led to the inaccurate deconvolved impulse response. However, this kind of inaccuracy of the deconvolution could be improved by obtaining more accurate system impulse.

In Figure II-4, the original waveform had noise in both edges that made the direct decomposition very difficult when we had no knowledge about the number of the reflecting objects along the pulse line. But, after removing the system response, the number and position of the peaks were evident. As shown in Figure II-4(b) and (c), the local peaks could also be clearly distinguished after deconvolution for more complicated waveforms. However, it should be noted that the deconvolution cannot break down pulses from surfaces that are too close which could result in one larger pulse as shown in the Figure II-4(b). This kind of peaks overlaying could be solved by increasing the number of iterations and repetitions in the deconvolution algorithm, but this may cause additional minor peaks that are adjacent to the major local peak as described in the study of Wu et al. (2011). Therefore, to set up the optimal number of iteration was critical to the detection of peaks. In this study, exploratory analysis helped us narrow down the range of parameters and final optimal combinations of iteration and repetition were (30, 4), (35, 5), and (40, 5) for the HF, SJER and OSBS sites, respectively.

2.3.2 Number of echoes

To provide a more comprehensive comparison, the quantitative results for the three study sites are shown in Table II-2 and Table II-3, respectively. For the deconvolution and decomposition method, the performance of the Gold algorithm was better than the RL algorithm from the perspective of the number of echoes detected. Additionally, the Gold algorithm mostly detected higher number of echoes with lower false detection rate. For instance, more waveforms were decomposed into three, four or five echoes for

individual waveform using the Gold algorithm than the RL algorithm for all study sites. This indicated that the Gold algorithm had a higher potential to detect the hidden peaks than the RL algorithm for complex waveforms. The direct decomposition's performance was similar to the RL algorithm in terms of the number of echoes detected. The direct decomposition was more capable of detecting a higher number of echoes than the RL algorithm for those study sites when we compared the number of waveforms with different echoes. However, the likelihood of detecting the false echoes was increased with the direct decomposition method compared with the deconvolution and decomposition method (Table II-2 and Table II-3).

Table II-2. Number of echoes estimated by direct decomposition method and deconvolution and decomposition method with different input data for HF site

Methods	Direct decomposition	Adjusted data Deconvolution +Decomposition			Raw data Deconvolution +Decomposition	
Number of echoes		RL algorithm(NA)[a]	Gold algorithm(NA)[a]	Gold algorithm(0)[b]	RL algorithm(NA)[a]	Gold algorithm(NA)[a]
	Number of waveforms					
1	5,533	6,057	4,261	3,160	8,438	7,454
2	4,313	3,088	2,783	2,990	3,164	4,024
3	2,398	3,493	5,049	5,626	1,336	1,438
4	685	322	765	1,088	150	169
5	133	116	221	214	4	7
6	23	13	13	14	0	0
7	7	3	0	0	0	0
False echoes	2,508 (10.05%)	370 (1.50%)	207 (0.71%)	296	56	102
Total echoes	24,945	24,679	29,217	31,524	19,394	20,527
Effective echoes	22,437	24,309	29,010	31,228	19,348	20,425

a is the intensity with an original value of 0 assigned to NA after deconvolution

b is the intensity of each pulse maintained as 0 after deconvolution

Interestingly, the false echoes rate of the SJER site was lower than the HF and OSBS sites for all approaches (Table II-2 and Table II-3). Comparing the three sites' decomposition results with their corresponding discrete-return LiDAR data demonstrated that the false echo detection rate of the SJER site was highest. The possible reason is that the SJER site was comprised of dense vegetation that could result in weak returns and overlapping echoes of the reflected waveforms, and potentially higher false echo detection rate was expected (Chauve et al., 2009). Another interesting finding was that

the deconvolution and decomposition method can reduce the false echoes rate significantly as shown in the Table 2 and Table 3.

Table II-3. Number of echoes estimated by direct decomposition method and deconvolution and decomposition approach with different input data for SJER and OBOS sites.

Methods	**SJER Adjusted data**		
	Direct decomposition	Deconvolution + Decomposition	
Number of echoes		RL algorithm	Gold algorithm
	Number of waveforms		
1	197,304	204,129	174,060
2	41,184	33,047	55,748
3	14,304	16,197	19,243
4	4,392	4,614	6,862
5	1,166	620	1,993
6	267	59	551
7	42	1	171
8	7	0	32
9	1	0	6
10	0	0	0
11	0	0	1
12	0	0	0
False echoes	17,913 (5.15%)	378 (0.11%)	1,223 (0.32%)
Total echoes	34,7943	340,731	385,522
Effective echoes	330,030	340,731	384,299

OSBS Adjusted data		
Direct decomposition	Deconvolution + Decomposition	
	RL algorithm	Gold algorithm
	Number of waveforms	
2,528,608	2,908,389	2,310,732
710,240	327,386	350,303
202,117	46,246	245,509
41,085	4,490	134,299
6,537	306	67,185
866	18	19,544
81	2	4,976
10	0	1,861
0	0	632
0	0	160
0	0	28
0	0	3
563,974 (11.85%)	14,709 (0.40%)	66,226 (1.38%)
4,758,307	3,736,220	4,861,828
4,194,333	3,721,511	4,795,602

Hence, based on the visual inspection and quantitative comparisons of different methods, we concluded that the deconvolution and decomposition method outperformed the direct decomposition method. The Gold algorithm was superior to the RL algorithm based on the number of echoes and false echo detection rate. We also found that pre-

processing of data significantly affected the echo detections; assigning zero-padded values of the return pulse to NA and utilizing the adjusted input data (the outgoing pulse, system response pulse and return pulse) could achieve better and more accurate results.

2.3.3 Position of echoes

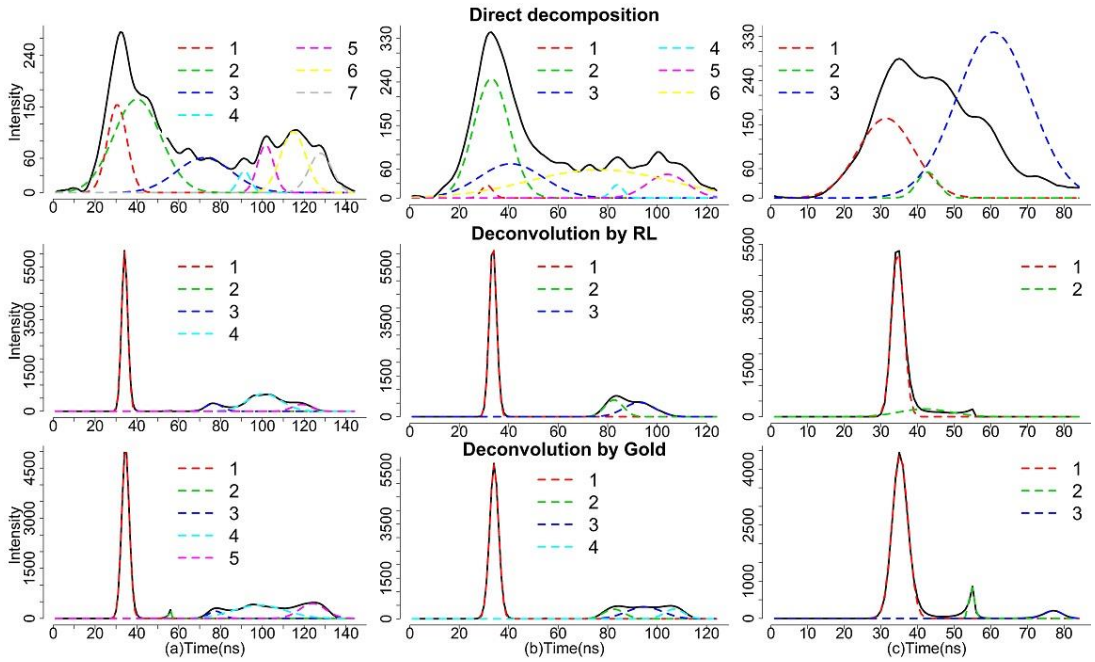


Figure II-5. Comparisons of the decomposition results with the direct decomposition approach, RL approach and Gold approach for three sample pulses (a, b, c). The solid black line is the original waveform. The colored dash lines are Gaussian components after decomposition.

To further demonstrate the performance of different approaches, we also explored the position of echoes. Figure II-5 shows that the three sample waveforms were decomposed by the three approaches: direct decomposition approach, RL approach and Gold approach. Here, the direct decomposition approach detected a higher number of echoes for each pulse than the other two deconvolution approaches, as shown in Figure II-5(a), (b) and (c) with 7, 6 and 3 Gaussian components, respectively. However, a closer examination revealed that the direct decomposition was likely to detect some unreasonable echoes. For example, the red Gaussian component in the direct

decomposition of Figure II-5(b) is almost overlaid with the green component; the yellow component looks like a supplement to the other echoes and neither agrees well with the reality. Furthermore, the blue component in the direct decomposition of the Figure II-5(c) is higher than the original waveform. These may explain why there was higher false echo detection rate for direct decomposition as shown in Table II-2 and Table II-3.

It was evident that deconvolved waveforms performed better on decomposition with explicit Gaussian components in terms of visual comparisons. The performances of the RL and Gold approaches were similar in our example and almost had the same peak positions and shape, but the Gold approach worked better when the original waveform was composed of many peaks with noise. As shown in Figure II-5, the Gold approach could detect more echoes for the same pulse and may reconstruct more accurate cross sections of vegetation and terrain. The RL approach was less capable of detecting the last echoes that may represent the ground when we interpreted the waveform LiDAR data. However, it was worthy to note that the Gold approach may cause the ringing effect as shown in the second Gaussian component (green part) of the deconvolution by the Gold algorithm (Figure II-5(a)). This kind of minor peak around the major local peaks may be caused by a wavelike artifact that resulted from the sum of remaining low-frequency components after the loss of high-frequency components (Wu et al., 2011).

2.3.4 Point clouds

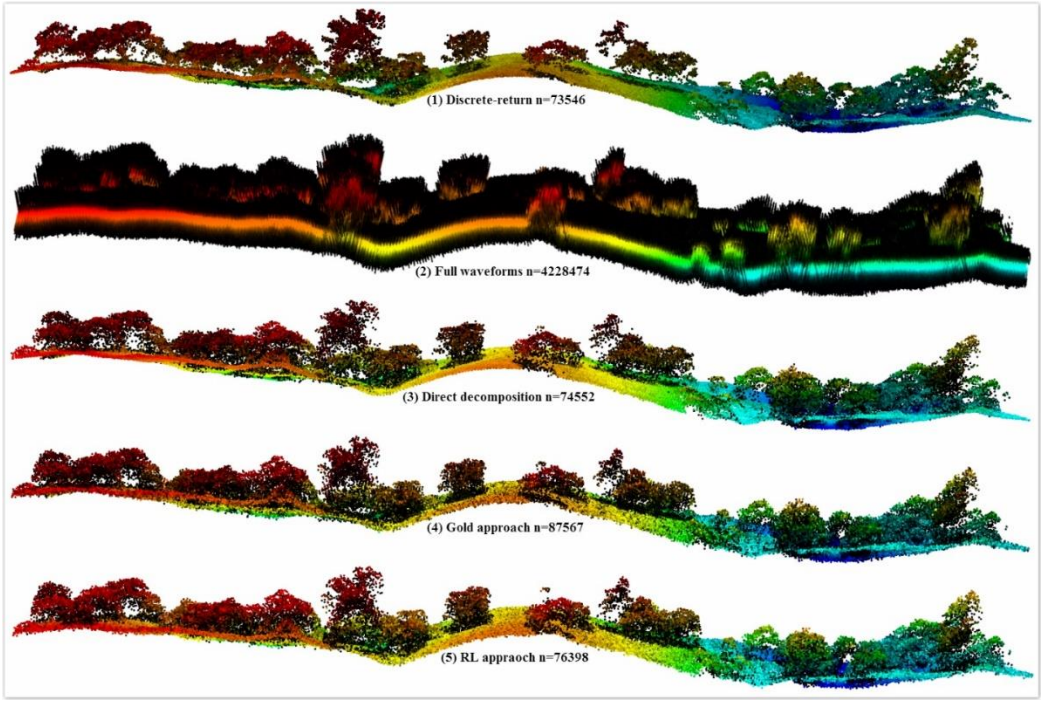


Figure II-6. Comparison of point clouds generated from discrete-return LiDAR data and waveform LiDAR data with different processing methods. (1) Point cloud derived from discrete-return LiDAR data. (2) Point cloud generated by geo-referencing every time bin of waveforms. (3) Point cloud derived from direct decomposition method. (4) Point cloud derived from Gold deconvolution and decomposition method (the Gold approach). (5) Point cloud derived from RL deconvolution and decomposition method (the RL approach).

After geo-referencing, the point clouds generated from full waveform LiDAR data using different methods with number of points (n) are shown in Figure II-5. As the figure shows, the point clouds derived from waveform LiDAR data (Figure II-6(3), (4), (5)) are denser than discrete-return LiDAR data (Figure II-6(1)), while they contained some noisy points in some vegetation parts. The point cloud generated by geo-referencing every time bin (Figure II-6(2)) could show a good shape of tree canopy with very high density points, but it was noisy with “false ground” and “false canopy height”. The point cloud generated from the direct decomposition (Figure II-6(3)) was comparable to that used the

deconvolution and decomposition (Figure II-6(4) & (5)) results. However, the number of points generated from these methods of the same study extent revealed that the Gold approach is larger than the RL approach and the direct decomposition approach. Overall, these results provide important insights into the selection of waveform processing methods, as some of these approaches may provide more information on the three-dimensional vegetation structure.

2.3.5 Accuracy assessment

Qualitative and quantitative accuracy assessments for the derived end products (DEM and CHM) were conducted in terms of visual and statistical comparisons. The results from SJER were used as an example to demonstrate the visual comparison results.

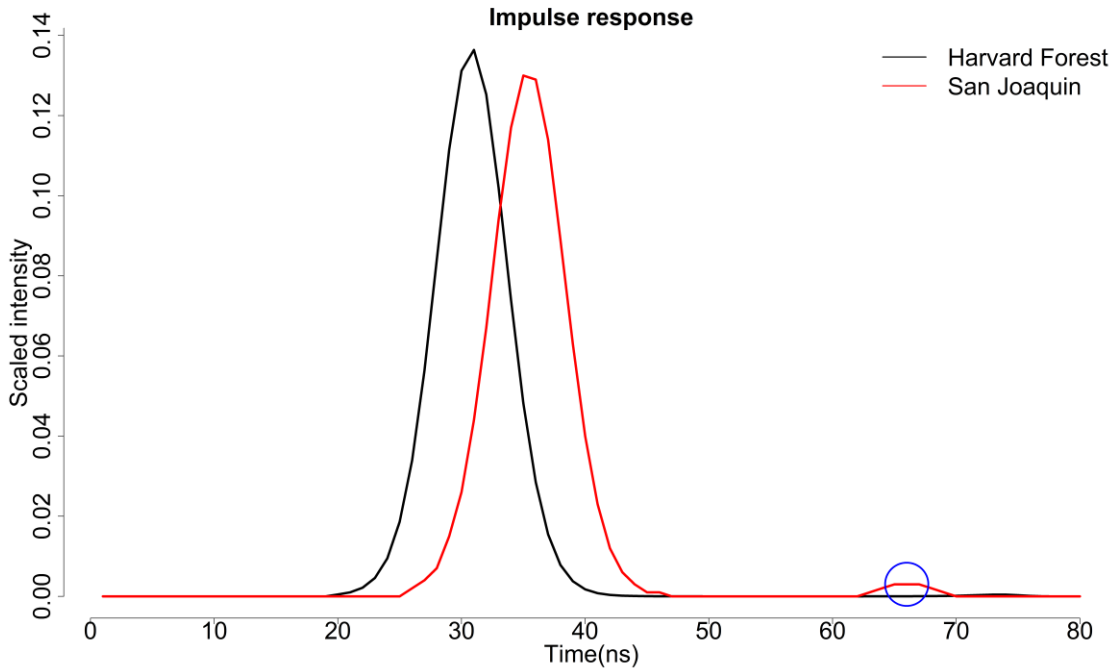


Figure II-7. The comparisons of the impulse response used for deconvolution provided by the NEON datasets

2.3.5.1 Digital Terrain Model

Figure II-8 shows that the DTMs generated from waveform LiDAR data with different approaches in the SJER study site almost have identical elevation distribution in comparison to the reference DTM. The range of the DTM derived from waveform LiDAR was also consistent with the reference DTM. However, a lower elevation range was observed for all approaches with the direct decomposition approach ranging from 381.9 to 424.2m, Gold approach from 381.3 to 423.2m and RL approach from 381.3 to 423.0m. It should be noted that there were some blank regions on the edges of DTMs, which may be due to no waveform or no information extracted from the sparse waveforms in those regions.

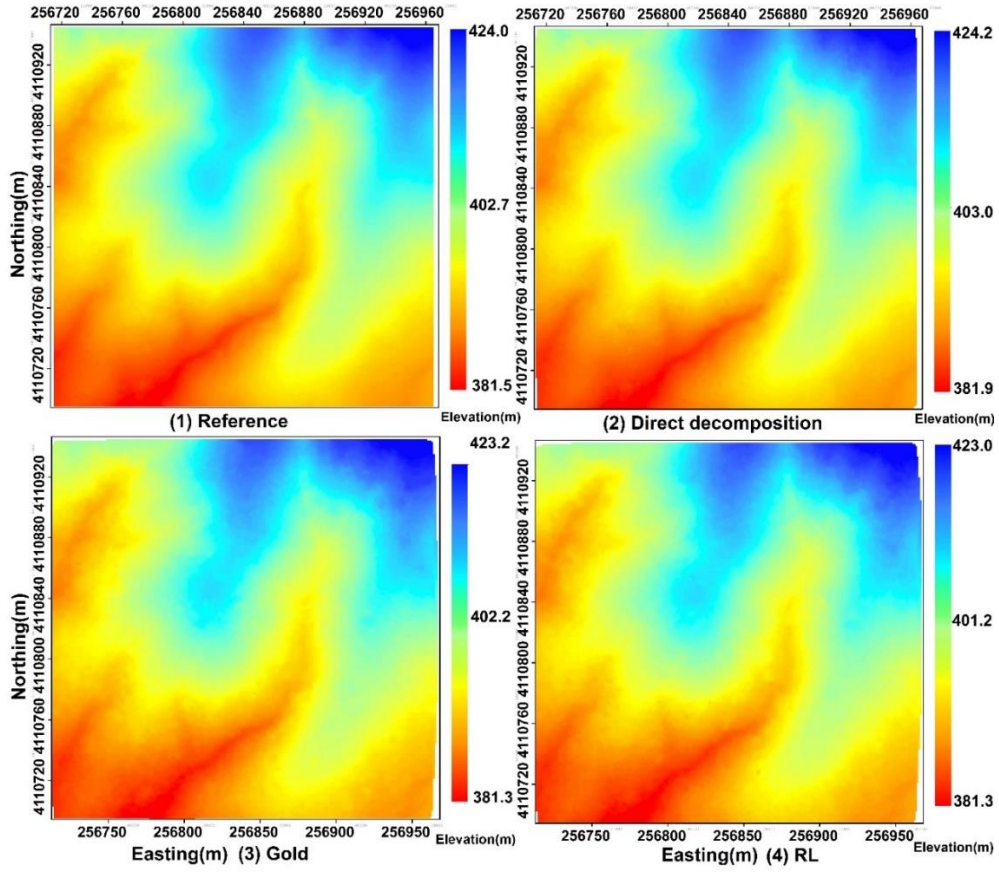


Figure II-8. Comparisons of (1) Reference DTM to waveform-based DTM generated from (2) the direct decomposition, (3) the Gold deconvolution and (4) the RL deconvolution approaches for the SJER site.

To further validate the performance of DTM results generated from different approaches, the statistical results for the DTMs are shown in Table II-4. It was noticeable that for all study sites, DTMs derived from waveform LiDAR were lower than the reference DTM from the perspective of range, and the direct decomposition approach outperformed the other two approaches generally. More precisely, the direct decomposition and gold approaches had similar performances for the HF study area in terms of root mean square error (RMSE) and mean difference (MD). Unlike the results of the HF site, the SJER and OSBS sites' direct decomposition approach outperformed the Gold and RL approaches. Especially for the SJER site, larger variances and RMSEs (Table II-4) were reported that probably resulted from the downward shift of detected peaks after we utilized the shifted estimated impulse response to deconvolve the waveforms. As shown in Figure II-7, the estimated impulse response of the SJER site (red) is shifted downward compared with the estimated impulse response of the HF site. Additionally, there is an extra peak at the end of the SJER site's impulse response which is an artifact, since the impulse response was reflected from flat ground that should have one peak. The deconvolution is sensitive to the input data such as impulse response and any shift of impulse response will propagate to the position accuracy of the detected peaks (Cawse-Nicholson et al., 2014). On the contrary, the impulse response for HF was more reasonable with only one significant peak. These may explain why the HF site's DTMs derived from the Gold approach and the RL approach were better than for the SJER site with the same approaches. However, through obtaining more accurate impulse response, the accuracy of the SJER DTM could be improved. Assuredly, the topography of the study areas, resolution of the DTM, and ground points' classification method may also contribute to larger variance of the results.

The percentage of the spatial elevation difference between DTMs derived from waveform LiDAR and reference data for the three study sites was also analyzed as shown in Table II-4. Almost all elevation differences (>94%) were within 0.5m when we used the direct decomposition approach for the SJER and OSBS study sites. The performance

of direct decomposition, Gold and RL approaches significantly varied on the SJER and OSBS sites when only the elevation difference within 0.5 (PW0.5) or 1m (PW1) was considered. However, the elevation difference within 2m (PW2) for this site was almost the same for these three approaches. As we explained above, the estimated impulse response may be the main reason for the lower performance of this site. For the deconvolution and decomposition method, the Gold approach had a better performance than the RL approach with a higher percentage of spatial difference located in the ranges from -0.5 to 0.5m and -1 to 1m for all study sites.

Table II-4. Summary statistics of DTMs (1m resolution) generated from the three approaches (Direct decomposition, Gold algorithm and RL algorithm) for the HF, SJER and OSBS sites

Approaches	Range (m)	SD (m)	MD (m)	RMSE (m)	PW0.5 (%)	PW1 (%)	PW2 (%)	PW3 (%)
HF								
Reference	313.5 - 318.0							
Discrete	313.9 – 318.1	0.18	-0.28	0.24				
Direct	313.7 - 317.9	0.23	-0.35	0.42	75.36	99.91	100	100
Gold	313.6 - 317.7	0.25	-0.35	0.31	68.82	97.68	100	100
RL	313.9 - 318.3	0.70	-0.21	0.70	60.43	80.45	93.40	100
SJER								
Reference	381.5 – 424.0							
Discrete	382.1 - 424.5	0.15	-0.12	0.15				
Direct	381.9 - 424.2	0.25	-0.24	0.26	93.87	99.70	100	100
Gold	381.3 - 423.2	0.30	-1.04	1.15	3.53	25.65	99.20	100
RL	381.3 - 423.0	0.31	-1.16	1.26	2.23	18.80	99.50	100
OSBS								
Reference	21.0 - 46.9							
Discrete	21.7 - 47.6	0.12	-0.08	0.15				
Direct	21.0 - 47.4	0.31	0.35	0.36	94.10	99.78	100	100
Gold	20.6 - 47.0	0.40	-0.14	0.42	60.56	94.56	99.88	100
RL	20.5 - 46.6	0.35	-0.54	0.62	49.68	92.47	99.94	100

Reference: Reference DTM; Discrete: Discrete-return LiDAR derived DTM; Direct: Direct decomposition approach; Gold: Gold approach; RL: RL approach. SD: standard deviation; MD: Mean elevation difference between DTM derived from waveform LiDAR and reference data; PW0.5: the percentage of difference within 0.5m (-0.5 - 0.5); PW1:

the percentage of difference within 1m (-1 - 1); PW2: the percentage of difference within 1m (-2 - 2); PW3: the percentage of difference beyond 2m (> 2.0 and < -2.0).

2.3.5.2 Canopy Height Model

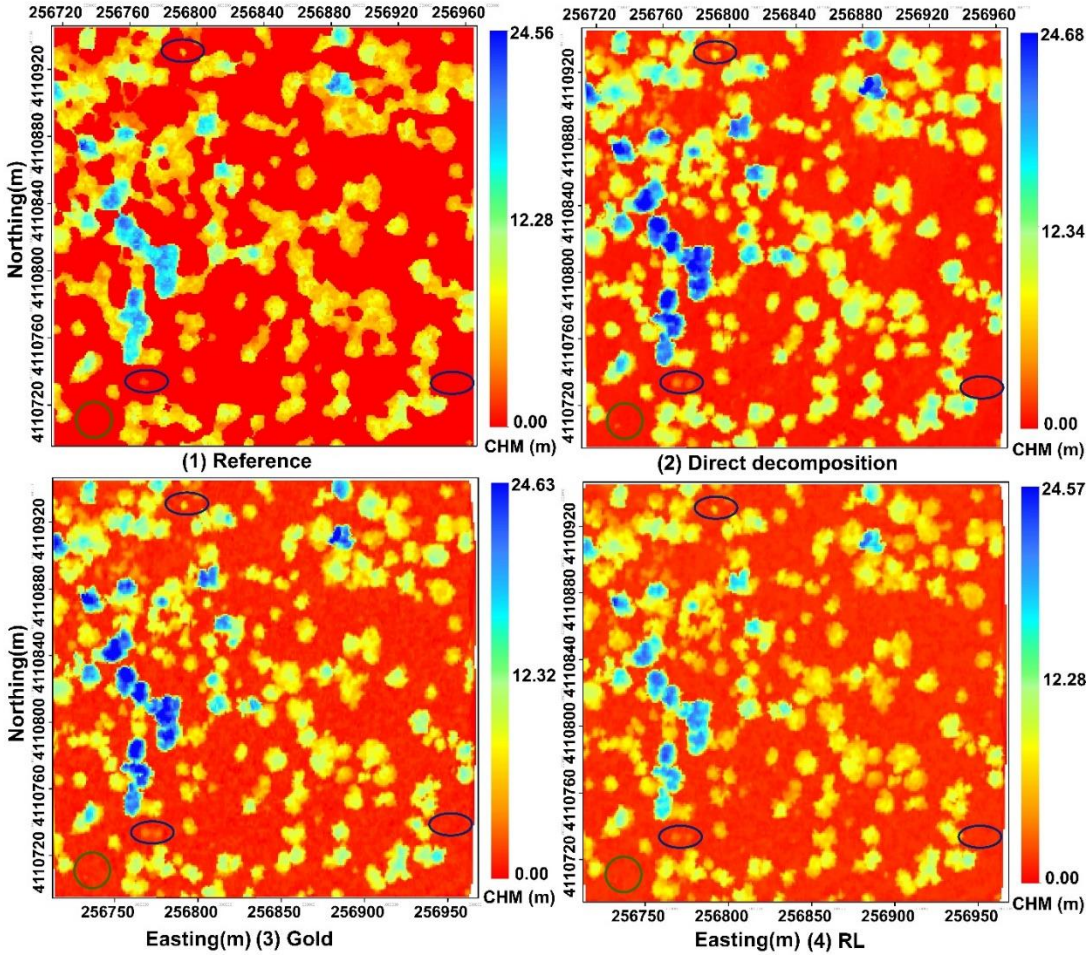


Figure II-9. Comparisons of (1) Reference CHM to waveform-based CHM generated from (2) the direct decomposition approach, (3) the Gold approach (4) and the RL approach for the SJER site.

Figure II-9 displays the CHMs generated from the direct decomposition, Gold and RL approaches for the SJER site. All approaches yielded similar results to our reference data in terms of the range and spatial distribution. It was worthwhile to note that the values in some regions of the reference data were zero (dark red), while the canopy height of the same region derived from waveform data was slightly higher as shown with ellipse shape.

As noted from the regions marked by the circle, the Gold and RL approaches' results give more detail information about low vegetation than the direct decomposition approach.

These differences may demonstrate that the waveform LiDAR data was more capable of detecting the low vegetation than the discrete-return LiDAR data, and the deconvolution and decomposition method had higher potential to detect low vegetation. The findings may provide insights into detecting understory layers below the forest canopy or grassland vegetation by using waveform LiDAR data. The range shift was not observed for CHMs using the Gold and RL approaches as noted for the DTMs, because the CHM was obtained by subtracting the DTM from the digital surface model (DSM) and the time shift was offset after subtraction.

In addition to the visual comparisons, the quantitative comparisons between waveform-based CHMs with different approaches and reference CHM data are shown in Table II-5. The Gold approach had the best performance with smallest standard deviation and RMSE for all sites. The result of the RL approach at the OSBS site had the smallest height difference but larger RMSE than direct decomposition approach, which may be primarily caused by the wider range of the height difference. However, it was still comparable to the result of direct decomposition approach. These observations may further indicate that the three approaches were reliable to extract the vegetation structure from the waveform LiDAR data and the Gold approach outperformed the other two approaches.

Analysis of percentage of the region's difference between maximum CHM derived from waveform LiDAR and the reference data further confirmed this conclusion (Table II-5). From the perspective of the percentage of height difference within 2m, about 96% of regions in HF site and 90% of region in other two sites, the performance of these three approaches were satisfactory. The majority of height differences larger than 2m occurred at the boundary of trees and ground. The reason behind this was that the small change of peak (t) not only lead to the slight change of height (z), but also resulted in the synchronized change of XY locations, as implied in the Eq. (II-8) and (II-9). When

compared with the reference data, the boundary of trees and ground most likely yielded larger height differences due to this kind of XY location shift.

However, the performance of the three approaches varied at different sites in terms of the height difference within 0.5 and 1m. The Gold approach worked best at the HF and SJER sites, while it did not work as well as the direct decomposition in the OSBS site.

Globally, the CHMs generated from the waveform LiDAR data using the three approaches were satisfactory compared to the reference data for these three study sites. The Gold approach worked slightly better with smaller standard deviation and RMSE for all sites. However, the direct decomposition approach outperformed the Gold approach with higher percentage of area located in the given spatial difference range, especially when the spatial difference range was within 0.5 and 1m at OSBS site. This may be because the HF study area was flatter than SJER study area and potentially proved that Gold approach could work well in regions with different topography.

Table II-5. Summary of comparison of CHMs (resolution 1m) generated from the three approaches (Direct decomposition, Gold approach and RL approach) for the HF, SJER and OSBS sites.

Approaches	Range (m)	SD (m)	MD (m)	RMSE (m)	PW0.5 (%)	PW1 (%)	PW2 (%)	PW3 (%)
HF								
Reference	9.10 - 23.06							
Discrete	9.19 – 23.06	0.42	0.25	0.51				
Direct	9.19 - 23.40	0.70	0.65	0.95	50.61	80.95	96.39	100
Gold	10.83 - 22.91	0.45	0.38	0.72	65.53	88.28	98.37	100
RL	9.95 - 23.67	0.98	0.28	1.02	42.45	69.31	95.28	100
SJER								
Reference	0.00 - 24.56							
Discrete	0.00 – 24.49	0.32	0.15	0.35				
Direct	0.00 - 24.68	1.40	0.50	1.65	60.87	75.08	88.03	100
Gold	0.00 - 25.12	1.28	-0.12	1.06	72.65	81.82	89.62	100
RL	0.00 - 24.57	1.39	0.35	1.51	34.56	80.34	88.89	100
OSBS								
Reference	0.00 - 27.35							
Discrete	0.00 – 27.77	0.28	0.11	0.31				
Direct	0.00 - 27.58	1.67	0.75	1.40	68.25	83.48	92.56	100
Gold	0.00 - 27.81	1.54	0.30	1.62	48.59	79.68	88.25	100
RL	0.00 - 28.50	1.93	0.13	1.87	38.25	68.37	86.87	100

Reference: Reference CHM; Discrete: Discrete-return LiDAR derived CHM; Direct: Direct decomposition approach; Gold: Gold approach; RL: RL approach. SD: standard deviation; MD: Mean height difference between CHM derived from waveform LiDAR and reference data; PW0.5: the percentage of difference within 0.5m (-0.5 - 0.5); PW1: the percentage of difference within 1m (-1.0 – 1.0); PW2: the percentage of difference within 1m (-2.0 – 2.0); PW3: the percentage of difference beyond 2m (> 2.0 and < -2.0).

2.3.6 Parameter uncertainty

2.3.6.1 Digital Terrain Model

Figure II 10 shows that the DTMs' spatial uncertainty of the SJER site for the three approaches mostly ranges from -1m to 1m, and the DTMs derived from the Lower datasets are smoother than the Upper datasets. More specially, the direct decomposition approach had the smallest variation of spatial uncertainty with more than 86% of the prediction error within 0.50m for the Lower dataset, and 89% of the prediction error was

located in the range from 0.00 to 0.49m for the Upper dataset. The RL approach yielded similar result as the Gold approach when using the Lower dataset and most of the uncertainty ranged from 0.51 to 1.00m. The spatial uncertainty using the Upper dataset with the RL approach had less variation compared to the Gold approach that was consistent with the standard deviation in Table II 6. It was worth noting that the spatial distribution of uncertainty derived from the Lower dataset was not consistent with that of the Upper dataset. This may be mainly attributed to the fact that the NEON data could provide x, y, z change per nanosecond for each waveform which can result in the synchronized change of points' x, y, z when the peak location with 95% uncertainty level was considered. In addition, the interpolation and the smoothing process may also lead to this kind of inconsistency.

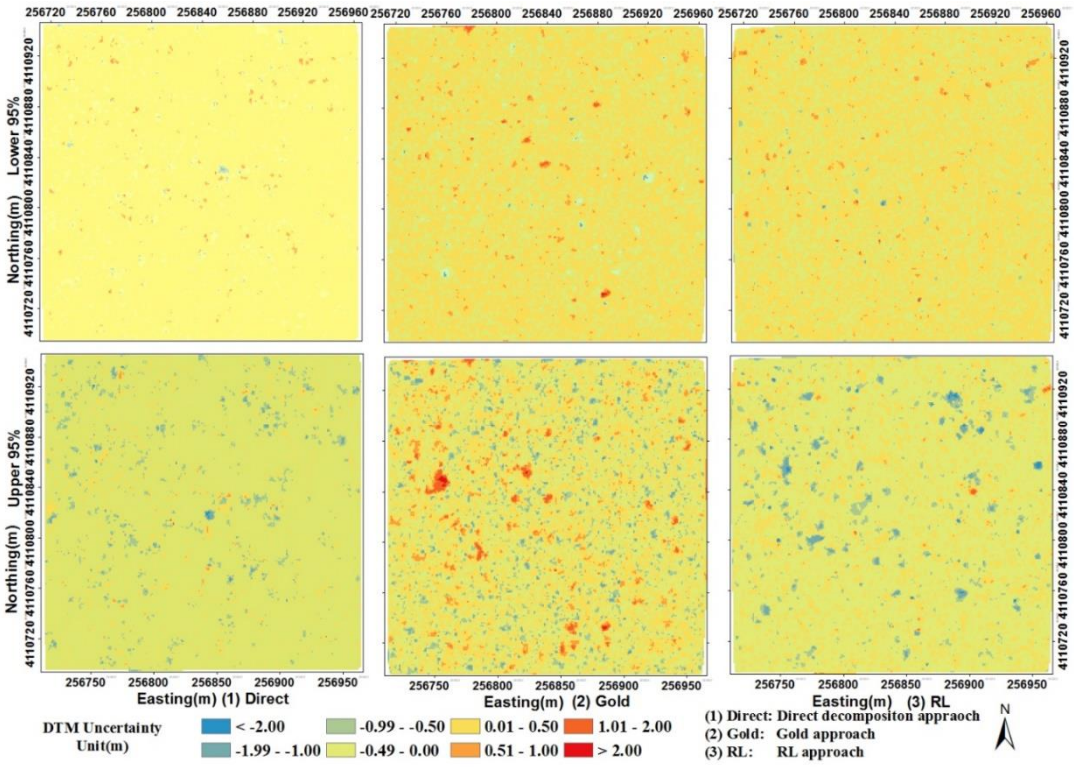


Figure II-10. The spatial uncertainty of the DTM caused by the parameter uncertainty in the SJER site using the direct decomposition approach (left), Gold approach (middle) and RL (right) approach, respectively. The above was the result from the corresponding Lower dataset and the bottom was the result from the corresponding Upper dataset.

Table II-6. Global statistics summarizing validation errors caused by parameter uncertainty for DTMs.

Approaches	Dataset	Range (m)	SD (m)	MU (m)	MinU (m)	MaxU (m)	RMSE (m)
HF							
Direct	low	313.89 - 318.15	0.15	0.23	-0.06	0.76	0.32
	up	313.69 - 317.50	0.18	-0.24	-0.79	0.60	0.32
Gold	low	313.40 - 318.00	0.07	0.04	-0.21	0.28	0.08
	up	313.31 - 317.97	0.09	-0.03	-0.49	0.21	0.09
RL	low	313.74 - 317.48	0.39	-0.15	-2.12	0.88	0.42
	up	313.35 - 316.2	0.52	-0.78	-3.05	0.22	0.94
SJER							
Direct	low	382.09 - 424.05	0.12	0.05	-1.24	1.53	0.21
	up	381.69 - 424.11	0.18	-0.10	-2.51	1.29	0.13
Gold	low	381.39 - 423.11	0.17	0.03	-1.61	1.43	0.17
	up	380.78 - 423.15	0.36	-0.04	-2.59	2.53	0.36
RL	low	381.39 - 423.11	0.14	0.03	-2.21	2.59	0.14
	up	381.35 - 422.92	0.26	-0.10	-3.77	1.89	0.28
OSBS							
Direct	low	21.88 - 47.10	0.16	0.08	-1.57	2.48	0.18
	up	20.96 - 47.02	0.26	-0.13	-2.27	1.42	0.29
Gold	low	20.94 - 45.46	0.11	0.01	-1.59	1.38	0.11
	up	20.99 - 45.61	0.18	0.02	-1.88	1.51	0.18
RL	low	20.48 - 46.67	0.10	0.01	-1.84	1.78	0.10
	up	20.99 - 46.63	0.11	0.00	-2.54	1.74	0.11

Direct: Direct decomposition approach; Gold: Gold approach; RL: RL approach. SD: standard deviation; MU: Mean uncertainty caused by parameters between DTM derived from uncertainty dataset and DTM derived from peak location dataset; MinU: Minimum change caused by parameter uncertainty; MaxU: Maximum change caused by parameter uncertainty; low: Lower dataset; up: Upper dataset.

Global statistics for the DTMs' parameter uncertainty is presented in Table II-6. For all study sites, all approaches' absolute mean spatial uncertainties were below 0.25m, except for the Upper dataset using the RL approach, indicating that biases caused by parameter uncertainty were relative low. The RMSE ranged from 0.08 to 0.36m, which was almost consistent with the standard deviation. From the statistics in Table II-6, there is no obvious difference among these three approaches in terms of the RMSE and SD.

However, the Gold approach outperformed the other two approaches when the MU was taken into account.

The RL approach had larger minimum error and maximum error in the HF site that resulted in larger RMSE and mean absolute errors at the HF site. The RL approach's performance enhanced substantially when it came to the SJER and OSBS sites with smaller RMSE. For instance, the RL approach's RMSE for the Lower and Upper dataset were 0.14 and 0.28m at the SJER site, when compared to 0.42 and 0.94m at the HF site. The only one flight line may contribute to higher SD and RMSE at the HF site since there was no overlap in the study region with less dense raw waveform data. It was surprising to find that the Upper datasets for all approaches have larger range and RMSE than the corresponding Lower datasets. It may be attributed to the fact that most of the Upper dataset's points were lower than reference DTM, which had a more weight on the effect of the DTM generation than using the Lower dataset.

The direct decomposition method worked well and consistently in these three sites, but the deconvolution and decomposition method (either the Gold or RL approach) was more likely to generate smaller RMSE than the direct decomposition approach.

To further identify areas where DTM surfaces of low quality with high uncertainty, and compare the performances of approaches under different conditions, the slope and the vegetation height were taken into account as important predictors for the categories of uncertainty. Here, the HF and SJER sites were selected as examples to demonstrate the effect of the slope and vegetation height.

The ANOVA analysis showed that vegetation height and slope had a significant effect on the uncertainty levels of DTMs for the three approaches at the HF site, with all p-values smaller than 0.05 (Figure II-11 and Figure II-12). The green line (median) increased with uncertainty levels for the three approaches, which demonstrated that larger slope and higher vegetation height were more likely to cause high uncertainty of DTM. For the SJER site, the uncertainty levels vs. slope showed that all approaches' p-values were zero except the Gold approach's p-value was 0.917. This indicated that the slope had no effect on the uncertainty levels and the Gold approach may be robust when dealing

with complex topography. The average slope and slope distribution of the three uncertainty levels for the other two approaches were similar to each other, even the ANOVA analysis showed that slope was a significant factor for determining the uncertainty levels. This may potentially imply that slope could influence the uncertainty levels, but its cause-effect relationship is not so strong. Unlike slope, the vegetation height's effect on the uncertainty levels was more significant with regard to mean, median and interquartile range (IQR) as shown in

Figure II-13 and Figure II-14. Additionally, the higher the vegetation height, the more likely for this area to have higher uncertainty level. It was worthy to note that the median was not consistent with the mean for the corresponding uncertainty level. Most of the low and medium uncertainty occurred on the ground with mean and median being zero for low uncertainty levels (Figure II-14).

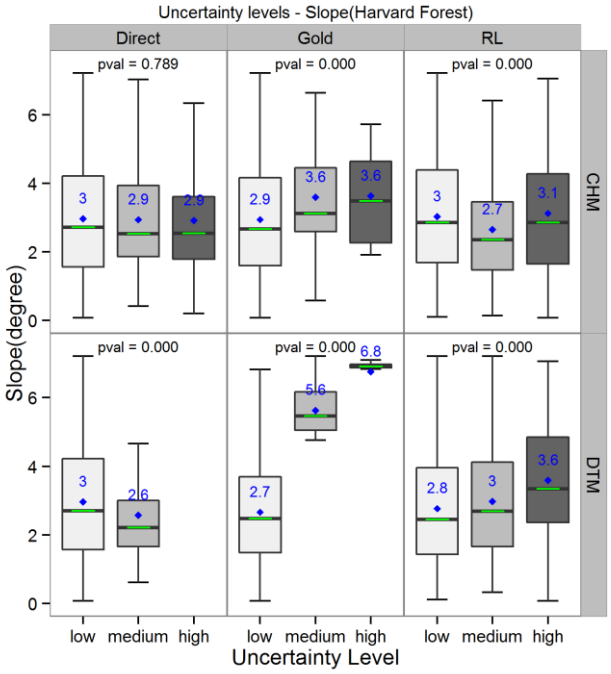


Figure II-11. Box-plot of DTMs and CHMs' Uncertainty levels vs. Slope for the HF site

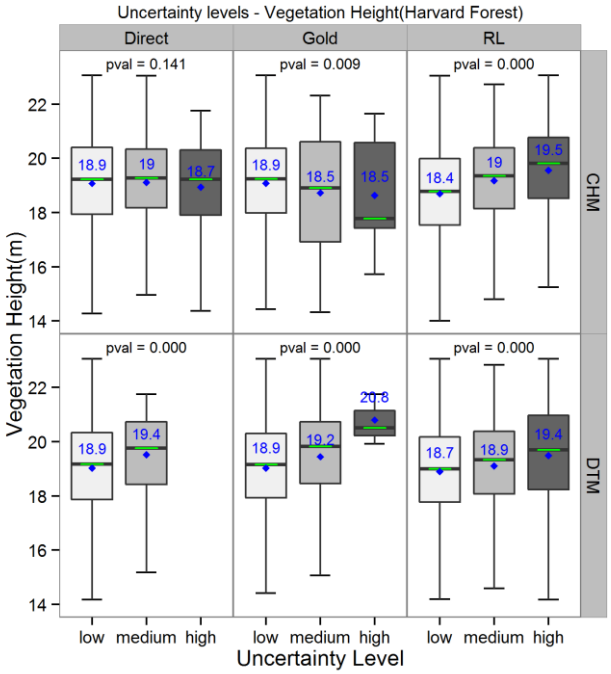


Figure II-12. Box-plot of DTMs and CHMs' Uncertainty levels vs. Vegetation Height for the HF site

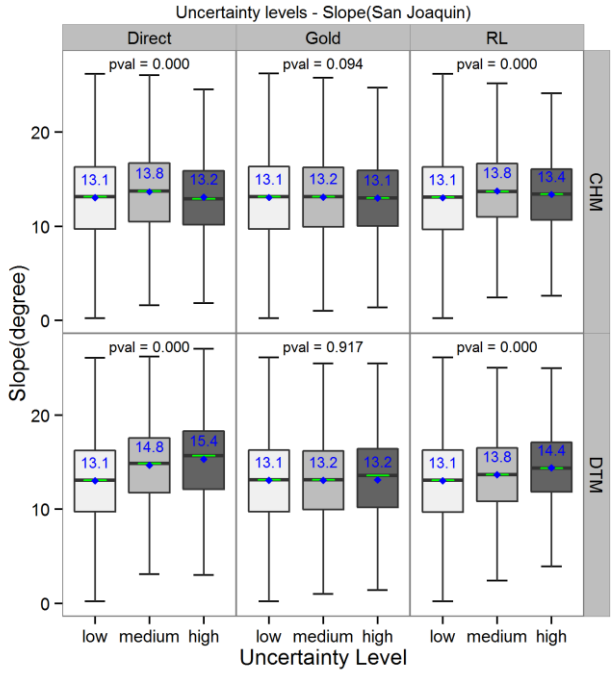


Figure II-13. Box-plot of DTMs and CHMs' Uncertainty levels vs. Slope for the SJER site

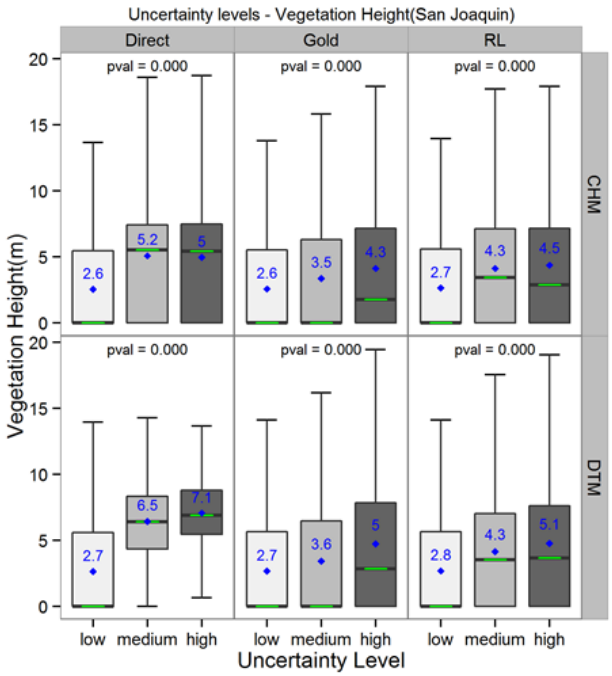


Figure II-14. Box-plot of DTMs and CHMs' Uncertainty levels vs Vegetation Height for the SJER site

*Green lines indicated median of dataset, the height of the box portion was given by the IQR of the dataset and the ends of the whiskers meant 1.5 IQR of lower quantile and 1.5 IQR of upper quantile. Blue points were the mean of the corresponding variable.

In summary, the analysis identified that high prediction uncertainty of DTM was more likely to occur at larger slope and higher vegetation for all approaches in flat topography with dense vegetation. The vegetation height's effect on the DTM's uncertainty levels was more significant than slope when it came to the complex topography. The ground was more prone to lower spatial uncertainty level which may result from that waveforms in the ground region were simpler than in the slope and vegetation regions.

2.3.6.2 Canopy Height Model

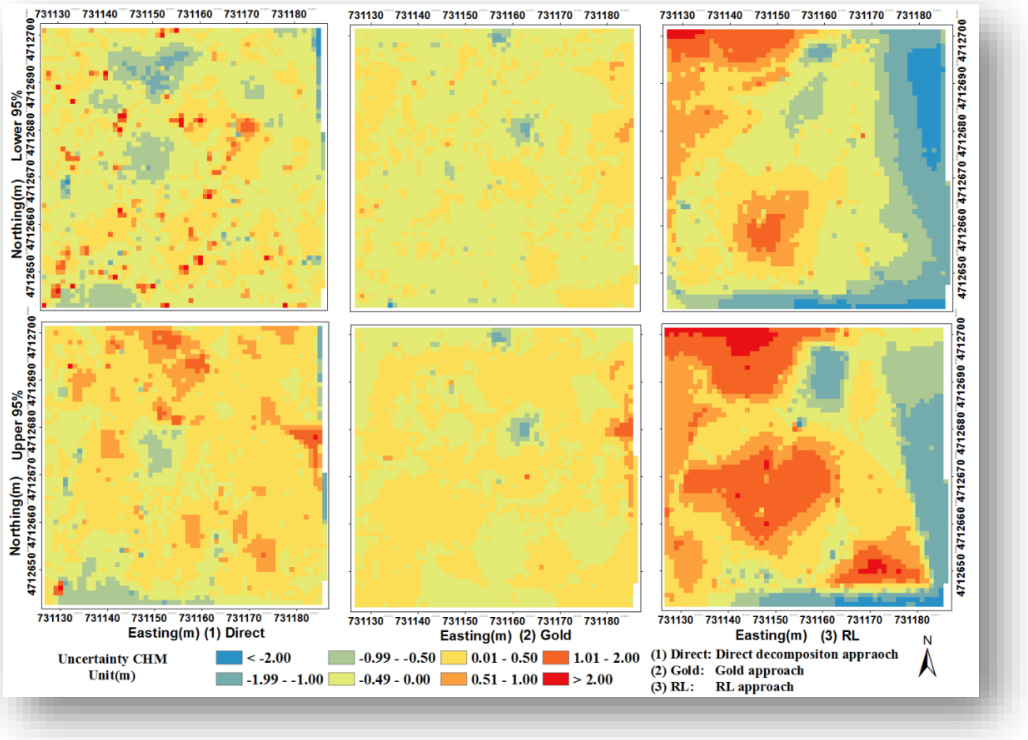


Figure II-15. The spatial uncertainty of CHM caused by the parameter uncertainty in HF region using the direct decomposition approach (left), Gold approach (middle) and RL (right) approach, respectively. The above was the result from the Lower dataset and the bottom was the result from the Upper dataset.

Visual comparisons of the maximum CHM of the HF site for the three approaches are shown in Figure II-15. With regard to different datasets for the three approaches, the spatial distribution of uncertainty was similar, but not identical. The Gold approach had the smallest variance with similar spatial distribution using the two different datasets. The RL approach had the largest variance, which was coincident with the global statistic of Table II-7. Somewhat surprisingly, the largest uncertainty level (dark red or dark blue) for the RL approach was more likely to occur at the edges of the HF site rather than at the boundary of trees and ground (Figure II-15). After a closer examination, we found that

the large uncertainty level was located in dense vegetation areas, which may further imply that the Gold approach was less suitable for dense vegetation regions.

The quantitative assessment of the CHMs' parameter uncertainty (Table II-7) yielded similar results to those of DTMs. The absolute mean spatial difference for all approaches ranged from 0.00 to 0.27m, which was smaller than the DTMs' result derived from the corresponding datasets. It was not surprising to see that the minimum uncertainty and maximum uncertainty for each dataset of CHMs were larger than corresponding DTMs' results, since DTMs were much flatter than CHMs and relatively lower uncertainty was expected. In addition, the CHM was generated with additional steps compared with DTM that may bring more error into the CHM products. Most of these large uncertainties occurred at the boundary of trees and ground, and a small shift of XY location from parameter uncertainty would result in a large difference of canopy height.

The Gold approach had the smallest RMSE for CHMs in the HF and OSBS sites, but it had the largest RMSE for the SJER site. By contrast, the RL approach yielded the opposite results with the smallest RMSE at the SJER site, and the largest RMSE and mean error for the HF site. These trends were consistent with the DTMs' results. It confirmed the previous conclusion that the Gold approach may be more suitable in flat terrain areas and RL approach tended to perform better in complex topography conditions. The direct decomposition method performed well in both study areas, but not as good as the deconvolution and decomposition method (either the Gold or RL approach) that may further indicate the advantages of the deconvolution. Assuredly, the relationship between vegetation, and topographic conditions and deconvolution results is complex, the simulated waveform data with different topographic and vegetation conditions will be the ideal datasets to further test the approaches and provide insights into selecting approaches under different conditions of topography and vegetation.

Table II-7. Global statistics summarizing validation errors caused by parameter uncertainty for CHMs (unit: m).

Approaches	Dataset	Range	SD	MU	MinU	MaxU	RMSE
HF							
Direct	low	11.44 - 24.13	0.46	-0.07	-2.85	4.82	0.47
	up	11.76 - 23.38	0.38	0.13	-1.74	2.33	0.40
Gold	low	10.71 - 22.75	0.18	-0.04	-2.34	0.92	0.19
	up	10.84 - 23.16	0.19	0.01	-1.64	1.71	0.20
RL	low	9.90 - 22.86	0.88	-0.27	-2.73	2.66	0.92
	up	10.19 - 23.71	0.95	0.25	-3.77	2.56	0.98
SJER							
Direct	low	0.00 - 24.85	0.43	0.06	-6.32	11.00	0.44
	up	0.00 - 24.70	0.40	0.01	-6.42	8.87	0.39
Gold	low	0.00 - 24.70	0.51	0.15	-7.06	10.22	0.53
	up	0.00 - 24.24	0.49	0.04	-8.86	9.26	0.49
RL	low	0.00 - 24.78	0.33	0.00	-9.44	8.70	0.33
	up	0.00 - 24.63	0.35	0.09	-6.19	5.50	0.37
OSBS							
Direct	low	0.00 – 27.36	0.61	0.00	-17.2	17.05	0.61
	up	0.00 - 27.73	0.69	0.07	-14.56	17.70	0.69
Gold	low	0.00 - 28.61	0.01	0.00	0.00	2.71	0.01
	up	0.00 - 28.59	0.58	-0.03	-13.71	16.22	0.57
RL	low	0.00 - 28.37	0.48	0.03	-18.18	19.20	0.49
	up	0.00 - 28.12	0.46	-0.01	-17.20	17.12	0.46

Direct: Direct decomposition approach; Gold: Gold approach; RL: RL approach. SD: standard deviation; MU: Mean uncertainty caused by parameters between DTM derived from uncertainty dataset and DTM derived from peak location dataset; MinU: Minimum change caused by parameter uncertainty; MaxU: Maximum change caused by parameter uncertainty; low: Lower dataset; up: Upper dataset.

The ANOVA analysis of CHMs at the HF site (Figure II-11) demonstrated that the median, mean and IQR of slope were similar for different CHMs' uncertainty levels using the direct decomposition approach. This result indicated that the slope had no effect on the CHMs' uncertainty levels for the direct decomposition approach in terms of the statistical perspective. However, this factor's effect on the uncertainty levels was significant ($p < 0.050$) for the Gold and RL approaches. Likewise, the vegetation height also had an impact on the uncertainty levels using the Gold and RL approaches, but not

for the direct decomposition approach at the HF site as shown in Figure 12. Generally, higher uncertainty level was more likely to occur at the higher slope and vegetation height for the Gold and RL approaches.

For the SJER site, the analysis showed that the uncertainty levels of CHMs were influenced by the vegetation height for all approaches ($p = 0.000$), and there was a large difference between mean and median of vegetation height for different uncertainty levels (Figure II-14Figure II-14. Box-plot of DTMs and CHMs' Uncertainty levels vs Vegetation Height for the SJER site). Most of the uncertainty was more likely to occur at the lower vegetation and ground, especially for the Gold approach. The analysis of the slope (Figure II-13) showed that the Gold approach's uncertainty levels were robust to slope changes ($p = 0.094$). The p-values of the other two approaches' were zero, which indicated that the uncertainty of CHM was likely to be affected by slope changes. However, these two approaches' median, mean and IQR were similar for different uncertainty levels which may imply that the slope did not play as an important role as vegetation height in determining uncertainty levels.

In summary, the analysis identified that higher uncertainty of CHM was prone to occur at higher vegetation for all approaches in the complex topography with less dense vegetation areas. For the flat topography, there was no obvious pattern for the Gold and RL approaches for different uncertainty levels. The direct decomposition approach outperformed other two approaches and was robust to the change of slope and vegetation height under such topography condition.

2.4 Conclusions

This study proposes a new waveform LiDAR deconvolution algorithm called the Gold algorithm as a preprocessing step, and comprehensively compares different methods of processing the waveform LiDAR data at three different ecological sites from both visual and quantitative perspectives.

Our work has demonstrated the advantage of the deconvolution and decomposition method with more echoes of waveforms detected and less false echoes generated, especially when the Gold approach is used. Furthermore, the accuracy assessment of the

end products (DTMs and CHMs) shows that the three approaches can generate satisfactory results, while the best performances vary when different criteria are used: the Gold approach has better performance with smaller RMSE, and the direct decomposition approach outperforms others in terms of the percentage of spatial difference within 0.5 and 1m. According to the parameter uncertainty of end products, the factors like the vegetation height and slope both have an effect on the robustness of approaches, while the slope becomes a less significant factor when it comes to the spatial uncertainty of CHMs. Specifically, the Gold approach tends to have better performance in the dense vegetation region and the RL approach works better in the sparse vegetation region. Therefore, the important contributions of this study lie in successfully introducing a novel deconvolution algorithm, the Gold algorithm, for waveform LiDAR processing, and providing a comprehensive comparison and a quantifiable basis selection of different waveform LiDAR processing methods for different topography and vegetation conditions. Potential future studies could use the proposed method to process waveform LiDAR data and extract semantical information, such as individual tree crown mapping, understory tree detection, and to estimate forest structure and biophysical parameters. In addition, future investigations could benefit from expanding the availability of new waveform LiDAR datasets to cover varied vegetation conditions in multiple ecosystem types and complex topography, urban areas, rangelands and grasslands.

CHAPTER III

BAYESIAN DECOMPOSITION OF FULL WAVEFORM LIDAR DATA WITH UNCERTAINTY ANALYSIS*

A thorough understanding of full waveform (FW) LiDAR data processing and associated uncertainty is critical to vegetation applications such as retrieving forest structure variables and estimating forest biomass. This paper applies the Bayesian non-linear modelling concept to process small-footprint FW LiDAR data (the Bayesian decomposition) collected at a study site of the National Ecological Observatory Network (NEON) to investigate its potential for waveform decomposition and uncertainty estimation. Specifically, several possible models suitable for fitting waveforms were assessed within the Bayesian framework, and the Gaussian model was selected to perform the Bayesian decomposition. Subsequently, we conducted performance evaluation and uncertainty analysis at the parameter, derived point cloud and surface model levels. Results of the model reasonableness show that the Gaussian model is superior to alternative models with respect to uncertainty, physical meaning and processing efficiency. After converting waveforms to discrete points, the model comparisons demonstrate that the Bayesian decomposition can be utilized for FW LiDAR data processing, and its results are comparable to the direct decomposition (DD), Gold and RL (Richardson–Lucy) approaches in terms of the root mean squared error (RMSE < 0.93 m) of the point distances between the waveform-based point cloud and the reference point cloud. Additionally, more points can be extracted from FW LiDAR data with these methods than discrete-return LiDAR data, especially at the mid-story of vegetation based on the results of height bins, percentile heights and canopy LiDAR density at the individual tree level. Moreover, uncertainty estimates from the Bayesian method enhance the credibility of decomposition results in a probabilistic sense to capture the true error of estimates and trace the uncertainty propagation along the processing steps. For example,

*Reprinted with permission from "Bayesian decomposition of full waveform LiDAR data with uncertainty analysis." by Zhou, Tan, and Sorin C. Popescu. Remote Sensing of Environment 200 (2017): 43-62. Copyright 2017 by Elsevier.

results of the surface model yield larger RMSE values (1.38 m vs. 0.65 m) with a wider credible interval than quantile point clouds with a more compact distribution. In contrast to commonly used deterministic approaches, the Bayesian decomposition method can produce an ensemble of reasonable parameter estimates with probability through Markov Chain Monte Carlo (MCMC) sampling from the posterior distribution of model parameters. These parameter estimates and corresponding derived products can be queried to provide meaningful interpretation of results and associated uncertainty. Both the flat priors and empirical priors can achieve good performance of the decomposition while the empirical priors tend to significantly speed up the model convergence. The Bayesian approach also renders an important insight into the uncertainty of the model performance evaluation using field data by generating reasonable prediction intervals to reduce inherent errors of field measurements.

Keywords: Waveform LiDAR, Bayesian inference, Decomposition, Uncertainty analysis, Tree canopy height, Model reasonableness

3.1 Introduction

Light Detection and Ranging (LiDAR) has been adopted as a valuable survey tool reducing the need for field measurement to accurately characterize vegetation structure for the past decades (Hancock et al., 2015; Popescu et al., 2003; Wulder et al., 2012; Zhao et al., 2009). Especially, the advent of the full waveform (FW) LiDAR system, which is capable of recording entire reflected energy along the pulse line, has enabled this advantage to become more conspicuous (Cawse-Nicholson et al., 2014; Wagner et al., 2006; Wulder et al., 2012). The FW LiDAR data primarily consists of two parts: a pulse part that keeps geo-reference locations derived from range measurement between the laser sensor and the reference location, and a wave part which fully stores digitized return energy starting from the reference location till the end of digitized samples. Within a forest environment, FW LiDAR energy could penetrate dense canopies through small gaps in the canopy and achieve a full time-versus-intensity profile. Consequently, more detailed vertical information of vegetation structure can be revealed through these data.

Generally, FW LiDAR data can be classified as large- (~50 m or larger), medium- (~10 - 30 m) or small-footprint waveform (< 1 m) based on their transmitted laser size (Wang and Weng, 2013). Some of the first FW sensors known as large footprint profilers included SLICER (Scanning LiDAR Imager of Canopies by Echo Recovery, 10 m footprint), LVIS (Laser Vegetation Imaging Sensor, 25 m footprint) and GLAS (the Geoscience Laser Altimeter System, 70 m footprint). All of them have been successfully applied to estimate various forest parameters and vegetation studies worldwide (Blair et al., 1999; Drake et al., 2002; Harding and Carabajal, 2005).

Recent advances of commercial LiDAR systems have promoted the availability of small-footprint FW LiDAR data from remote sensing industry providers. However, extensive applications of such systems to characterize forest structure and biomass are limited (Wulder et al., 2012). There are three main reasons behind this: (1) there is no standard format of FW LiDAR data, (2) large data volume required for storing the information, which leads to difficulties of data distribution and processing, (3) high cost of data acquisition, with added cost compared to discrete-return (DR) LiDAR data that hinders their adoption for many potential applications (Pirotti, 2011). In addition, while there are many software packages and applications available for processing DR LiDAR data, only few software developments are currently available for processing FW LiDAR. Therefore, the development of robust and dedicated methods and non-proprietary software for processing small-footprint FW LiDAR data are urgently needed.

Existing methods for FW LiDAR processing can be mainly categorized into two types: the decomposition method and the combined deconvolution and decomposition method. The most commonly used approach for the decomposition method is the direct decomposition (DD) which models the waveform with a mixture of Gaussian functions. Typical approaches such as Non-linear least-squares (NLS) (Hofton et al., 2000) or maximum likelihood estimation using the Expectation-Maximization (EM) algorithm (Persson et al., 2005) have been developed for fitting the waveform to extract 3D points and related parameters. However, these approaches are sensitive to the initialization of unknown parameters. Another popular method for recovering the true cross section of

objects along the pulse line is the combined deconvolution and decomposition method. Multiple deconvolution algorithms, such as the Gold, Richardson-Lucy (RL), Non-negative least squares (NNLS), and Wiener Filter (WF) (Cawse-Nicholson et al., 2014; McGlinchy et al., 2014; Neuenschwander, 2008; Roncat et al., 2010; Rowe, 2013; Wu et al., 2011; Zhou et al., 2017) have been successfully introduced for reconstructing the differential backscatter cross section. One practical issue of these methods is that they are pertinent to the choice of proper parameter combinations for the deconvolution, which typically requires parameter optimization before data processing (Zhou et al., 2017). Although these methods have been proven to be able to generate sufficient fitting models, we cannot characterize the uncertainty with these models. They are calculated based on the deterministic models, which only seek a point value for the parameter of interest (Edwards et al., 2003). Without uncertainty analysis, the models are less informative or even useless when they are applied to the real-world problems (Zhao et al., 2011).

In the domain of LiDAR applications, the observations or data are inherently subject to various errors such as system setting, system calibration, and range measurement errors (Griewank and Walther, 2008). Additionally, the LiDAR vendors often do not clearly state what errors are considered when the data are provided. Thus, uncertainty of "truth" is ubiquitous and inherently present in the realities of LiDAR data modeling. Some studies associated with uncertainty mainly focus on the DR LiDAR data applications (Chauve et al., 2009; Chen et al., 2015; Frazer et al., 2011), while few published studies have explored FW LiDAR data's uncertainty for vegetation characterization. Furthermore, the models used here are based on the non-linear functions that generally suffer from problem of non-uniqueness (Sen and Stoffa, 1996), which can generate different parameter combinations given the same observational data and model, or several models can fit observational data at the expense of violating the physical meaning and theoretical assumption of the "real" model. These problems are more evident for the sophisticated models with multiple peak components in the waveform decomposition. Thus, estimating model uncertainty is imperative for an in-depth understanding of information derived from data and the estimation accuracy. This kind of

uncertainty analysis has been inadequately addressed in many preceding studies due to the great diversity of remote sensing-based information retrieval procedures (Zhao et al., 2011) and the absence of efficient and universal methods to capture the uncertainty of data modeling.

One strategy to overcome these problems is to adopt a Bayesian paradigm of statistical inference by considering the model parameters to be realizations of random variables. Within a Bayesian framework, we combined prior information about unknown parameters with observed data using the Bayes' rule. Using a Markov Chain Monte Carlo (MCMC) (Gelfand and Smith, 1990) algorithm, we were able to sample from the posterior distribution of the unknown model parameters of interest. Through the posterior distribution, uncertainty bounds on the resulting model parameters and model reasonableness can be measured. This approach generally takes more time to reach a solution, but the non-uniqueness of the model parameter estimates can be avoided by describing these parameters in terms of probability density functions (PDFs) in the model space (Hong and Sen, 2009; Sen and Stoffa, 1996). Additionally, computational advances and the introduction of the more efficient Hamiltonian Monte Carlo (HMC) algorithm (Neal, 2011) have contributed enormously to the growing interest in applying Bayesian approaches to remote sensing data processing.

Recently, Bayesian analytical approaches have been applied to diverse domains related to waveform data as an alternative to traditional deterministic techniques. For example, Qin et al. (2016) analyzed ground-penetrating radar (GPR) data to detect the defect of underground structure using a Bayesian inversion method. Roonizi et al. (2016) elaborated how the electrocardiogram (ECG) waveform separation was conducted in a Bayesian framework to evaluate cardiac health status. In the geophysical field, the Bayesian approach was used for marine seismic waveform data to characterize subsurface reflectivity (Ray et al., 2013). For LiDAR applications, Bayesian methods are mainly used in a spatial modeling context to predict and map forest variables (Finley et al., 2013) or image construction (Hernandez-Marin et al., 2008). However, employing Bayesian

approaches to decompose FW LiDAR data for vegetation studies is rarely reported in the current literature.

Therefore, the overall goal of this paper is to explore a Bayesian statistical method with the HMC algorithm to process small-footprint FW LiDAR data and quantify the uncertainty from data and models. More specifically, we attempt to (1) evaluate the reasonableness of models suitable for FW LiDAR data within a Bayesian framework; (2) develop a robust and dedicated Bayesian decomposition method to process FW LiDAR data for vegetation, and implement thorough comparisons with other FW LiDAR data processing methods; and (3) obtain reliable estimates of error and uncertainty in different steps (the parameter estimates, point cloud and surface models) using the Bayesian decomposition method. The motivation for the first objective is to check the validity of previous studies' underlying assumptions that the Gaussian model is sufficient for FW LiDAR data decomposition (Mallet and Bretar, 2009; Pirotti, 2011; Wagner et al., 2006), and to further reduce the uncertainty caused by the theoretical model error. The innovative aspects of this study consist of (1) integrating the nonlinear Bayesian concept with waveform data to provide a novel decomposition method for small-footprint FW LiDAR data; and (2) generating a consistent, transparent knowledge learning framework to quantify the uncertainty emerging from data and trace the uncertainty propagation along the processing steps. In this study, we did not intend to use the Bayesian decomposition as a proof of concept, but instead we applied this approach to processing millions of waveforms in our study sites to provide insights into model justification and derive a benchmark for the uncertainty quantification of FW LiDAR data.

3.2 Materials and methods

3.2.1 Study site and data

3.2.1.1 Study site

The study site is located at the San Joaquin Experimental Range (SJER), which is in the foothills of Sierra Nevada Mountains, about 32 km north of Fresno, California. Two study regions were investigated as shown in Figure III-1. One waveform sample region (SJER1) is about 6.25 ha (250 m × 250 m) with the center at 256,840.0 Easting,

4,110,820.0 Northing, and UTM Zone 11N. The SJER1 is composed of vegetation dominated by blue oak (*Quercus douglasii*), interior live oaks (*Quercus wislizeni*) and digger pine (*Pinus* sabiniana) with scattered shrubs and a nearly continuous cover of herbaceous plants. This study region was used mainly to develop the proposed model of the present paper and to compare its performance with existing approaches such as the DD, Gold and RL. Another study region (SJER2) covers approximately 136 ha with the center at 255,977.6 Easting, 4,110,780.2 Northing, and UTM Zone 11N, which primarily aims to demonstrate that the proposed model can be applied to a relatively large area, instead of a small concept-area. This region consists of mixed patches of vegetation structure and heterogeneous land cover types, including grassland, forest, water body, open ground and road.

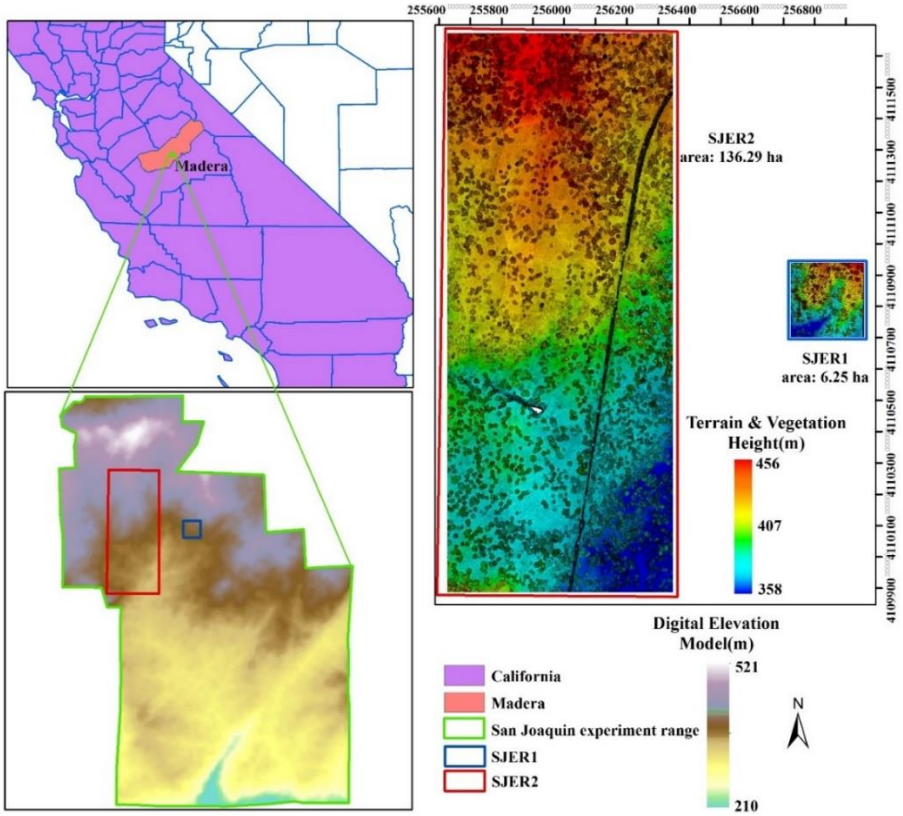


Figure III-1**.** Map of the San Joaquin Experimental Range (SJER) with location of study regions in California (left panel) and discrete-return LiDAR point image with terrain and vegetation height (right panel)

3.2.1.2 LiDAR data

The LiDAR data were collected through the National Ecological Observatory Network (NEON) Airborne Observation Platform (Kampe, 2010) which carried sensors such as a hyperspectral imaging spectrometer, a FW LiDAR sensor and a DR LiDAR sensor flying at about 1,000 m above ground level. This design can achieve sub-meter to meter scale ground resolution of study sites. Detailed technical specifications of data can be found in the study of Zhou et al. (2017).

In this study, two regions were chosen as displayed in Figure III-1. The SJER1 region had 258,667 waveforms with two flight lines, while the SJER2 region had 20,040,883 waveforms with four flight lines. The original waveform is composed of 500 time bins with 1 ns temporal resolution. Each time bin stores the digital number (DN) or intensity of corresponding backscattered pulse. The time bins with non-recorded values are assigned as zero (zero-padded) to keep the length of the waveforms constant. For geolocation of corresponding waveform, 16 basic geolocation information attributes associated with waveforms are provided. Among them, the dx, dy, and dz are the pulse direction vector that can measure position change per nanosecond. In the subsequent analysis, eight items were used for calculating the geolocation of desired time bin after decomposition. The eight items are the Easting of first return x_0 (m), the Northing of first return y_0 (m), the height of first return z_0(m), dx (m), dy (m), dz (m), the outgoing pulse reference bin location (leading edge 50% point of the outgoing pulse), and first return reference bin location (leading edge 50% point of the first return).

The DR LiDAR data for corresponding study regions were also collected with the maximum horizontal accuracy of 0.4 m and maximum vertical accuracy of 0.36 m based on the NEON'S LiDAR Algorithm Theoretical Basis Document (ATBD) (Keith and Tristan, 2015).

3.2.1.3 Field data

As part of the NEON's data collection efforts, extensive annual ground measurements of vegetation structure were conducted by the NEON Airborne Observation Platform and

the Terrestrial Instrument System (TIS) programs. The plot design followed the protocol of the NEON Terrestrial Observation System, and each plot is restricted to a 20 ×20 m region. The plot locations have been established by NEON's Field Sentinel Unit for long-term plant, insect and soil measurements. There were six field plots with 151 individual trees collected during June 2013, which were available for these two study regions. One field plot including 16 trees was located in the SJER1 study region and the other five field plots with a total of 135 individual trees were located in the SJER2 study region. The key vegetation structure variables for each tree were measured such as the location (Easting, Northing), the maximum height, and the tree species.

3.2.2 Bayesian decomposition

3.2.2.1 Theoretical background

In a Bayesian statistical framework, deterministic models are specified via mathematical equations, e.g., linear or nonlinear functions, and unknown model parameters are treated stochastically with various probability distributions.

Based on Bayes' rule, the unknown parameters of a statistical model can be written as:

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \quad \text{(III-1)}$$

where **y** is a vector of observed data which has a probability distribution depending on an unknown vector of parameters denoted as $p(\boldsymbol{y}|\boldsymbol{\theta})$, which is also known as the likelihood function. The prior distribution of model parameter vector $\boldsymbol{\theta}$ is a probability distribution that represents the experimenter's beliefs about unknown parameters prior to observing the data and was denoted as $p(\boldsymbol{\theta})$ (Hoff, 2009). $p(\boldsymbol{\theta}|\boldsymbol{y})$ is the posterior distribution of the unknown vector of parameters.

Eq. (III-1) is fundamental to understanding that the posterior distribution of unknown parameters is proportional to the prior belief about unknown model parameters, $p(\boldsymbol{\theta})$, and the probability distribution of observed data given $\boldsymbol{\theta}$ ($p(\boldsymbol{y}|\boldsymbol{\theta})$). In this way, the

posterior distribution expresses the experimenter's updated beliefs about $\boldsymbol{\theta}$ in light of the observed data **y**.

The main controversy in the Bayesian approach lies in the preparation of prior information which is subjective (Ulrych et al., 2001). Three legitimate arguments for this subjectivity are: (1) it is natural that one's conclusion is affected by one's prior opinions, (2) the priors have little effect on the posterior when a large amount of data are available, and (3) the non-informative priors can be used to express ignorance about the unknown parameters which can be assumed to be objective (Hoff, 2009). Generally, there are two kinds of prior distributions frequently used. One is the non-informative priors that are commonly used when nothing is known about the value of parameters. Another is the informative priors or the empirical priors, which can be obtained from the previous evidence or empirical data.

Based on the previous study (Wagner et al., 2006), an individual waveform can be modelled with a mixture of Gaussian distributions. Therefore, it is natural to assign the distribution of the waveform as $p(\boldsymbol{y}|\boldsymbol{\theta})$ with a mixture of Gaussian distributions, which can be also interpreted as the likelihood of model. The shape of the waveform is determined by the parameters of the Gaussian distribution that can be easily obtained as the prior distribution of parameters through the peak identification algorithm (Zhou et al., 2017). The core of this algorithm is to identify the peaks by comparing the three adjacent intensities of the waveform and then selecting peak(s) when the corresponding peak intensity is higher than one-fifth of maximum intensity of the given waveform. Meanwhile, the number of peaks for the waveform is also obtained through this process.

Once we formulated the prior distribution of the parameters and the likelihood function, the concept of Eq. (III-2) can be used to derive the posterior distribution of model parameters through MCMC simulation (Gelfand and Smith, 1990). MCMC is a crucial technique for the rapid expansion of the Bayesian inference in science. There are cases that some parameters' posterior distributions are difficult or impossible to sample when the non-conjugate priors are used or the integration of parameters is conducted over a high dimensional parameter space (Hoff, 2009). In such conditions, the MCMC method

can be helpful by approximating the true posterior distribution using the joint distribution $p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})$ instead of directly sampling from the integration of posterior distribution for parameters of interest $p(\boldsymbol{\theta}|\boldsymbol{y})$.

3.2.2.2 Waveform decomposition application

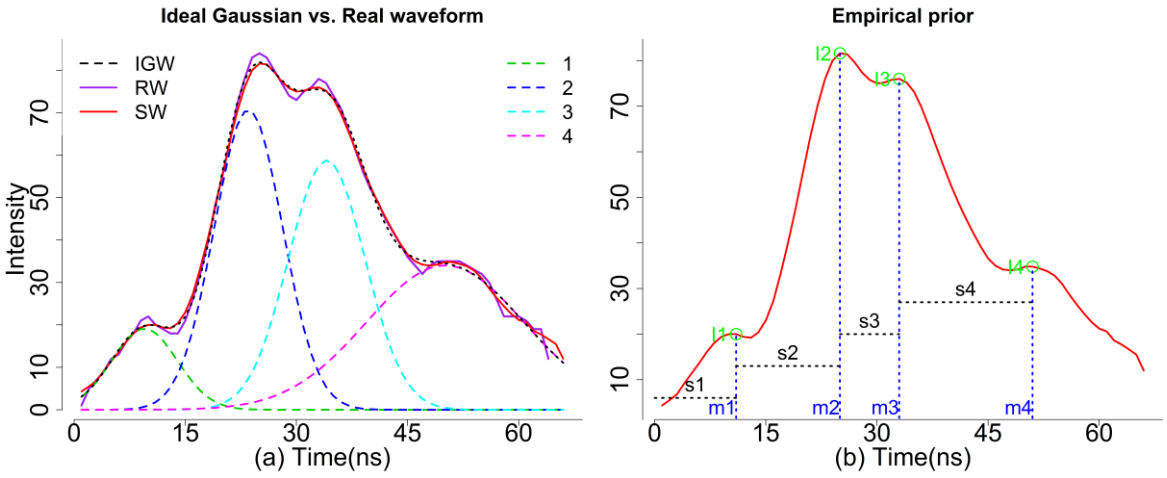


Figure III-2. (a) Illustration of ideal Gaussian distribution waveform (IGW, black dash) vs. real waveform (RW, purple) and smoothed waveform (SW, red). The number (1, 2, 3 and 4) represent the individual Gaussian components. (b). Empirical priors derived from the SW through peak identification algorithm.

In a Bayesian context, the nonlinear model can be formulated in the following form

$$y_i = f(x_i, \boldsymbol{\theta}) * \epsilon_i \quad \text{(III-2)}$$

where y_i is the observed data, $f(x_i, \boldsymbol{\theta})$ is a nonlinear function with parameters $\boldsymbol{\theta}$ and predictor x_i, ϵ_i is an independent error with $\log\epsilon_i \sim N(0, \tau^2)$, τ is the standard deviation of $\log\epsilon_i$, and x_i is the ith time bin of the waveform. Each waveform is reconstructed in terms of the main model part ($f(x_i, \boldsymbol{\theta})$) with a set of parametric functions and the error part. For the main model part, the coherence between proposed configurations and the real waveforms is measured. The multiplicative error is used here mainly because FW LiDAR data are typically restricted to be nonnegative (Gelman et al., 2015), and the multiplicative format is also convenient to formulate distribution for the error part. These formulations are consistent with the real FW LiDAR data as demonstrated in Figure

III-2(a). There is a difference between the ideal Gaussian waveform (IGW) $f(x_i, \boldsymbol{\theta})$ (black dash line) and the raw waveform (RW) $f(x_i, \boldsymbol{\theta}) * \epsilon_i$ (purple line) that corresponds to error part $(\epsilon_i - 1)^* f(x_i, \boldsymbol{\theta})$. The ideal Gaussian waveform is derived from Eq. (III-3) by summing four Gaussian components (j = 1, 2, 3 and 4) as shown in Figure III-2 with dash lines with different colors.

$$f(\boldsymbol{x}, \boldsymbol{\theta}) = \sum_{j=1}^{n} A_j \exp\left(-\frac{(\boldsymbol{x}-u_j)^2}{2\delta_j{}^2}\right) \quad \text{(III-3)}$$

where *n* is the number of the Gaussian components, A_j, δ_j and u_j are the amplitude of the peak, the standard deviation and the time location of the peak for j[th] waveform component, respectively. Eq. (III-3) gives rise to total 3*n parameters associated with the number of Gaussian components. A_j, u_j and δ_j are restricted to nonnegative values.

To reduce the impact of detected "false" peaks of the raw waveform resulted from noise, especially at the beginning and tail of the waveform, a mean filter was conducted prior to subsequent processing. We called the waveform after filtering as the smoothed waveform (SW) (red line) that would be employed for subsequent analysis. A visual inspection showed that the smoothed waveform was nearly overlapping with the ideal Gaussian distribution waveform (Figure III-2 (a)) which may justify the use of the Gaussian model for fitting FW LiDAR data. We explored more details of model choices in the Section 2.4 to quantitatively test whether this assumption was valid. The number of Gaussian components, *n*, for each waveform varies depending on the number of peaks detected along the pulse line. The unknown parameters $\boldsymbol{\theta}$ include $A_1, u_1, \delta_1, \ldots, A_n, u_n, \delta_n$.

After log transformation of the Eq. (III-2), the log-likelihood of y_i can be written as:

$$log\varepsilon_i = logy_i - log\, f(x_i, \boldsymbol{\theta}) \quad \text{(III-4)}$$

$$p\,(logy_i | \boldsymbol{\theta}, \tau) = N(log\, f(x_i, \boldsymbol{\theta}), \tau^2) = \frac{1}{\sqrt{2\pi\boldsymbol{\tau^2}}} \exp\left(-\frac{(logy_i - \log f(x_i, \boldsymbol{\theta}))^2}{2\tau^2}\right) \quad \text{(III-5)}$$

To compare the influence of prior information on the model performance, the non-informative priors and the empirical priors derived from the raw waveforms were assigned to the model. Here, the non-informative priors indicated that assigning equal

probabilities to all possible values of parameter space, named it the flat priors in subsequent analysis.

Based on the statistical summary of raw FW LiDAR data, we narrowed down the reasonable range of these parameters to [10, 150], [15, 100] and [4, 15] as the parameter space of A_j, u_j and δ_j, respectively. Hence, a uniform distribution (U) was assigned to each parameter to express the ignorance of the effect of parameters' prior distribution on the outcomes, and the prior distribution for each parameter followed $A_j \sim U(10, 150), u_j \sim U(15,100)$ and $\delta_j \sim U(4,15)$. Another option for specifying the priors were to use the empirical priors that were derived from the corresponding SWs through peak identification algorithm (Zhou et al., 2017). This algorithm is mainly to estimate the number of Gaussian components n and approximated peak locations. According to Figure III-2(b), the time bin for the peak of one waveform component m_j (m_1, m_2, m_3 and m_4) was associated with the location of the peak that corresponds to u_j. The corresponding intensity I_j (I_1, I_2, I_3 and I_4) at the peak was related to A_j in Eq. (III-4). δ_j was much more difficult to interpret, therefore we used a third of the difference between consecutive peaks (s_j/3) to roughly represent prior information of δ_j. To sum up, we specified the prior distribution of A_j, u_j *and* δ_j to follow the normal distribution (*N*) with $A_j \sim N(I_j, 10^2)$, $u_j \sim N(m_j, 5^2)$, $\delta_j \sim N(s_j/3, 3^2)$ and $\tau \sim N(0, 0.5^2)$.

A log posterior distribution of the model was obtained through all prior information about parameters of interest and the data distribution $p(\mathbf{y}|\boldsymbol{x}, \boldsymbol{\theta}, \tau)$. For the flat priors, the posterior distribution was the likelihood of data (Eq. (III-6)). For the empirical priors, the posterior distribution of model was expressed in Eq (III-7):

$$p(\boldsymbol{\theta}, \tau|\boldsymbol{y}) \propto \text{p}(\mathbf{y}|\boldsymbol{\theta}, \tau)\, p(\boldsymbol{\theta})\, p(\tau)$$

$$\propto \prod_{i=1}^{m} \frac{1}{\boldsymbol{\tau}} \exp\left(-\frac{1}{2\tau^2}(logy_i - logf(x_i, \boldsymbol{\theta}))^2\right) * p(\boldsymbol{\theta}) p(\tau)$$

Flat priors: $$\propto \frac{1}{\tau^m} \exp\left(-\frac{1}{2\tau^2}\sum_{i=1}^{m}(logy_i - logf(x_i, \boldsymbol{\theta}))^2\right) \quad \text{(III-6)}$$

Empirical priors:

$$\propto \frac{1}{\tau^m} \exp\left(-\frac{1}{2\tau^2}\sum_{i=1}^{m}(logy_i - logf(x_i, \boldsymbol{\theta}))^2\right) \prod_{j=1}^{n} p_1(A_j)\, p_2(u_j)\, p_3(\delta_j)\, p(\tau) \quad \text{(III-7)}$$

$$p_1(A_j) \sim N(I_j, 10^2); p_2(u_j) \sim N(m_j, 5^2); p_3(\delta_j) \sim N(s_j/3, 3^2); p(\tau) \sim N(0, 0.5^2)$$

It is reminded that m is the number of observations for each waveform and n is the number of Gaussian components of the corresponding waveform.

3.2.3 Model implementation

Our aim was to decompose the waveform data using the above models to perform inference about quantities of the unknown parameters of interest. The model was implemented in R using the brms packages (Buerkner, 2016) which can fit generalized non-linear mixed models using Stan by performing the Bayesian inference and the optimization for the user-specified model (Gelman et al., 2015). Stan is a C++ program to perform Bayesian inference which is composed of four main blocks: variable declarations, parameter statements, transformed parameters and model blocks. Detailed descriptions of model structure and procedures are given in Appendix.

Model dialogistic. We measured the model's convergence using the potential scale reduction factor, named Rhat ($\hat{R}$), which is a statistical criterion to test how well the Markov Chains are mixing, or moving around the parameter space. $\hat{R}$ close to one indicates convergence, while high $\hat{R}$ value implies that we should run a longer chain to improve convergence to the stationary distribution. The effective sample size was also generated to represent the equivalent number of independent iterations of the chain. It is a criterion for the estimation efficiency. Generally, the higher the effective sample size, the more reliable estimates can be achieved (Gelman et al., 2014).

Model inference. The total samples may have divergent values before the chain reaches a stationary state, therefore the model inference was conducted on the posterior samples after dropping burn-in samples. Generally, through these posterior samples, the distribution of parameters ($\boldsymbol{\theta}$) and summary measures, such as mean, mode, standard deviation and percentiles for each parameter can be derived. The main advantage of simulating from posterior samples is that we can generate as many values as we wish and thereby minimize errors in approximating quantities of interest. For each parameter, these draws could be used to approximate credible interval (CI) as shown in Figure III-3(b).

3.2.4 Model reasonableness

The Gaussian function has been mostly used to decompose the waveform LiDAR data, under the implicit assumption that the Gaussian model is capable of reconstructing detected objects based on waveform shapes. However, few studies have quantitatively or statistically justified the reasonableness of this assumption.

According to the previous study (Mallet et al., 2009), several models that could be used to fit waveform LiDAR data. In this study, we employed three representative models to explore the reasonableness of models in a Bayesian context: the Weibull model, the Adaptive Gaussian model and the Gaussian model. The Gaussian model is most frequently used model for waveform decomposition. The Adaptive Gaussian distribution has the form as Eq. (III-8) which can minimize the residual of the model by introducing another variable which is also known as rate parameter (λ).

$$f_{AG}(\boldsymbol{x}, \boldsymbol{\theta}) = \sum_{i=1}^{n} A_i \exp\left(-\frac{(x-u_i)^{\lambda}}{2{\delta_i}^2}\right) \quad \text{(III-8)}$$

In this study, the rate parameter λ, as a stochastic variable, was assigned to follow normal distribution N~ (2, 0.25), because the rate parameter in the Gaussian model is 2 and the Adaptive Gaussian model's rate parameter should be close to this value.

The Weibull model was introduced since it enables us to simulate either symmetric or asymmetric peaks with four unknown parameters. This model has been successfully applied for Synthetic Aperture Radar (SAR) image processing (Tison et al., 2004). The Nakagami and Bur functions (Mallet et al., 2009) are also capable of simulating the waveform shape with four parameters as the Weibull function. These three functions share the same feature that all can simulate asymmetric and symmetric waveforms with the same number of parameters. Here, the Weibull function was selected to represent this class of potential models for the waveform decomposition. The Weibull distribution function used here can be written as:

$$f_W(\boldsymbol{x}, \boldsymbol{\theta}) = \sum_{i=1}^{n} A_i \frac{\mathrm{k}}{\delta_i}\left(\frac{x-u_i}{\delta_i}\right)^{k-1} \exp\left(-\left(\frac{x-u_i}{\delta_i}\right)^{k}\right) \quad \text{(III-9)}$$

where A_i is the amplitude, k (>0) is the shape parameter that controls the behavior or the shape of the distribution, and δ_i (>0) is the scale parameter that controls the spread of the distribution. The shape parameter can capture the asymmetry or skewness of the

waveforms that overcomes the disadvantage of the Gaussian function, which is only suitable for symmetric distributions. u_i is a location parameter in the Weibull model. However, this parameter is not useful in waveform interpretation and subsequent geolocation transformation, since it does not have any physical meaning in our case.

The predictive accuracy of models is generally measured with the deviance information criterion (DIC) for the Bayesian model selections. In the present study, a more relevant criterion, the Watanabe-Akaike information criterion (WAIC), was adopted to provide a basis for model assessment and model selection (Vehtari et al., 2016). WAIC is a fully Bayesian criterion that uses posterior distribution of existing simulation draws rather than a point estimate to approximate leave-one cross validation for estimating pointwise out-of-sample prediction accuracy.

One hundred forty waveforms with a different number of waveform components that represented different levels of model complexity were randomly selected as samples to check the model reasonableness of waveform decomposition. Each waveform was fitted with the above three models ($f(\boldsymbol{x},\boldsymbol{\theta}), f_{AG}(\boldsymbol{x},\boldsymbol{\theta}), and\ f_W(\boldsymbol{x},\boldsymbol{\theta})$) and WAIC, uncertainty bounds and residual standard error of model (SE) were reported. The model choice would be based on these criteria, physical meaning of corresponding parameters and processing efficiency.

3.2.5 Model efficiency

There are several commonly used MCMC methods such as Gibbs sampling, Metropolis algorithm, and Metropolis-Hastings (MH) algorithm. In this study, the MCMC simulation was achieved using the Hamilton Monte Carlo (HMC) algorithm to enhance the model efficiency.

The HMC is a relatively new MCMC algorithm that applies the concept of Hamiltonian dynamics to Metropolis update for simulating a Markov chain (Neal, 2011). The No-U-Turn Sampler (NUTS) (Hoffman and Gelman, 2014) was used to implement the sampling procedures of HMC and more details can be found in Appendix.

In addition, the effect of different priors on the efficiency of the model was also explored using the same sample waveforms as used in Section 2.4. Each waveform was

decomposed twice using the flat priors and empirical priors with the Gaussian model. The computation time for each model was recorded and the average time for given number of waveform components was reported. In this step, the number of iterations and burn-in had been assumed to be adequate to derive the accurate approximation of parameters no matter which prior was used. According to Table III-5, the empirical priors converge faster. As a result, we used the empirical priors to process all other waveforms.

3.2.6 Geolocation transformation

The Bayesian decomposition can provide estimated quantiles (u_{ij}, from 1% to 99% of posterior samples including mode estimate u_{i50}) for possible target time bin locations (u_i), corresponding standard error, effective sample size and $\hat{R}$ of A_i, u_i, σ_i for each waveform. The 3D point clouds were generated by combing the original georeferenced data such as x_0, y_0, z_0, dx, dy and dz provided by the NEON datasets with the estimated time bin (u_i, $u_{i1} \ldots u_{i99}$). The leading edge position for each detected peak was used to compute the geolocation of desired time bin by incorporating the full width at half maximum (FWHM).The detailed calculation processes are given in (Zhou et al., 2017).

3.2.7 Performance evaluation

Both the DD and Bayesian decomposition methods can be classified as the decomposition method instead of the combined deconvolution and decomposition method in terms of processing steps. However, the DD method is different from the Bayesian decomposition method which belongs to the probabilistic approach, but the DD, Gold and RL methods are the deterministic approach which only has one estimate for a parameter or possible target position. Bayesian decomposition method here generates the distributions for the parameters, from which multiple possible estimates for the parameter values could be obtained e.g., using the quantiles. In our case, we generated 99 possible quantile estimates for the parameter u_i with probability as shown in Figure III-3(b), which resulted in 99 possible point clouds after geolocation transformation. To compare the performances of the Bayesian decomposition method with deterministic methods, the

point cloud with highest probability using the mode estimate (u_{i50}) for geolocation transformation was selected to conduct the performance evaluation.

The Bayesian decomposition cannot converge for the extremely irregular or noisy waveforms, which results in some noisy points after geolocation transformation. With the aid of LAStools (Isenburg, 2012), the noisy points were deleted before conducting the method comparisons of the point cloud. In this study, the point cloud obtained using mode estimates (u_{i50}) with Bayesian decomposition method was compared with the DD, Gold and RL approaches from the previous study (Zhou et al., 2017) at two different levels: point cloud and individual tree's metrics such as the number of points at various height bins, percentile heights and canopy point density.

Point cloud comparisons. Point clouds are the primary result of waveform decomposition and their accuracy significantly influences the quality of their derived products such as percentile heights and Digital Terrain Models (DTMs). Thus, we computed the Hausdorff distances (Mémoli and Sapiro, 2004) between the waveform-based point cloud derived from different methods and the DR LiDAR point cloud. This comparison was named C2C in subsequent analysis. The thrust for the Bayesian decomposition method is to avoid the error brought by interpolation and variability of area based products such as DTM caused by various grid cell sizes. In addition, the point cloud comparison is a natural and direct way to evaluate the surface representation without adding any intermediate step (Mémoli and Sapiro, 2004).

The principle of the Hausdorff distance is that for each point of a compared cloud, the nearest neighbor method is used to search the nearest point in the reference cloud (the DR point cloud) and then compute their Euclidean distance. This process was implemented in the CloudCompare software (Girardeau-Montaut, 2015). Meanwhile, all points' X, Y, Z differences were also generated for the Bayesian decomposition, DD, Gold and RL approaches.

Individual tree metrics comparisons. In addition to the comparisons for the grand picture such as point clouds, the comparisons at the individual tree level were also explored to present a more comprehensive comparison of methods' performances.

Individual tree dimension's LiDAR metrics such as percentile heights, median height and crown density are crucial for characterizing canopy structure and estimating biomass (Falkowski et al., 2009; Popescu, 2007; Zhao et al., 2011). We randomly selected 121 trees from the SJER to compare their total number of points, the number of non-ground points (the elevation larger than reference DTM), the non-ground canopy point density, the percentile heights and median height using the Bayesian decomposition method with corresponding DR LiDAR data results. Additionally, the results of individual tree level from the DD, Gold and RL approaches were also incorporated into comparisons that had been done in the previous study.

Due to fewer points at lower heights corresponding to the understory, the height bin width was 2 m when the tree height was below 4 m, and it became 1 m when the tree height was larger than 4 m. The total number of points and corresponding ratio of points in each height bin were summarized. To further compare the performance of different waveform processing methods, the normalized percentile heights (subtracting the minimum elevation for each tree boundary) were calculated. Given the point cloud of one tree, ten height metrics including 10^{th}, 20^{th}, 30^{th}, 40^{th}, 50^{th}, 60^{th}, 70^{th}, 80^{th}, 90^{th} and 100^{th} (maximum height) percentile heights were extracted to demonstrate the vertical structure of vegetation based on the height of LiDAR points. These metrics not only help to predict biomass, but also can quantitatively measure the waveform LiDAR data's penetration advantage.

The canopy point density for individual trees was also analyzed for different approaches since it was beneficial to map tree stem and crown (Lee and Lucas, 2007). We randomly selected 21 trees from 121 trees to show detailed results of the comparisons. Here, the canopy point density was represented by the number of above ground points per square meter.

Field data calibration. What the above comparisons share is that DR LiDAR data were adopted as reference data that had been successfully applied in many studies (Allouis et al., 2013; Chen, 2007; Zhou et al., 2017) for its high accuracy of measuring height. In this study, we not only compared the point cloud and individual trees' metrics

derived from FW LiDAR data to corresponding DR LiDAR (reference) provided by the NEON, but also used the field-measured tree height to evaluate results. To facilitate comparisons of different methods' performances, the average bias, standard deviation and Root Mean Squared Error (RMSE) of their differences were computed. According to 151 field-measured trees' locations, the values from waveform-based Canopy Height Models (CHMs) using u_{i50} were extracted. To make our extracted values more representative, we generated a 1m buffer for each location and then averaged the values fell in each buffer.

3.2.8 Uncertainty analysis

3.2.8.1 Uncertainty propagation

The essential feature of the Bayesian approach is the explicit quantification of uncertainty introduced by incorporating multiple levels of randomness or various sources of errors. Through estimating the predictive parameter uncertainty, rigorous error propagation along the processing steps can be quantified. In this study, there were different sources of uncertainty originating from data themselves and along the processing steps, which would accumulate and propagate to the final products. Most of the previous studies are based on the deterministic models or approaches that only seek a single value and ignore the noise or error inherent in the data and approaches. To this end, we conducted a comprehensive uncertainty analysis to quantify the uncertainty of results at different steps including parameter estimates, point cloud generation and surface model generation (Figure III-3).

Through the probability distribution of inference using the Bayesian approach, the summary measures of unknown parameters such as mode, percentiles and standard deviation are obtained instead of a single estimate. Unlike the deterministic method, the Bayesian credible region is characterized by the 95% highest posterior density (HPD) region rather than 95% confidence interval. The main difference is that the HPD region can be discontinuous when the parameter's posterior density is multimodal or asymmetric (Hoff, 2009), while the confidence interval region is always continuous. At this stage, the Gaussian model had been chosen to fit the waveforms that gave us suitable physical interpretation and accurate estimation of parameters. The posterior density of the

individual parameter was generally used to obtain their CIs. The fitting functions used here followed the normal distribution which was symmetric. Thus, we directly used the empirical quantiles of the posterior samples to approximate the uncertainty of the peak locations (blue dash) as displayed in Figure III-3(a). The value of peak location was regarded as a realization of MCMC process. The distribution of possible peak locations with probability from the posterior samples after Bayesian decomposition was shown in Figure III-3(b). The estimated quantile peak locations u_{ij} starting from 1% to 99% of posterior samples were chosen to conduct geolocation transformation to derive possible points located along the blue arrow (Figure III-3(c)). We used the mode estimate u_{i50} of possible peak locations to generate the point cloud as the background in Figure III-3(c), and several points with different length of arrows were selected as examples to demonstrate the uncertainty of the corresponding point when we used their quantile estimates ($u_{i1} \dots u_{i99}$). However, quantifying uncertainty of these individual points separately was not useful to some extent, since most of the subsequent studies were conducted on products derived from these points instead of individual points, such as the point clouds (Figure III-3(d)), DTM and CHM (Figure III-3(e)). The u_{ij} quantile point cloud was composed of the points derived from the geolocation transformation using u_{ij} for all waveforms located in the region. Figure III-3(d) demonstrates the u_{i1} and u_{i99} point clouds to demonstrate the possible maximum uncertainty of point cloud products in a sample region. Due to the relatively small differences among 99 point clouds, we showed only two point clouds. As mentioned earlier, the Hausdorff distance was employed to calculate the distance of these waveform-based u_{ij} quantile point clouds and the DR LiDAR point cloud. For each u_{ij} point cloud, the average bias (mean), standard deviation and RMSE of X, Y, Z and point distances were derived. The uncertainty of point cloud distances was characterized by the distribution of average bias and RMSEs derived from these quantile point clouds. A similar method was also conducted on the waveform-based surface models such as DTM and CHM. More specifically, the DTM for each quantile point cloud were generated through the *lasground* and *las2dem* implemented in LAStools after deleting noisy points. Regarding the CHM, we followed the steps described by

Khosravipour (2014) and implemented these steps in LAStools to obtain the 99 quantile CHMs. To further quantify the performance of these methods, we compared the waveform-based DTMs and CHMs from quantile point clouds with corresponding reference data and computed the evaluation criteria such as the average bias and RMSE. The graphical and statistical methods were employed to analyze the comparison results and their corresponding uncertainties.

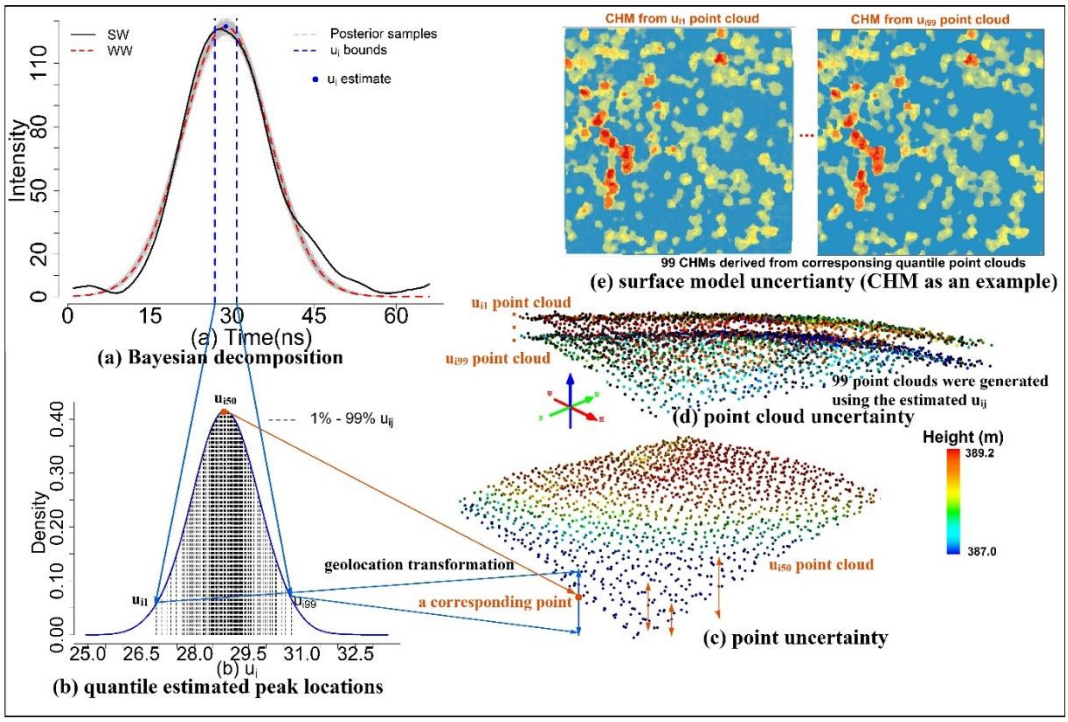


Figure III-3. Illustration of uncertainty propagation from data to the parameter estimates, point, point cloud and surface model such as CHM using Bayesian decomposition method. (a) The uncertainty of peak location (parameter uncertainty) using Bayesian method to fit the waveform. SW (black) represents the original waveform after smoothing, WW (red dash) represents the waveform using the mode estimates of Bayesian method, and the gray shadow represents the possible solutions for fitting the waveform. (b) The 99 quantile estimates of the possible peak locations from the u_i posterior distribution. (c) The point uncertainty propagated from the parameter uncertainty through geolocation transformation at a sample region with the background of u_{i50} point cloud. (d) Possible point clouds generated from 1% quantile estimate and 99% quantile estimated peak locations as examples. (e) Possible surface models such as

Canopy Height Models (CHMs) generated from u_{i1} quantile point cloud and u_{i99} quantile point cloud as examples.

3.2.8.2 Uncertainty of accuracy assessment with field data

A limited number of field observations and imprecise field-measured data significantly affect the calibration results and accuracy assessment (Yao et al., 2012). Both require us to conduct uncertainty analysis of field data to enhance the credibility of calibration. However, the uncertainty of field data is difficult or impossible to quantify since data providers do not clearly state which error sources were considered. In this study, we assumed the field data were "true", and the uncertainty of estimated tree height using the Bayesian decomposition method was analyzed. Specifically, 1 m buffers generated from field-measured individual tree locations (X and Y) were first used to extract possible tree points from the u_{i50} point cloud. For each buffer, the Z values of the points above 95th percentile height were selected as the possible tree height and these points were employed to identify the waveform(s) that fell in the tree region. The uncertainty of these individual trees' height was quantified through the 95% CI of peak locations from selected waveform(s) after the Bayesian decomposition. Moreover, we calculated the uncertainty of RMSE for the estimated tree height based on the nearest possible value and farthest possible value from each tree's 95% CI to the field-measured data. An overview of the proposed methodology was generated to summarize the major steps implemented in this paper as shown in Figure III-4.

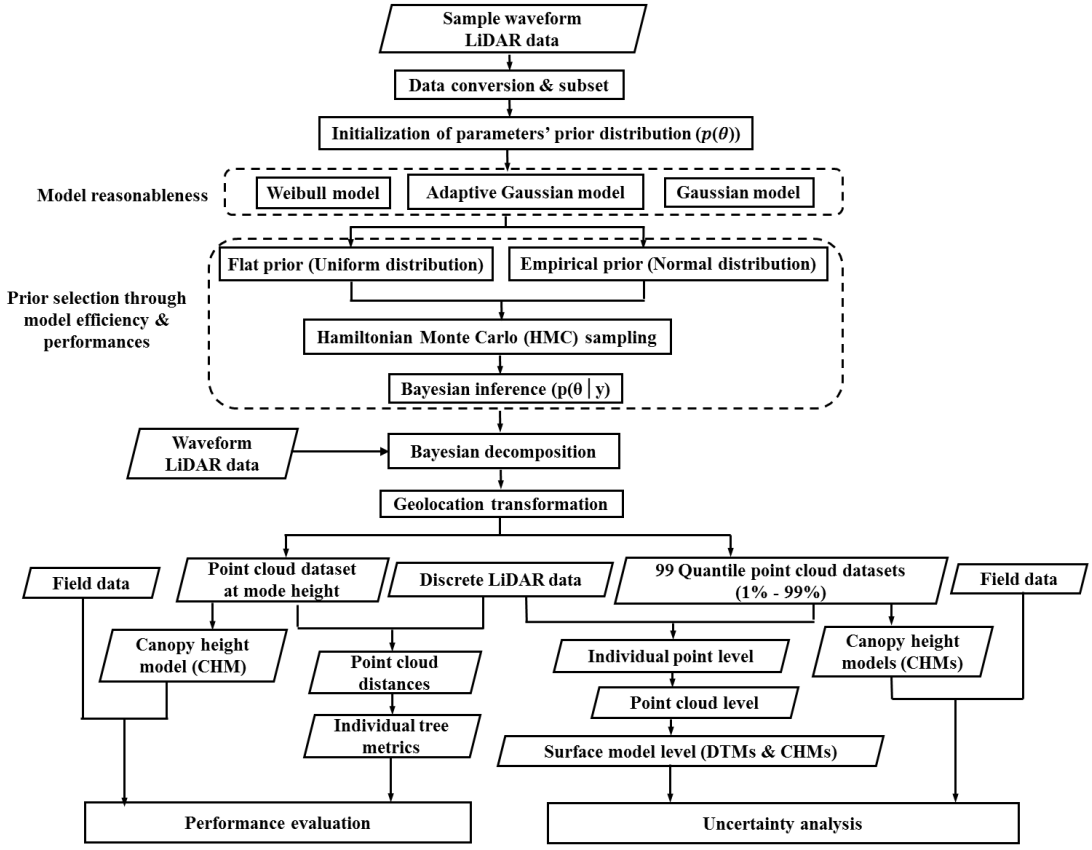


Figure III-4**.** Flowchart for the Bayesian decomposition of waveform LiDAR data with uncertainty analysis

3.3 Results

3.3.1 Model reasonableness

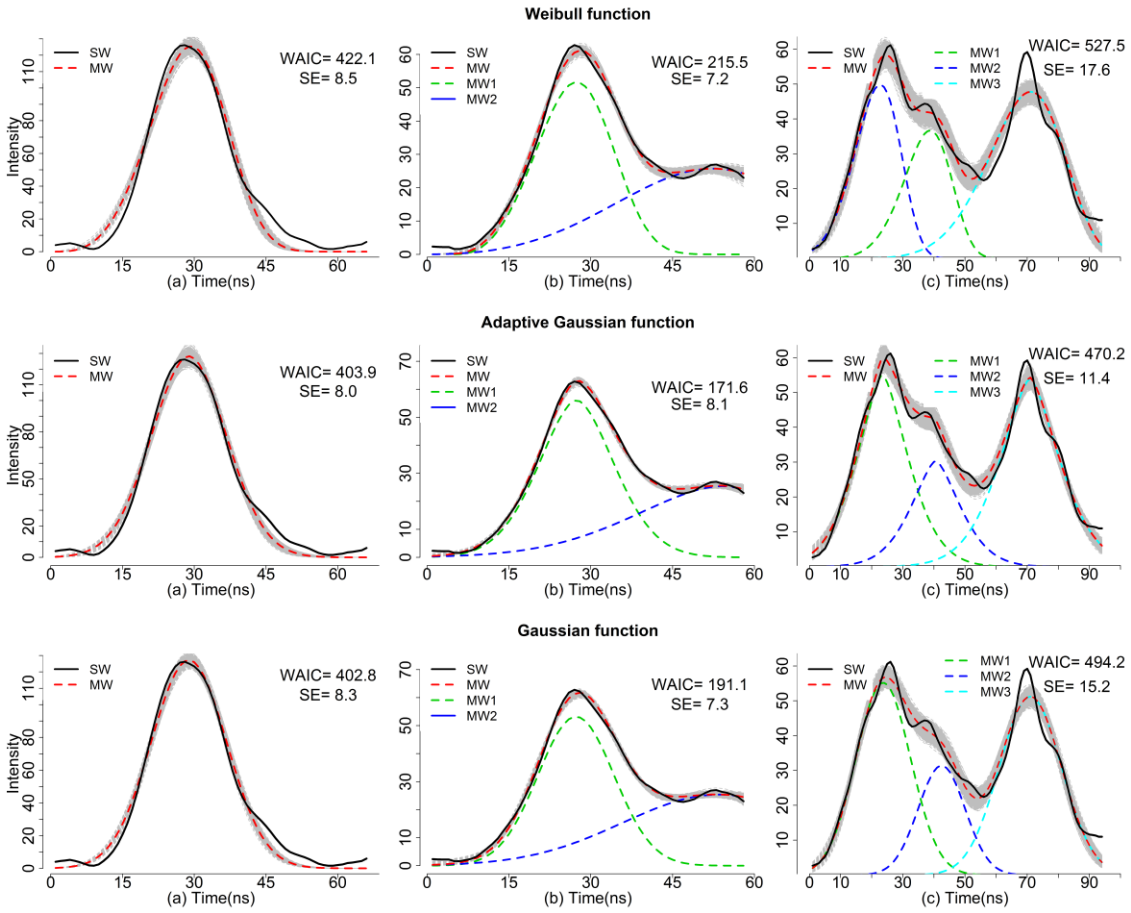


Figure III-5. Examples of Bayesian decomposition of FW LiDAR data using three models (the Weibull, Adaptive Gaussian and Gaussian models) and their corresponding uncertainties (gray shadow). Black solid line represents the smoothed waveform (SW) using a mean filter, red dash line represents the modeled waveform (MW), and other color dash lines represent modelled waveform components. WAIC is the Watanabe-Akaike information criterion and SE is the residual standard error of the model.

We modeled 140 sample waveforms with different complexities (the number of components) using the Weibull, Adaptive Gaussian and Gaussian models within the Bayesian framework, and three representative examples of decomposition results including waveform components (colored dash lines), uncertainty (gray shadow), the

corresponding WAIC and SE of model are demonstrated in Figure III-5. As expected, all models worked well with a relatively small difference when we visually inspected the SW and MW. There was no noticeable loss of the fitting accuracy using these three models when the waveform was relatively simple.

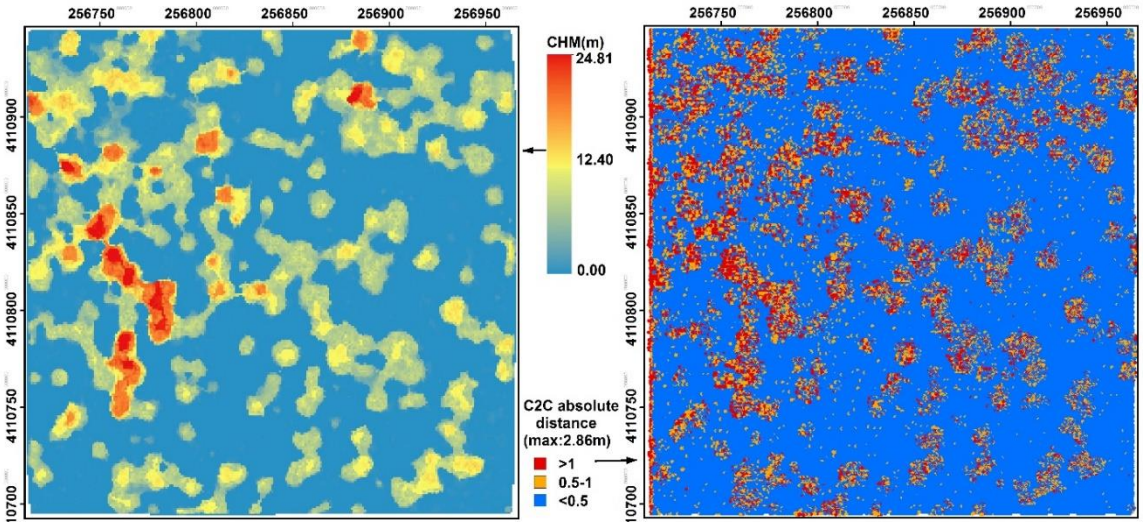
However, the Weibull function model became less fitted and yielded a wider gray shadow with larger uncertainty when the waveform had higher number of Gaussian components. The statistics summary of the models further confirmed this inspection, as the Weibull model always achieved the largest WAIC and SE given the same waveform. Moreover, most of the waveforms in the study site were close to symmetric distributions, which diminishes the advantage of the Weibull model that is capable of modeling the asymmetric waveforms. Among these three models, the smallest WAIC was achieved using the Adaptive Gaussian model to fit complex waveforms with more than one peak. However, this model might face the problem of overfitting, since the rate parameter (λ) of the Adaptive Gaussian model for each waveform was adjusted mainly for minimizing residuals of the fitted model. Consequently, the noise contained in the data might be considered as the main model part ($f(\boldsymbol{x_i}, \boldsymbol{\theta})$) instead of the error part. The experiment of sample waveforms showed that 68 out of 95 waveforms with one component could generate the smallest WAIC using the Gaussian model. Additionally, over 75% of all waveforms in the SJER study site were considered to be one component according to the inspection of the waveform components (n). These gave us more confidence to utilize the Gaussian model instead of the Adaptive Gaussian model to perform waveform decomposition in terms of model accuracy and uncertainty.

In addition, the physical interpretation of the parameters was also limited when we applied the Weibull model or Adaptive Gaussian model. For the four-parameter Weibull model, the estimates can't be employed as a location like Gaussian model's peak location to calculate the geolocation of the extracted points. It was meaningless or difficult to interpret the rate parameter of the Adaptive Gaussian model from the decomposition perspective and geolocation transformation. Furthermore, the experiment of these waveforms showed that the Weibull and Adaptive Gaussian models took more time to

find optimized parameters and reach convergence for an additional parameter. Therefore, we concluded that the Gaussian distribution model was the most suitable model for FW LiDAR data decomposition based on the accuracy, uncertainty, physical meaning and processing efficiency.

3.3.2 Performance evaluation

3.3.2.1 Point cloud comparisons



(a) (b)

Figure III-6(a). The canopy height model (CHM) generated from Bayesian approach (Left). (b). The spatial distribution of the distance between waveform-based point clouds using Bayesian decomposition method and the DR point clouds (C2C) at SJER1 region (Right).

The C2C validation of the SJER1 study region was executed as displayed in Figure III-6(b). To associate the C2C distances across the study site with the vegetation distribution, the CHM derived from the Bayesian decomposition method (Figure III-6(a)) was plotted against the C2C distances' spatial distribution pattern. It was worthy to note that larger C2C absolute distances (>0.5 m) were more likely to occur at the vegetation part with higher CHM values. The ground region was the most accurate portion with relative small distance when we compared waveform-based point cloud to the corresponding DR LiDAR point cloud.

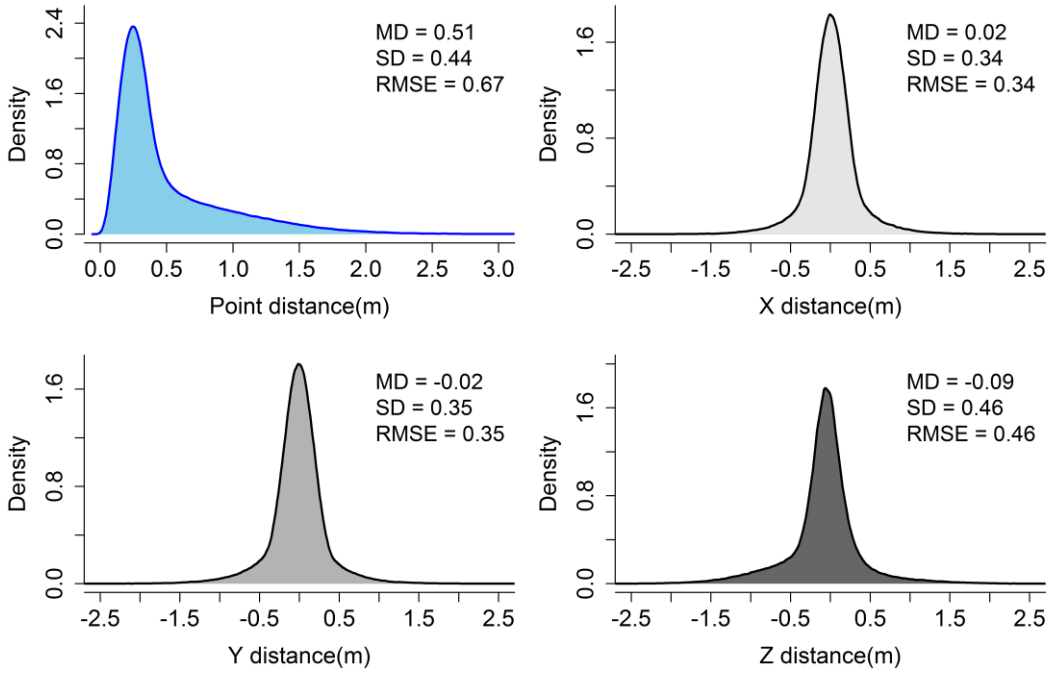


Figure III-7. Probability distribution of the C2C point distances (sky blue), horizontal (X, Y) and vertical (Z) distances (gray) using mode dataset with the Bayesian decomposition method at the SJER site. MD: Mean distances between waveform-based point clouds and DR point clouds. SD: Standard deviation. RMSE: Root mean square error.

Figure III-7 depicts the distribution of point cloud distances, horizontal (X, Y) and vertical (Z) distances between the waveform-based point cloud using the mode dataset with the Bayesian decomposition method, and the DR point cloud. When examining these three 1D distances (gray), it was evident that the horizontal (X and Y) and vertical (Z) distances' distributions were symmetric around 0 m with almost the same distribution. However, a closer examination revealed that compared with the horizontal distances, the vertical distances were greater, but not markedly greater, with larger SD and RMSE. These three 1D distance's distribution together contributed to the point distances' distribution (3D) with the mean point distances and their corresponding RMSE were 0.51 m and 0.67 m, respectively. As expected, the absolute MDs and RMSEs of X, Y, Z coordinates' C2C distances were all smaller than point distances.

To further illustrate, more detailed C2C point cloud comparisons between waveform-based point clouds using four methods (the Bayesian decomposition, DD, Gold and RL) and reference point cloud were summarized in Table III-1. All methods generated satisfactory point clouds with acceptable mean point distances (< 0.84 m) and RMSEs of point distances (< 0.93 m) compared to the DR point cloud. Upon closer examination of results with different methods, the DD method outperformed the other three methods with the smallest average point distances, X, Y, Z distances, and corresponding RMSEs. Especially for the horizontal distances (X and Y), there was a tiny difference of the four methods in terms of the MD and RMSE. However, a relatively larger difference of the methods' MD and RMSE occurred at the vertical direction (Z) that consequently leads to the same pattern of four methods' point distances.

Table III-1. Summary statistics of the distances between waveform-based point clouds with the four methods (Bayesian decomposition (Bayesian), DD, Gold and RL) and the reference point cloud at the SJER site.

Items	Methods	MD	SD	RMSE	MAX	MIN
Point distances(m)	Bayesian	0.51	0.44	0.67	5.07	0.00
	DD	0.48	0.37	0.61	3.98	0.01
	Gold	0.78	0.36	0.86	4.74	0.00
	RL	0.84	0.40	0.93	4.57	0.01
X distances(m)	Bayesian	0.02	0.34	0.34	3.74	-3.73
	DD	-0.01	0.30	0.30	2.45	-3.60
	Gold	0.00	0.36	0.36	3.64	-3.68
	RL	0.00	0.39	0.39	3.47	-3.44
Y distances(m)	Bayesian	-0.02	0.35	0.35	3.69	-3.81
	DD	0.02	0.29	0.29	2.59	-3.00
	Gold	0.05	0.34	0.35	3.23	-3.59
	RL	0.05	0.38	0.38	3.71	-3.46
Z distances(m)	Bayesian	-0.09	0.46	0.46	4.37	-4.62
	DD	-0.17	0.42	0.45	3.24	-3.07
	Gold	-0.48	0.51	0.70	3.98	-4.61
	RL	-0.47	0.59	0.75	3.56	-4.23

*MD: the mean point distances between waveform-based point clouds and DR point clouds. SD: the standard deviation of distances. RMSE: the root mean square error of distances. MAX: the maximum of distances. MIN: the minimum of distances.

3.3.2.2 Individual trees' metrics

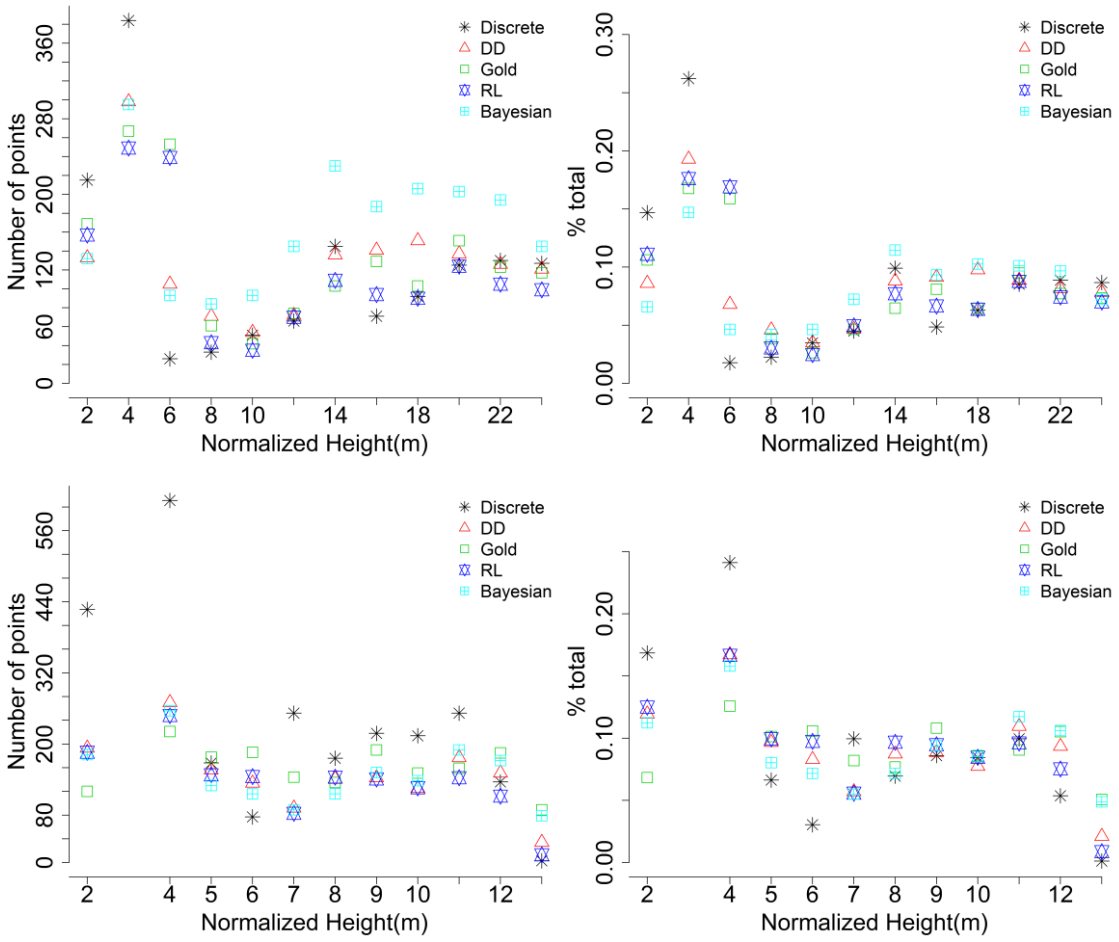


Figure III-8. Two representative examples (Top: Tree1, Bottom: Tree2) of individual trees' height bin vs. absolute point frequency and normalized frequency using DR LiDAR data and FW LiDAR data with the four waveform processing methods including the Bayesian decomposition, DD, Gold and RL approaches. Discrete represents results from DR LiDAR data. DD, Gold, RL and Bayesian represent results from the DD, Gold approach, RL approach and Bayesian decomposition method, respectively.

This section provided a comparison and evaluation of the FW LiDAR processing methods at the individual tree level. To demonstrate the robustness of approaches and reduce the selection bias, 121 randomly selected trees' height bins, percentile heights and canopy point densities were generated (see examples in Figure III-8 and Table III-2).

There were two patterns observed with respect to the point density of individual trees. Thus, two representative trees derived from waveform-based point clouds using the Bayesian decomposition, DD, Gold and RL approaches were chosen to demonstrate these patterns: one tree with higher point density (Tree1), and another tree with lower point density (Tree2) were compared to the DR LiDAR data. Overall, about 91% of all selected waveform-based tree point clouds had more dense point clouds than their corresponding DR LiDAR data. As shown in Table III-2, 16 out 21 trees' waveform-based point densities were higher than the corresponding region's DR LiDAR data, which followed the pattern of the Tree 1. Figure III-8 shows the absolute and normalized point frequency in each height bin of FW LiDAR data using four approaches and DR LiDAR data. As expected, more points were extracted from FW LiDAR data with these four approaches when examining the Tree 1's middle height bins from 6 to 22 m. The Gold and Bayesian decomposition methods outperformed other methods from the perspective of the number of points extracted from the mid-story of the tree.

This trend was not so obvious for the Tree 2, where higher point frequency in most of the height bins was observed using DR LiDAR data rather than FW LiDAR data, especially the absolute point frequency. While the normalized point frequency demonstrated advantages of FW LiDAR data for characterizing the mid-story of the tree, but the evidence was not as strong as for the Tree 1. The common feature of all trees shares was that DR LiDAR can detect more points in the lower part of their height or on the ground than FW LiDAR. This might be attributed to the fact that the tree had a dense canopy and the transmitted energy rarely reached the ground.

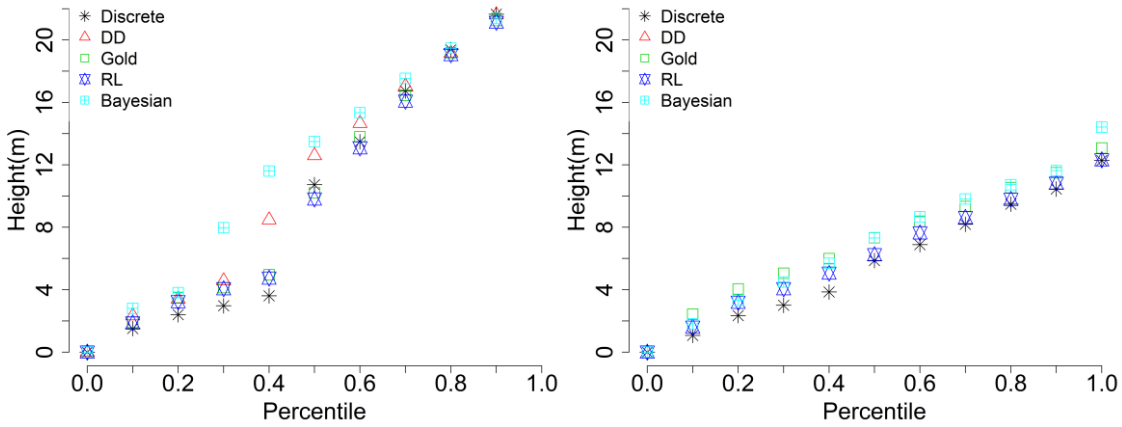


Figure III-9. The individual trees' percentile heights (Left: Tree 1; Right: Tree 2) using DR LiDAR data (Discrete), and waveform LiDAR data with the DD, Gold, RL and Bayesian decomposition (Bayesian) methods.

The percentile heights' results for these two examples demonstrated similar trend as height bin results (Figure III-9). For Tree1, the significant differences among different processing methods occurred around the median percentile, especially for the Bayesian decomposition and DD methods. Surprisingly, the four approaches and DR LiDAR data's percentile heights reached an agreement after 80th percentile height. This may indirectly imply that more points were extracted from the waveforms at the mid-story of tree height.

To reduce the ground points' effect on the comparisons, the trees' non-ground part was used to compare different methods' performances. We used 21 representative trees from the SJER1 to demonstrate comparison results. According to Table III-2, only 5 out of 21 trees' DR LiDAR data results (with bold) yielded higher canopy point density than waveform LiDAR data using these four methods. Additionally, the DR LiDAR results detected more ground points than waveform-based results as shown in Table III-2. There was no evident trend for the other three methods, however, the canopy point density results of all methods indicated that waveform LiDAR data can provide more non-ground points than DR LiDAR data. A closer examination of five individual trees with fewer points than corresponding DR LiDAR data shows that three of them (the Tree 2, Tree 12 and Tree 20) are located in the regions with fewer flight lines overlaid which may mainly contribute to the reduction of points detected.

Table III-2. Summary of *the total number of points/ number of non-ground points/ non-ground canopy point density* for 21 individual trees using DR LiDAR data (Discrete), and waveform LiDAR data with the DD, Gold, RL and Bayesian decomposition (Bayesian) methods.

Tree index	Discrete	DD	Gold	RL	Bayesian
Tree1	1465/867/7.3	1554/1114/15.0	1607/980/13.2	1429/831/12.7	2008/1593/21.3
Tree2	**2534/1337/10.8**	**1616/947/7.4**	**1761/1014/8.0**	**1490/838/6.6**	**1982/1367/10.7**
Tree3	827/353/7.0	858/448/18.1	879/425/19.1	833/419/15.9	1035/569/28.7
Tree4	887/327/7.1	898/432/15.7	1053/399/18.0	903/439/16.5	102/499/19.3
Tree5	**3029/1216/16.7**	**1128/610/13.2**	**1190/569/13.1**	**1085/546/10.7**	**1326/744/15.0**
Tree6	**3873/1211/11.2**	**2306/1059/9.8**	**2315/1079/10.0**	**2257/1046/9.7**	**2312/1202/11.2**
Tree7	2186/678/6.6	1880/880/8.6	1877/737/7.2	1859/862/8.5	1816/872/8.6
Tree8	1698/999/7.2	2256/1335/9.7	2385/1243/9.0	2422/1448/10.5	2559/1681/12.2
Tree9	543/295/7.5	659/445/11.5	669/432/11.0	652/430/10.6	671/448/11.3
Tree10	552/316/7.2	405/283/6.4	558/322/7.3	569/358/8.1	563/387/8.8
Tree11	499/193/8.1	444/184/7.8	521/239/10.1	509/268/11.3	500/243/10.2
Tree12	**1945/649/17.7**	**893/445/12.1**	**946/386/10.5**	**885/398/10.9**	**897/412/11.2**
Tree13	773/331/7.9	834/454/10.8	883/433/10.3	876/500/11.9	832/465/11.1
Tree14	1278/718/7.2	1484/951/9.5	1563/964/9.7	1553/953/9.5	1609/1101/11.0
Tree15	1275/647/7.4	1748/988/11.2	1910/1161/13.2	1853/1090/12.4	1988/1467/16.7
Tree16	1788/1071/6.5	2098/1407/9.0	2439/1596/10.2	2428/1512/9.7	2624/2086/13.4
Tree17	708/329/7.4	625/315/7.1	813/327/7.4	765/393/8.8	749/377/8.5
Tree18	770/303/7.1	691/309/7.3	747/297/7.0	753/356/8.4	746/356/8.4
Tree19	1847/902/7.2	2437/1696/13.6	2687/1678/13.5	2737/1700/13.6	3027/2423/19.5
Tree20	**1116/575/11.4**	**758/531/10.5**	**829/525/10.4**	**824/531/10.5**	**811/548/10.9**
Tree21	1660/805/6.8	2004/1182/9.9	2041/1104/9.3	2042/1100/9.2	2267/1589/13.4

3.3.2.3 Field data calibration

The accuracy of the maximum height derived from DR-based and waveform-based CHMs were assessed by comparing the field-measured data at the SJER1 study region. Overall, there was no significant difference among four waveform processing methods with regard to the average bias, standard deviation and RMSE. All of these methods generated comparable and acceptable results compared with the field-measured data. Specifically, the comparison between DR LiDAR data and the measured data produced the smallest RMSE with 1.11 m, and the DD method consistently yielded the least

accurate results with the largest average bias and RMSE. The superior waveform processing method varied with different statistics criteria used: the RL method was superior with the smallest RMSE (1.35 m), and the Gold method outperformed others with the smallest average bias.

Table III-3. Summary of the comparisons between the field-measured tree height and maximum tree height derived from the CHMs using DR LiDAR data, and FW LiDAR data with the DD, Gold, RL and Bayesian decomposition (Bayesian) methods.

Study site	Statistics criteria	Discrete (m)	DD (m)	Gold (m)	RL (m)	Bayesian (m)
SJER	Average	-0.24	0.73	-0.15	-0.32	-0.26
	Standard deviation	1.11	1.64	1.52	1.36	1.48
	RMSE	1.11	1.75	1.49	1.35	1.50

3.3.3 Uncertainty analysis

3.3.3.1 Individual parameter uncertainty

Figure III 10 shows an example of the trace plot for three parameters' sampling processes and the marginal posterior distribution of the parameters (A, u, δ) using the flat priors and empirical priors. From the trace plot (Left), we can see an obvious difference at the beginning of the parameter sampling and the flat priors (sky blue) takes more steps to reach the stable status, especially for A and δ. Given the same number of iterations, the distributions of the parameters were nearly symmetric following the normal distribution with the empirical priors (gray), while the results of the flat priors (sky blue) were not symmetric which implied the posterior sampling of the parameters did not reach the stationary stage. The Model efficiency experiment using different parameters' prior distribution showed that the performance of the flat priors can reach the same level as the empirical priors with more burn-in steps and total iterations.

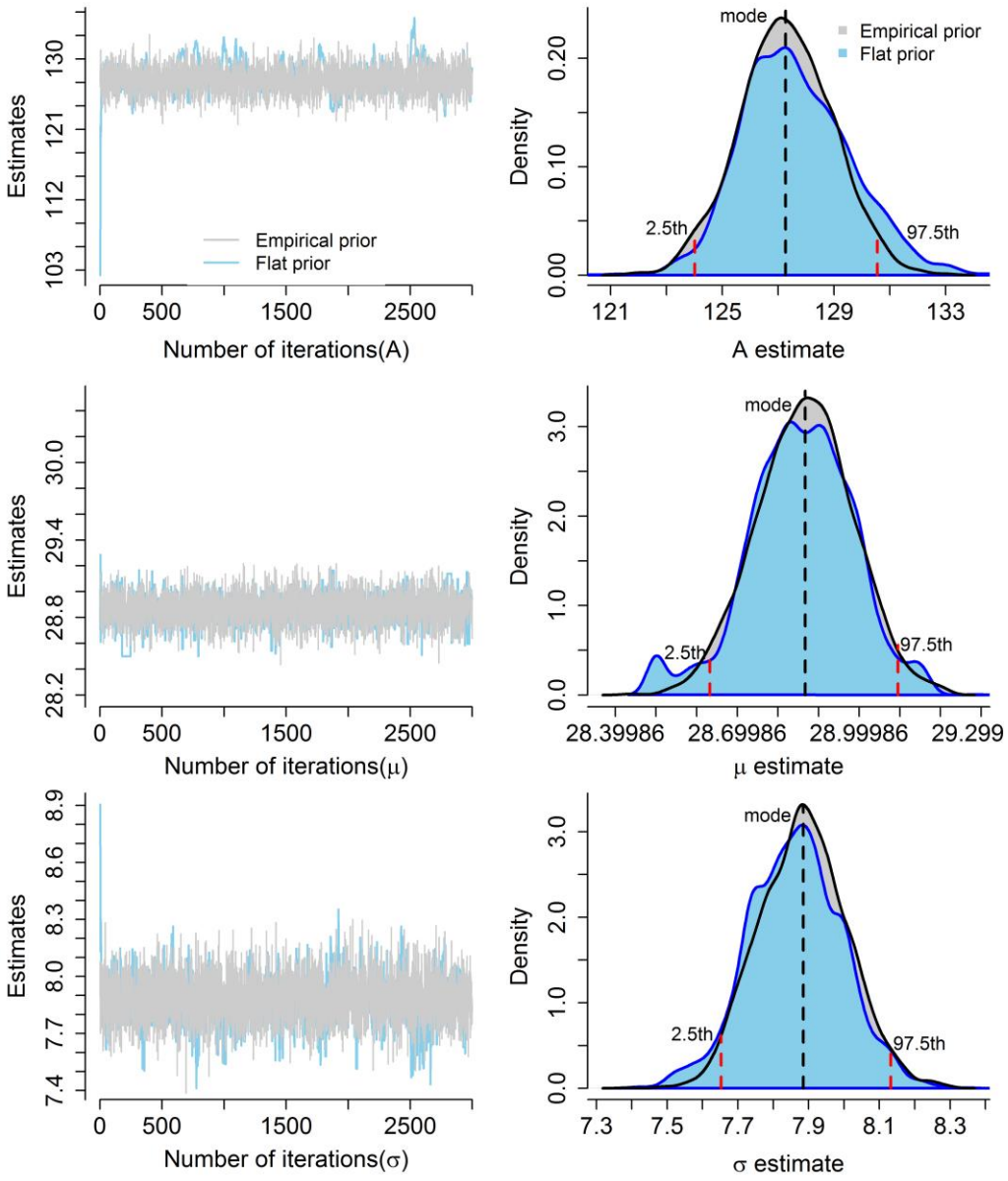


Figure III-10. An example of parameter estimates derived from the Gaussian model within the Bayesian framework using the flat priors and empirical priors with the same number of iterations. Left: The trace plot of model parameters A, u, δ using the flat priors

and the empirical priors. Right: The probability density distribution of estimated parameters A, u, δ using the flat priors and the empirical priors.

Table III-4. Average processing time for a single waveform with different number of peaks.

Number of components per waveform	Number of waveforms	Average time (flat priors, seconds)	Average time (empirical priors, seconds)
1	95	2.43	1.91
2	21	12.40	10.49
3	10	67.38	45.36
4	6	235.60	114.70
5	4	364.65	162.00
6	2	462.60	267.00
7	2	511.20	276.90

From Table III-4, a general trend has emerged that more time is taken to process the waveform with a larger number of components no matter which priors were used. Interestingly, there was an abrupt rise of time cost when the waveform's number of components became four. The comparison between the flat priors and the empirical priors demonstrated that there was little difference of time cost when the waveform was relatively simple with one or two waveform component(s). However, the computation time using the empirical priors was much shorter than when using the flat priors for complex waveforms. This reduction of time could make a substantial difference when millions of waveforms need to process and the efficiency of using the empirical priors will become more evident. Without the consideration of the computation cost and time, the impact of different priors on the parameter estimates and the performance of the Bayesian models was negligible.

3.3.3.2 Point cloud uncertainty

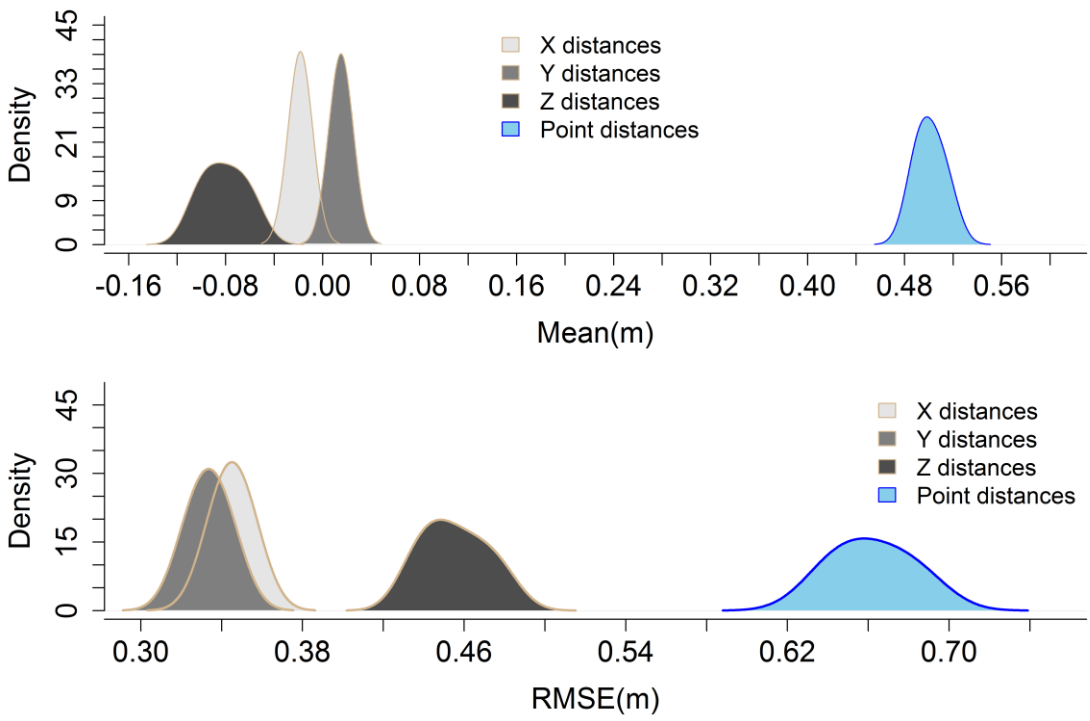


Figure III-11. Uncertainty of average distance (Mean) and corresponding RMSE between reference point cloud dataset and waveform-based LiDAR quantile point cloud datasets using the Bayesian decomposition method at the SJER site

The PDFs of descriptive statistics (Mean and RMSE) for point distances (sky blue), X, Y and Z distances (gray) through comparing the waveform-based LiDAR quantile point cloud datasets with the DR dataset are displayed in Figure III-11.

The uncertainty of average bias and corresponding RMSE for point distances were larger than the other three individual coordinates' descriptive statistics when examining the mode and range of corresponding distribution. For example, the mode of average distances of the X, Y and Z centered around -0.02, 0.02 and -0.08 m, respectively. The counterpart of point distances was much larger which centered around 0.51 m with more flat distribution. The RMSE demonstrated a similar pattern that the distribution of point distances was less compact than other three coordinates' distribution with a larger mode. Interestingly, there was no significant difference between the distribution of X distances

and Y distances, while the distributions of Z distances varied with the X and Y distances with larger absolute mode and uncertainty.

3.3.3.3 Surface model uncertainty

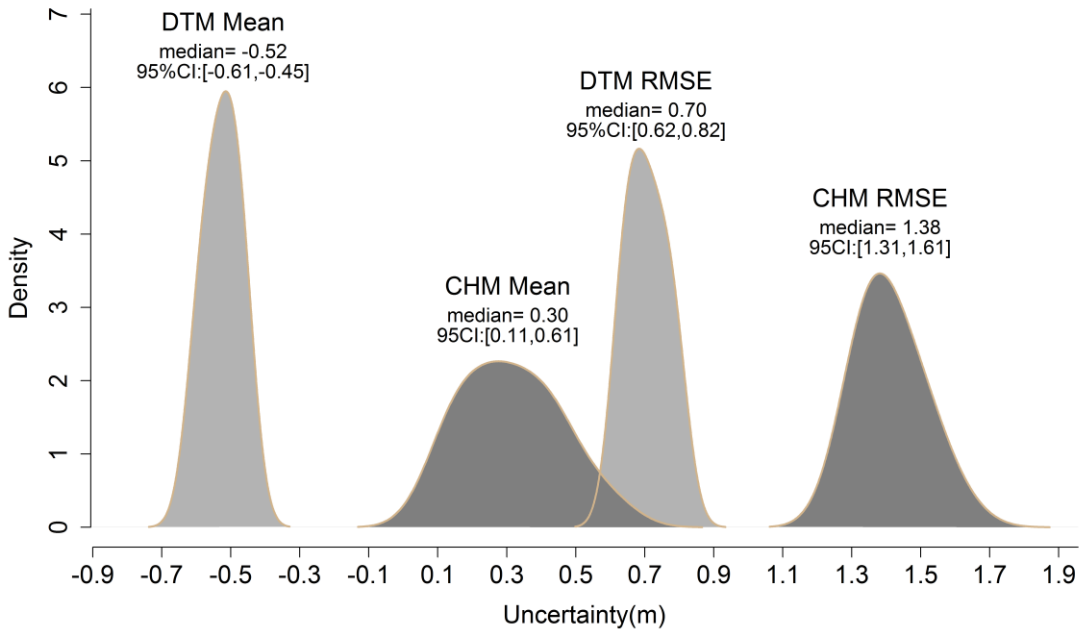


Figure III-12. Uncertainty of the average difference and RMSE between waveform-based surface models (the DTM and CHM) and reference surface models using Bayesian decomposition method at the SJER site. CI: Credible interval.

The uncertainty of surface models derived from these quantile point clouds using the Bayesian decomposition method is demonstrated in Figure III-12. The distribution of the CHM's mean bias and RMSE (dark gray) tended to be wider than DTM counterparts. For instance, the estimated RMSE for the CHM was 1.38 m and 95% CI ranged from 1.31 to 1.61 m, both of which were larger than the RMSE of the DTM. The distribution of the average difference for DTM was likewise smaller than CHM's average difference from the perspective of 95% CI. Therefore, it was evident that the uncertainty of DTM was smaller than the uncertainty of CHM in terms of their corresponding distribution of the average bias and RMSE.

Compared to the point clouds' uncertainty (Figure III-11), the surface models yielded larger uncertainty with wider CIs, which could represent the propagation of true estimation error along the processing steps. Using the RMSE of CHM as an example, it demonstrated a larger median and more flat pattern distribution than point clouds' RMSE (Figure III-11).

3.3.3.4 Uncertainty of accuracy assessment with field data

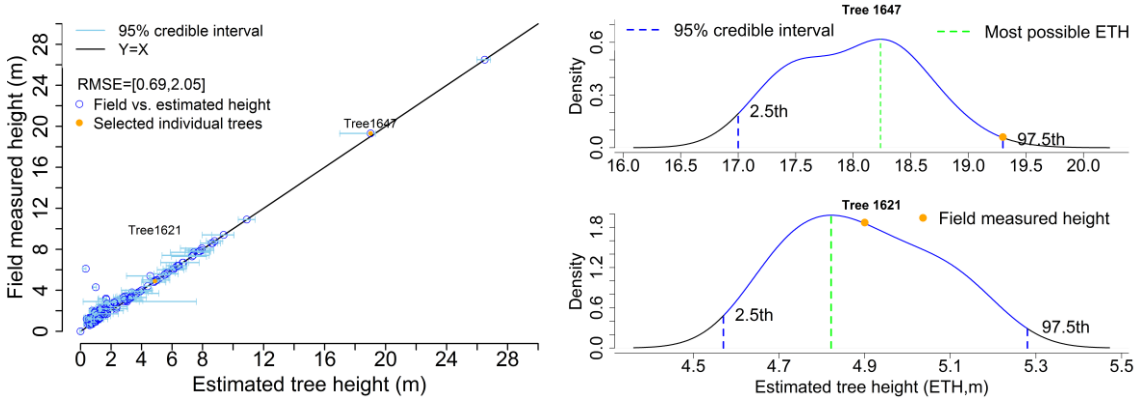


Figure III-13. Uncertainty of the individual trees' height as obtained from the Bayesian decomposition method vs. field-measured tree height at the SJER study site. The left panel refers to all individual trees' 95% credible interval (sky blue), and the right to the distribution of two possible estimated tree heights with 95% credible interval (blue).

Estimated tree height's uncertainty yielded from the Bayesian decomposition method at individual tree level against the field-measured data is displayed in Figure III-13. It was observed that data points (blue circle) of estimated tree height (ETH) versus field-measured tree height mostly followed tightly around Y=X line that indicated the Bayesian decomposition method can accurately capture the height of these individual trees. In addition, almost all trees' 95% CIs of ETH (145 out 151 individual trees), as indicated by the horizontal error bar (sky blue), were intercepted by Y=X line. The uncertainty of ETH's RMSE confirmed good agreement with the field-measured height that resulted in the uncertainty of RMSE ranges from 0.69 to 2.05 m. To avoid the proliferation of figures, two individual trees (Tree 1621 and Tree1647, orange) were used as examples to further demonstrate final results of tree height estimation using the

Bayesian decomposition method as shown in the right panel of Figure III-13. This figure provided more than just one value of ETH that generally obtained from the deterministic methods, and yielded more informative estimates of tree height with probability for the individual tree.

3.4 Discussion

3.4.1 Model reasonableness

Several potential models such as the Generalized Gaussian, Burr and Weibull models have been proposed to fit the airborne LiDAR waveforms (Mallet et al., 2009), while most of them are unrealistic models which may violate theoretical considerations or misinterpret their real-world's physical meanings. Thus, the choice of model is crucial for the subsequent FW LiDAR data processing and outcome interpretation.

The results of model reasonableness in the present study justify that the Gaussian model is more suitable for reconstructing the waveform LiDAR signals in terms of efficiency, physical interpretation and uncertainty. Three representative models suitable for FW LiDAR data processing including the Weibull, Adaptive Gaussian, and Gaussian models were explored. In our case, the Adaptive Gaussian model better fitted waveforms compared to the Gaussian model. However, the Adaptive Gaussian model is prone to overfit the model and incorrectly considers data noise as real waveform signals by adjusting the rate parameter to pursue the smallest residual of model fitting. The Weibull model is more difficult to constrain than the other two models with larger CIs or uncertainty. Moreover, there is no explicit way to explain the physical meaning of the corresponding parameters. Hence, the Gaussian model is a suitable trade-off between accuracy and meaningful solution to extract useful information form FW LiDAR data.

Simultaneously, the quantification of uncertainty using these models demonstrates the attractive features of the Bayesian method and allows us to interpret results in a more natural way with posterior distribution of estimates. Moreover, we can explicitly trace the uncertainty inherent in any inference and monitor the uncertainty propagation through the use of quantiles for different parameters, quantile datasets and models (Figure III-3).

3.4.2 Performance evaluation

Point cloud. The calibration and comparison conducted on the point clouds eliminate the reliance on the underlying surface derived from local parametric estimates. Additionally, we compare every point of the waveform-based point cloud with the corresponding reference point cloud instead of mainly comparing a subset of the total points in the specified grid cells when the area based method (DTM or CHM) is used. This can avoid the error introduced by interpolation and poor choice of gridding size, and ultimately make the comparisons become more convincing and comprehensive. The C2C results show the Bayesian decomposition method is applicable for extracting information from FW LiDAR data to characterize the vegetation structures. The contribution made by the Z direction to the point distances is much larger than the X and Y directions. A reasonable explanation for the difference is that pulse direction vector of the Z direction (~ -0.15 m/ns) is much larger than the X and Y directions (~ -0.02 and -0.01 m/ns). As a consequence, the relative change in the Z direction is much higher than in the X and Y directions given the same change of time. This may also suggest that FW LiDAR data's horizontal precision is higher than the vertical precision. The uncertainty of the point cloud (Figure III-11) further substantiates this finding with larger average distance and RMSE in the Z direction.

Overall, the C2C method's results support that the four waveform processing methods can generate relatively satisfactory results with all methods' RMSE values being less than 0.93 m (Table III-1). However, a closer examination of the point clouds comparison results reveal that the decomposition method (the DD and Bayesian decomposition) may generate more accurate point clouds than the combined deconvolution and decomposition method (the Gold and RL) in terms of the descriptive statistics such as the MD and RMSE. Specifically, the DD method is prone to outperform other three methods with slightly smaller average distance and corresponding RMSE. This conclusion backs up the results using the area based method of Zhou's study (2017) that the DD method potentially yield more accurate DTMs and CHMs than the Gold and RL methods.

Individual tree metric. The C2C comparison yields a general view about the waveform processing methods on a large scale. Examining metrics derived from different methods at the individual tree level further reveal advantages and disadvantages of these methods. Most of the individual trees derived from waveform LiDAR data can obtain more points - by as many as 80% of trees (16/21) at the SJER1 study region - compared with the corresponding DR LiDAR data. Two illustrative examples of individual tree's point distribution at different height bins, percentile heights and canopy point density depict a similar trend that FW LiDAR data are capable of extracting more points at the mid-story of the vegetation. Surprisingly, by comparing the DR LiDAR data results, FW processing based methods are less likely to detect ground points that maybe the disadvantage of using FW LiDAR data.

Field data calibration. Various factors could affect the field calibration results, such as the number of sample plots, the accuracy of measured instruments and the mis-registration error between LiDAR data and field plots (Zhao et al., 2011). Although field calibration is necessary for LiDAR applications, caution should be exercised when using calibration results, given the survey error was inherently unavoidable and ubiquitous. Table III-3 demonstrates that DR LiDAR data (RMSE<= 1.1 m) can be an alternative for field-measured data when they are not available. This conclusion is also consistent with the previous studies' claims that DR LiDAR data can accurately measure the tree height (Chen, 2007). In addition, the uncertainty of ETH (Figure III-13) is generated to indicate the true magnitude of estimation error and enhance the credibility of field calibration results by reducing the error introduced by the field measurements. For the Bayesian decomposition method, the calibration results are not just represented by one RMSE value like the DD, Gold and RL methods yielded (Table III-3), but all possible RMSE values with probability are provided. This situation renders an intuitive and reasonable way to address the real-world model calibration problem.

3.4.3 Uncertainty analysis

Quantification of uncertainty is one of the notable advantages for adopting the Bayesian method. The uncertainty is retained throughout all processing steps including

parameter estimates, geolocation transformation and surface model generation, a full disclosure of uncertainty of these processing steps is crucial for the result interpretation and the potential applications. In the present study, a thorough uncertainty quantification method is conducted on parameters, quantile point clouds and surface models (DTMs and CHMs) to ensure the quality of the results. The Bayesian approach gives us all possible estimates with probability for unknown parameters (Figure III-3) instead of just providing a single value within a frequentist framework. This enables the researchers to view where the most probable locations of the waveforms to represent the illuminated object along the pulse line. Furthermore, the probability distribution of the estimates also essentially precludes non-uniqueness problem, which is common in the waveform decomposition. Actually, this problem has been faced by the users who employed advanced statistical models to tackle inversion problems in various fields (Gouveia and Scales, 1998; Oh and Kwon, 2001).

The Bayesian method has been subjected to criticism for its subjectivity by introducing prior information. In the present study, the results of using different priors showed that this subjectivity can be overcome by using the flat priors (non-informative priors). The nearly identical performance as empirical priors can be achieved at the expense of more computation cost and time using the flat priors, which agrees well with Ellison's conclusion (Ellison, 2004). While this conclusion is not consistent with Denham's study (Denham et al., 2009), using the empirical priors tends to increase the precision of the parameters. There are factors that contribute to this inconsistency, such as data noise, setting of non-informative prior and the main model used in the simulation. However, all results demonstrate that the Bayesian approach is capable of data analysis such as parameters extraction from waveforms with reasonable accuracies. To reach the same performance of the empirical priors, the computation time using the flat priors is markedly longer (Table III-4). This indirectly reflects that the empirical priors are capable of reducing the complexities of Bayesian decomposition method with accurate outputs.

The Bayesian decomposition method may appear to require more computation cost than the other three methods, however, it can overcome the parameter initialization

problem of the DD method, and avoids the step of parameters optimization for deconvolution when the combined deconvolution and decomposition method is used. Results of this study demonstrate the potential and advantages of using the Bayesian approach to characterize uncertainty of parameter estimates. Moreover, it allows researchers to trace the propagated error and uncertainty explicitly through PDFs of corresponding evaluation criteria from parameters to points, point clouds and surface models (Fig 3). Assuredly, the unsuccessful Bayesian decomposition may occur when the waveform is extremely noisy or irregular. Assigning the number of Gaussian components (n) as a random variable is expected to be one potential solution which could be one aspect of further research.

The uncertainty results for the quantile point cloud datasets are represented by a probability distribution, which can characterize the variability, extent of average distance and RMSE in a probabilistic sense to assess the performance of the Bayesian method. As expected, the horizontal precision of FW LiDAR data is better than their vertical precision and the main uncertainty of the point clouds come from the vertical direction.

The surface model's uncertainty results reveal that the DTM has smaller variance and uncertainty than the CHM, which yields a larger average bias and a wider range of the RMSE (Figure III-12). Various factors can degrade the accuracy of the DTM, such as sampling size, point density, terrain conditions and processing methods (Vincent et al., 2012). To some extent, the quality of the DTM significantly affected the CHM quality. Consequently, larger error or uncertainty of CHM accrued might mainly due to the fact that CHM was generated by subtracting DTM from Digital surface model (DSM) and an additional error was likely to be introduced with this step. This kind of uncertainty propagation was also observed through comparing the average bias and corresponding RMSE of the point clouds (Figure III-11) with the surface models (Figure III-12). More specifically, the point clouds' RMSE varied from 0.62 to 0.71 m, while the uncertainty of the RMSE for the waveform-based DTM and CHM were 0.62 - 0.80 m and 1.31 - 1.61 m, respectively. It is evident that the error and uncertainty are increasing along the processing steps and an additional processing step brings more error and uncertainty. The

uncertainty analysis captures the fact the error introduced into the estimation steps and provide more informative information which consequently gives us more confidence to interpret and apply results to real-world problems such as tree species identification.

Actually, most of the remote sensing applications such as variable extraction and biomass estimation using various sources of remote sensing data are inadequately addressed or overlooked the uncertainty analysis of their results (McRoberts et al., 2010; Zhao et al., 2011). The Bayesian concept or approach can be transplanted to these applications to quantify the estimate of error or uncertainty by employing statistical inference of posterior distribution for the subjects of interest. At the current stage, the Bayesian approach appears complicated which requires users to have knowledge about the concepts and procedures, and needs extensive computation time (De Lannoy et al., 2014) for MCMC sampling. These aspects have hindered the broad applications of Bayesian approaches, however, the advances in computation and development of generalized tool such as BUGS, JAGS and Stan will likely contribute to the popularity of Bayesian models in the foreseeable further.

3.5 Conclusion

This paper has incorporated the Bayesian concept with waveform decomposition to develop an innovative method to extract information from FW LiDAR data and conduct a thorough uncertainty analysis along the processing steps. Built upon the deterministic method for waveform decomposition of Zhou et al. (2017), a more comprehensive exploration of the extant methods and the proposed Bayesian method are conducted at the point, point cloud and surface model levels.

The Bayesian method contributes to the waveform decomposition in several ways. As demonstrated, the solutions of decomposition using the Bayesian method are represented by a probability distribution over parameter estimates instead of producing a single value for each parameter using the deterministic approach. Additionally, it permits the users to interpret the results in a probabilistic sense that is stable enough to provide a feasible solution to the decomposition problem. Moreover, the adoption of Bayesian analytics generates a systematic and transparent knowledge learning framework to

estimate uncertainty emerging from parameters, points, the point cloud and the surface model via the use of quantiles of the probability distribution. Meanwhile, the uncertainty propagation can be explicitly traced and observed from data to the parameter estimate, the point, the point cloud and the surface model (Figure III-3). We also explored the non-uniqueness problems of waveform decomposition within Bayesian framework to reduce the theoretical error from the model itself and justify the reasonableness of using the Gaussian model for FW LiDAR data decomposition in terms of uncertainty, physical meaning and processing efficiency. The superior method of FW LiDAR data processing varies from the perspective of different criteria and waveform-derived products. Results of point cloud comparisons demonstrate that the Bayesian decomposition method can achieve comparable accuracy as the DD, Gold and RL methods. Results from the individual tree level highlight that FW LiDAR data can characterize more detail of the mid-story of vegetation with more dense points. The combined deconvolution and decomposition method (the Gold and RL approaches) outperforms the decomposition method (DD and Bayesian decomposition) in terms of the surface model's results. In addition, field data for calibration results using the Bayesian method provide a reasonable way to reduce the effect of field measurement error on the model calibration. Further, the Bayesian method is expected to become more efficient and user-friendly with the aid of computational advances and convenient implementation tools. In addition, more research efforts are still needed to apply FW LiDAR data for vegetation studies, such as tree species identification and biomass estimation over extensive regions.

3.6 Appendix

3.6.1 Model implementation & model structures

The models were built on the Stan inference engine. It consists of four blocks: variables declarations, parameter statements, transformed parameters and model blocks The data block is functional as input data declaration; the parameters block is to introduce all unknown parameters of interest; the transformed parameters block is mainly for the conversion of data and parameters to the readable way, and the model block is to compute the log posterior density. Stan first translated a model into C++ and then compiled the

code for each waveform. Different waveforms may have a different number of Gaussian components, which generates a different posterior distribution and model for each waveform. The challenge is that the model always changes based on the number of Gaussian components which requires users to compile a new model for each waveform. Meanwhile, a new dynamic link library (DLL) was generated to store the model. This would exceed the maximum number of DLLs can be loaded on a computer after the model fitting a certain amount of waveforms. In order to solve this problem, the *update* function of brms was used to avoid recompilation issues. This function requires the structure of the input model or the posterior distribution should be exactly the same format. Therefore, we grouped the waveforms with the same number of peaks, and then employed the *update* function for fitting these waveforms. This strategy could save the compilation time and make the processing much more efficient. In addition, the parallel computing is also automatically implemented in brms package by specifying the number of clusters and chain. In this analysis, we assigned them both as 2. All of these components in brms package make HMC converge much faster to a target distribution.

3.6.2 Model convergence

The model convergence is measured with $\hat{R}$ and is computed as follows:

$$\widehat{Var}(\theta) = \left(1 - \frac{1}{l}\right) W + \frac{1}{l} B \quad \text{(III-10)}$$

$$\hat{R} = \sqrt{\frac{\widehat{Var}(\theta)}{W}} \quad \text{(III-11)}$$

where W is the mean variances of stationary distribution for each chain, B is the variance of stationary distribution at the between-chain level, l is the number of draws in each chain, and $\widehat{Var}(\theta)$ is the variance of the stationary distribution as a weighted average of W and B. Here, the chain of $\hat{R}$ was the split chain that discarded the burn-in iterations.

3.6.3 Model efficiency

Recently, the inference engine Stan (Gelman et al., 2015) has been introduced to implement HMC sampling. Stan employs a reverse mode automatic differentiation rather than a numerical differentiation to compute the gradient (Griewank and Walther, 2008).

Furthermore, the No-U-Turn Sampler (NUTS), a variant of HMC, was used to automatically tune two parameters in the leapfrog method. Specifically, the step length L is achieved by means of a recursive algorithm with doubling procedure devised by Neal (Neal, 2003) for slice sampling, and step size ϵ via an adaptation of dual averaging algorithm of Nesterov (Nesterov, 2009).

These enable NUTS to run more efficiently than other MCMC algorithms and to become desirable for those who have little experience of tuning HMC without user intervention. Figure III-14 displays the relationship between the log posterior of the model (x-axis) and acceptance rate (y-axis) using NUTS samplers to demonstrate the efficiency of HMC or NUTS. Our experiment demonstrates that the sampling acceptance rate of NUTS is much higher (around 85 - 95%) than the desired acceptance rate of Metropolis algorithm which generally ranges from 23% to 50% (Roberts et al., 1997) (Figure III-14).

The exploration of these waveforms also demonstrated that assigning the proper number of and total iterations of MCMC was critical to the model efficiency. The above sample waveforms were also used to explore optimized combinations of the number of burn-in samples and total iterations. To balance the processing time and accuracy, the optimization of these parameters was conducted. $\hat{R} < 1.1$ was employed to judge the number of total iterations and samples were enough to obtain the acceptable results. The summary of these parameter combinations after optimization was shown in Table III-5. The number of chains used in this analysis was 2. To save the computation time and reduce the autocorrelation of draws in MCMC simulations, we saved every third iteration of each chain by thinning posterior samples.

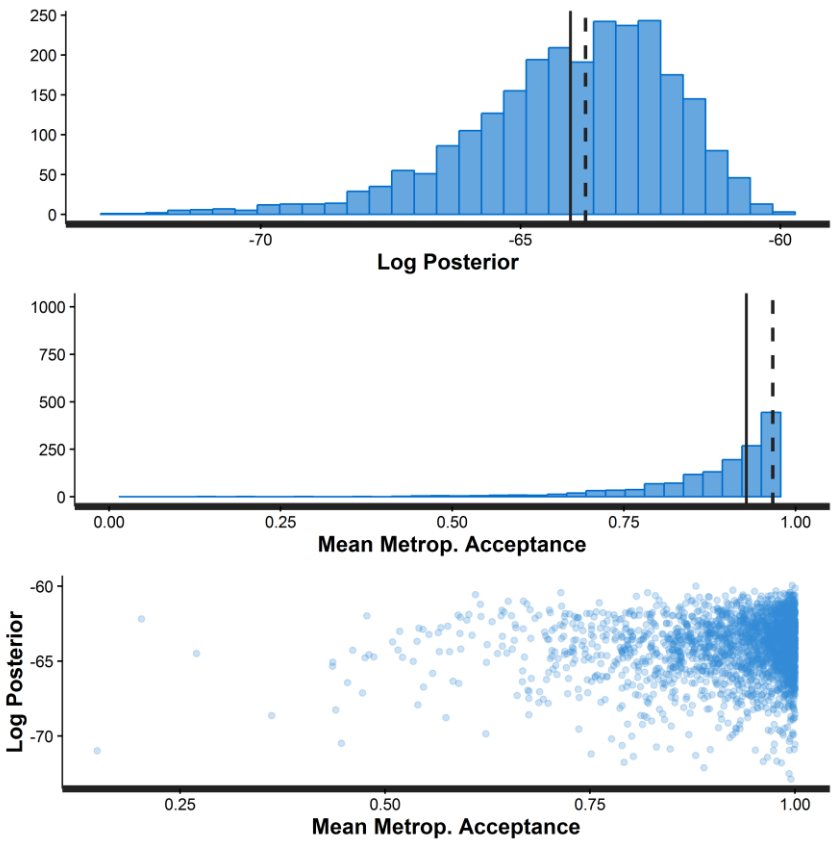


Figure III-14. An example of histogram for log posterior and final Metropolis acceptance rate, and distribution of log posterior vs. Metropolis acceptance rate using HMC algorithm

Table III-5. The final parameters of MCMC simulation for different waveform components

Number of component per waveform	Number of total iterations	Number of burn-in
1	9,000	2,000
2	10,000	2,500
3	12,000	3,000
4	15,000	4,000
5	18,000	5,000
6	20,000	6,000
7	22,000	7,000
>8	25,000	8,000

CHAPTER IV
BAYESIAN AND MACHINE LEARNING METHODS FOR TREE SPECIES CLASSIFICATION WITH WAVEFORM LIDAR DATA

A plethora of information contained in full-waveform (FW) LiDAR data offers prospects for characterizing vegetation structures. This study aims to investigate the capacity of FW LiDAR data alone for tree species identification through the integration of waveform metrics with machine learning methods and Bayesian inference. Specifically, we first conducted automatic tree segmentation based on the waveform-based canopy height model (CHM) using the local maximal (LM), watershed algorithms and the combination of the LM and watershed algorithms. Subsequently, the Random forests (RF) and Conditional inference forests (CF) models were employed to identify important individual tree-level waveform metrics derived from the waveform-based point cloud, raw waveforms and composite waveforms. Further, we discriminated tree and shrub species by the RF, CF and Bayesian methods using identified important waveform metrics. Results of the tree segmentation demonstrate that the combination of the LM and watershed algorithms outperforms other algorithms for delineating individual tree crowns. The accuracy of delineated tree crowns greatly affects the waveform metrics' effectiveness for feature selection. The CF model overcomes waveform metrics selection bias caused by the RF model which violates the implicit null hypothesis and favors correlated metrics, and enhances the accuracy of subsequent classification. We also found that composite waveforms are more informative than raw waveforms and waveform-based point cloud for characterizing tree species in our study area. Both machine learning methods (the RF and CF) and the Bayesian method generate satisfactory overall accuracy (72.7% for the RF, 73.8% for the CF and 84.4% for the Bayesian method), and the Bayesian method slightly outperforms the other two methods. However, these methods suffer from low individual classification accuracy for the blue oak, which is prone to being misclassified as interior live oak due to similar characteristics of the blue oak and

interior live oak and insufficient accuracy of field data. Uncertainty estimates from the Bayesian method compensate this downside by providing classification results in a probabilistic sense and rendering users with more confidence in interpreting and applying classification results to real-world tasks such as forest inventory. Overall, this study highlights the recommendation of the CF method for feature selection and implies that the Bayesian method can be a superior alternative to machining learning methods.

4.1 Introduction

Successful tree species classification with remote sensing data is of considerable value to forest inventory and ecosystem management (Schlerf et al., 2005; Treitz and Howarth, 2000). Over the past decade, the emergence of Light Detection and Ranging (LiDAR) techniques has facilitated the development of tree species identification in forest ecosystems (Heinzel and Koch, 2011; Holmgren and Persson, 2004). Especially with advances of waveform digitization in the commercial LiDAR systems, this state-of-the-art technology renders promising potential for reconstructing the structure of objects such as trees (Hollaus et al., 2009b; Reitberger et al., 2009). Full waveform (FW) LiDAR data enable users to "see" the whole recording process instead of a black box as discrete-return (DR) LiDAR data which internally interpret the pulse by the system itself (Reitberger et al., 2008; Zhang et al., 2015). Additional information such as geometric and reflection characteristics along the pulse line can also be obtained through recording the whole pulses. Moreover, different tree species generally have different shapes of tree crowns and internal structures (Heinzel and Koch, 2011), which gives rise to a different number of waveform components (peaks) and intensities. Such advantages unarguably suggest that FW LiDAR data are theoretically well-suited to classify tree species from the tree morphology perspective, which, on the other hand, connotes the pressing needs of developing methods to extract useful FW information and applying them to vegetation studies.

For the past decade, most attention has been paid to applying DR LiDAR data and optical images to discriminate tree species. Previous studies have demonstrated that DR LiDAR data alone (Holmgren and Persson, 2004; Vaughn et al., 2012) or in conjunction

with ancillary data such as multispectral imagery (Heinzel et al., 2008; Holmgren et al., 2008; Ke et al., 2010; Leckie et al., 2003), and hyperspectral imagery (Jones et al., 2010; Zhang and Qiu, 2012) can classify tree species by using variables such as percentile heights and spectral features extracted from these data. Segmentation of individual tree crowns is the general preprocessing step of tree species identification. The accuracy of segmentation is crucial for precisely extracting other tree structural attributes such as crown width, tree height, basal area and subsequent tree species classification (Chen et al., 2007; Koch et al., 2006; Li et al., 2012; Popescu and Wynne, 2004). Most of these segmentations are based on the LiDAR-derived Canopy Height Model (CHM) using different algorithms such as the LM filtering algorithm (Popescu et al., 2002), watershed algorithms (Chen et al., 2006) and pouring algorithm with empirical geometric shape of trees (Koch et al., 2006). Other studies segment tree crowns directly from the point cloud (Li et al., 2012; Morsdorf et al., 2003; Wang et al., 2008). The accuracy of tree crown segmentation is influenced by many factors such as forest types and point density (Vauhkonen et al., 2011), while the tree segmentation method has been proven to be the primary factor in determining the accuracy of individual tree crown extraction (Kaartinen et al., 2012).

The segmentation step is also pivotal to delineate individual tree crowns using FW LiDAR data, but this procedure falls outside the scope of this research, and thus, no detailed discussion of tree segmentation algorithms is provided. Briefly, the CHM derived from DR LiDAR or FW LiDAR data is mainly employed to conduct individual tree crown segmentation for FW LiDAR related studies (Reitberger et al., 2008; Reitberger et al., 2009; Yao et al., 2012). Specific examples of emerging FW LiDAR data for tree species identification mainly use variables extracted from the waveform decomposition, such as echo width, amplitude, backscatter cross section, inner point density and the number of peaks per waveforms (Hollaus et al., 2009a; Reitberger et al., 2008; Vaughn et al., 2012; Yao et al., 2012; Yu et al., 2014). The main advantage of the decomposition method is that the general view of shape for each waveform can be obtained by deriving additional information such as the echo width and amplitude which provides promising prospects

for their application in tree species classification (Hollaus et al., 2009a; Reitberger et al., 2008). However, the rich information contained in the waveforms may automatically be lost by just using metrics from the decomposition. Several studies have explored the variables directly from FW LiDAR data instead of the corresponding point cloud, named FW metrics, to investigate their potential for tree species discrimination. The FW LiDAR metrics are firstly utilized in the large footprint FW LiDAR data to characterize the vegetation structures and land covers (Drake et al., 2002; Harding, 2005; Hermosilla et al., 2014a), and these metrics have been transplanted to small footprint FW LiDAR data for forest studies such as estimating forest structure parameters (Hermosilla et al., 2014b) and identifying tree species (Cao et al., 2016; Yu et al., 2014) in recent years. Few studies have, however, investigated the waveform metrics derived from composite waveforms (Hermosilla et al., 2014b) which contained absolute vertical information of intensities and overcome the possible negative effects caused by the off-nadir angle of flight.

Regarding the tree species classification with FW LiDAR data, the common challenges are high dimensional variables with sparse field data, known as “small n large p” problem, and inherent correlations among variables further require users to conduct feature selection to obtain useful variables before the classification. Traditional ways to maintain parsimonious metrics and avoid multicollinearity are achieved with statistical measures such as variance inflation factor and Akaike information criterion. However, the high dimensionalities resulting from preserving as much information as possible in the raw data and model-specific assumptions preclude their potentials for precisely determining useful metrics for subsequent classification. Advanced statistical models and algorithms with increasing computational resources have enabled researchers to construct complex and high-structured models that are previously considered intractable to tackle these problems. For example, machine learning (ML) techniques have been proven to be a valuable tool capable of solving these problems in various domains through capturing the implicit and complex relationship between variables and the subject of interest (Zhao et al., 2011), especially for the Random forests (RF) method which has gained popularity in measuring variable importance in various scientific fields (Belgiu and Drăguţ, 2016;

Cao et al., 2016; Strobl et al., 2007). However, feature selection results are misleading when the potential variables vary in scale or number of categories (Strobl et al., 2007), and correlated predictors are preferred to obtain more importance because the scheme of the RF's permutation importance is constructed (Strobl et al., 2009a).

In the tree species classification domain, several ML techniques such as linear discriminant analysis (Heinzel and Koch, 2011), artificial neural network, support vector machine (SVM) (Vaughn et al., 2012), and random forest (RF) (Cao et al., 2016; Yu et al., 2014) have been employed to demonstrate their advantages over classical methods in tree species identification with a large suite of FW metrics. The limited number of available field sample data as training data is another concern for applying these methods to achieve high accuracy results (Acevedo et al., 2009). Generally, the training data obtained through field measurements in LiDAR remote sensing are scarce, which require users to meticulously construct the model or classifier by making the best use of available data. Furthermore, these machine learning methods are mainly based on the deterministic concept that can't capture the inherent errors and uncertainty of classification in a probabilistic sense and may further degrade the methods' predictive accuracy.

Recent developments in Markov Chain Monte Carlo (MCMC) and advances in computation have contributed to the emergence of the Bayesian approach as an effective tool to meet these challenges. Through the Bayesian approach, the belief of the subject of interest is updated with new information and prior knowledge of model parameters, which renders an intuitive way to interpret the results and errors with probability distribution. Even for the sparse training data, useful posteriors and model results are still achievable with some reasonably informed priors for model parameters. Consequently, the Bayesian method is capable of capturing uncertainties related to parameters and models which further enhances the credibility of model prediction. In LiDAR remote sensing, Bayesian methods have been employed to decompose waveform LiDAR (Zhou and Popescu, 2017) and predict forest variables and biomass (Babcock et al., 2016; Finley et al., 2013; Patenaude et al., 2008), yet the capacity of Bayesian methods for tree species classification using FW LiDAR data is rarely exploited.

The overall goal of this study is to integrate advanced statistical methods such as the RF, Conditional inference forest (CF), and Bayesian inference with FW metrics to conduct tree species identification using FW LiDAR data alone. In particular, three specific objectives are formulated: (1) to explore tree segmentation using a waveform-based CHM with the combination of the LM and watershed algorithms; (2) to investigate significant metrics derived from waveform-based point cloud, composite waveforms and raw waveforms using traditional RF and CF methods, and (3) to introduce the Bayesian method to perform tree species classification and compare it to the RF and CF methods, and further quantify the uncertainty of classification with the Bayesian method. Two main innovative aspects emerge from this study: (1) introducing the CF model to conduct FW metrics selection to cope with "small n large p" problem and conquer the inherent bias of the RF model, and (2) integrating Bayesian inference with FW metrics to classify tree species with uncertainty using FW LiDAR data alone. The framework and methods developed in the present study can be easily transferred to vegetation characterization and biomass mapping.

4.2 Materials and Methods

4.2.1 Study site

Our study site is an ecosystem research experimental area, named the San Joaquin Experimental Range (SJER, UTM Zone 11N, Easting 257,600, Northing 4,109,300), which is situated at the foothills of Sierra Nevada Mountains, about 32 km north of Fresno, California (Figure IV-1). Within in the SJER, the vegetation is mainly composed of interior live oak *(Quercus wislizeni),* blue oak (*Quercus douglasii*), gray pine (*Pinus sabiniana*) and scattered shrubs with a nearly continuous cover of herbaceous plants. The whole area is characterized by the complex topography including coarse, large hills and valleys with elevation ranging from 210 to 521 m above sea level with a mean elevation of 366 m.

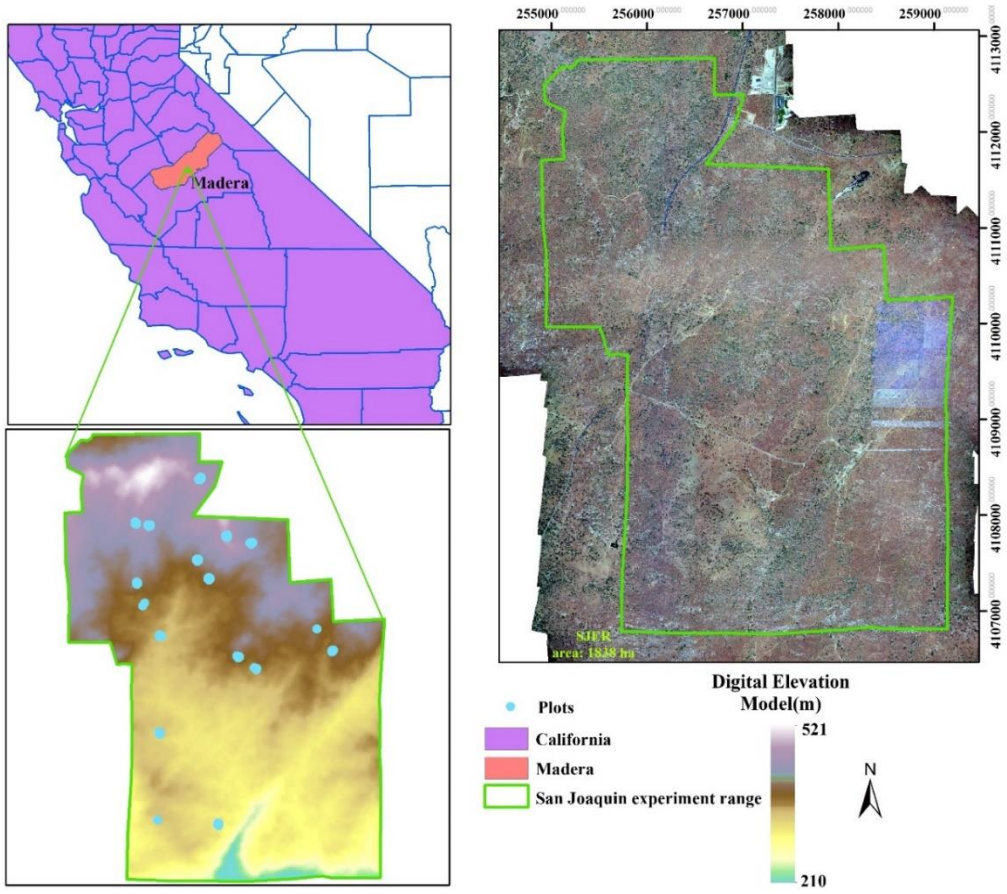


Figure IV-1. An overview of study area with filed-measured plots in the San Joaquin Experimental Range (SJER).

4.2.2 Data

4.2.2.1 LiDAR data

FW LiDAR data of the study site were acquired during leaf-on season in June 2013 with an Optech Gemini instrument flying at approximately 1,000 m above the ground level. The flight campaign was conducted by the National Ecological Observatory Network (NEON) Airborne Observation Platform, which provided FW LiDAR data with a 0.8 m diameter footprint, and a spacing of about 0.524 m in the across-track direction and 0.5 m in the along-track direction. Compared to DR LiDAR data, the reported range and vertical accuracies of FW LiDAR are 0.06 - 0.28 m and 0.15 m, respectively. The

entire study area was covered by two perpendicular direction flight lines, with 12 and 19 flight lines in an east-west direction and north-south direction, respectively. Major technical specifications of the flight campaign have been reported in in the study of Zhou et al. (2017).

In this study, two pieces of waveform related information were used: a return pulse with 500 time bins and corresponding reference geolocations. The temporal resolution of the return pulse is 1 ns and the zero-padding is applied to non-recorded values of the return pulse to keep the length of waveforms constant. For the reference geolocation of the corresponding waveform, 8 basic geolocation information attributes associated with the waveform were obtained. Specifically, the attributes consist of the Easting of first return x_0 (m), the Northing of first return y_0 (m), the height of first return z_0 (m), the pulse direction vectors (dx (m), dy (m) and dz (m)), the outgoing pulse reference bin location (leading edge 50% point of the outgoing pulse), and first return reference bin location (leading edge 50% point of the first return).

4.2.2.2 Reference data

Field-measured data were collected across 17 plots in the study site during June 2013. An overview of the plots is displayed in Figure IV-1 with blue points. The locations of the plots were established by the NEON's Field Sentinel Unit for measuring long-term plant, insect and soil related properties. According to the protocol of the NEON Terrestrial Observation System, the size of the plot is a square region with 20 ×20 m. A various number of trees were measured in each plot, and in total 181 trees and shrubs were observed. Several key vegetation structure variables were recorded and only locations (Eating and Northing) and tree species were used in this study.

4.2.3 Methods

The first step of automated tree species classification using LiDAR data is to segment individual tree crowns. While exploring the capability of FW LiDAR data to segment trees is an important research area, it is beyond the scope of this study. For the completeness of the research, we briefly described the process of the tree segmentation in *Section 4.2.3.2*. Previous studies have demonstrated that descriptive metrics can be

extracted from the point cloud through waveform decomposition (Heinzel and Koch, 2011; Yao et al., 2012), raw waveforms (Allouis et al., 2013) and composite waveforms (Cao et al., 2016). However, it is unclear which source of metrics is more useful for characterizing vegetation such as tree species discrimination. Therefore, we extracted variables from these three sources and conducted a feature selection process using RF and CF methods. Subsequently, the Bayesian, RF and CF methods were adopted to investigate the capability of FW LiDAR data alone on tree species identification using selected metrics from the feature selection step.

4.2.3.1 Waveform decomposition

As part of the preprocessing step of tree segmentation and metrics extraction, we first applied the Gaussian decomposition (Eq. IV-1) for waveforms to obtain a point cloud as described in our previous study (Zhou et al., 2017). The present study will not go into details of how the point cloud is obtained, while major steps of the procedure are the preprocessing of waveforms, the fitting waveform with a mixture of Gaussian functions, and geolocation transformation. Once we obtained the point cloud, we employed LAStools (Isenburg, 2012) to classify ground points with *lasground* and then generated a CHM with 1 m resolution using the remaining non-ground points. Simultaneously, the echo width (u_i) and amplitude (A_i) of each waveform were derived for subsequent metrics extraction from the point cloud.

$$f(x) = \sum_{i=1}^{n} A_i \exp\left(-\frac{(x-u_i)^2}{2{\delta_i}^2}\right) \quad \text{(IV-1)}$$

where n is the number of Gaussian components, A_i is the amplitude of peak at ith waveform component, δ_i is the standard deviation of ith waveform component, and u_i is the time location of the peak at ith waveform component.

4.2.3.2 Tree segmentation

We first explored two commonly used algorithms including the variable window filter (Popescu and Wynne, 2004) and watershed algorithms (Chen et al., 2006) to segment tree crowns using the CHM derived from the waveform decomposition. The variable window filter was implemented in the TreeVaw package to identify the tree tops using local maximal (LM) algorithm and then to measure crown width (Popescu et al.,

2003) as shown in Figure IV-2(a). However, individual tree crown geometry was not well delineated since the shape of tree crowns is not always circular in the real-world. To reduce the bias caused by the tree crown segments that could significantly affect the accuracy of extracted metrics, the watershed algorithm was employed to detect individual trees (Figure IV-2(b)). The geometries of individual tree segments are more representative of shapes of real individual tree crowns while over-segmentation (green circles) is obviously observed with the watershed algorithm (Figure IV-2 (b)) as compared to tree tops identified by the TreeVaw. Thus, we integrated the LM algorithm with a continuously varying filter, as implemented in TreeVaw, with the marker-controlled segmentation algorithm (Chen et al., 2006) which is a variant of the watershed algorithm to conduct individual tree detection to overcome the problems suffering from the above two methods. According to experiments, some detected tree tops were close which resulted in over-segmentation in tree crown delineation. Thus, we combined detected tree tops within 5.7 m to obtain final tree tops. The whole process was implemented in R with the aid of the ForestTools package (Plowright, 2017). The marker-controlled segmentation algorithm delineated tree crowns based on the tree tops we identified in the above steps. The relationship between tree height and crown width, and the threshold of combing tree tops were determined with some sampling data collected from the point cloud and then refined with a trial-and-error method. In this section, we focused on describing the process conceptually and thus presented with less details of algorithmic developments, which can be found in the study of Beucher and Meyer (1992). Key parameters used for identifying tree tops were summarized in Table IV-1. To evaluate the accuracy of tree segmentation, we generated a 2 m buffer around the position of each delineated tree top and compared these buffers to the field-measured data. Once one field-measured data was present in the buffer, we assumed the tree was correctly detected, otherwise, it would be treated as “false detected tree”. If more than one detected tree were presented in the reference tree buffer, it will be assumed as over-segmentation tree. The nearest detected tree was kept for subsequent feature extraction and tree species classification in over-segmentation condition.

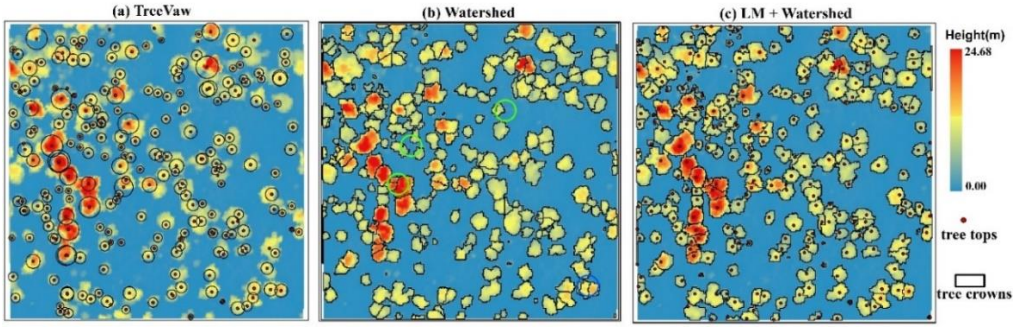


Figure IV-2. An example of tree segmentation with three approaches including the TreeVaw, watershed and LM + watershed algorithms.

Table IV-1. Key parameters used in the tree segmentation approaches.

Approaches	Function	Minimum height (m)	tolerance	extent
TreeVaw	0.804x + 3.67	4	NA	NA
Watershed	NA	3.5	1	2
LM + watershed	0.1 x + 2.15	2.5	NA	NA

***x** is the relative height (m); **Minimum height**: a value below this threshold will not be a crown; **tolerance:** The minimum height of the object in the units of image intensity between its highest point (seed) and the point where it contacts another object (checked for every contact pixel). If the height is smaller than the tolerance, the object will be combined with one of its neighbors, which is the highest: **extent**: Radius of the neighborhood in pixels for the detection of neighboring objects.

4.2.3.3 Feature extraction

To preserve a plethora of original information from FW LiDAR data and improve the use efficiency of FW LiDAR data in statistical models, we retrieved waveform metrics from three sources: the point cloud, raw waveforms and composite waveforms which are converted from raw waveforms through waveform voxelization.

Regarding the metrics from the point cloud, we conducted waveform decomposition of these selected waveforms and then summarized the average, standard deviation and maximum of the amplitude, echo width and peak time bins for each tree

crown segments. Additionally, the first peak and last peaks' average amplitude and echo width were also generated.

For raw waveforms, we first selected waveforms that fell into the delineated tree crowns and filtered out waveforms with one peak to reduce the bias of tree crown segments and to ensure waveforms selected were from vegetation. Within each tree crown, we averaged the metrics extracted from individual waveforms which have been proven to be successfully applied for characterizing vegetation such as height of median energy (HOME), roughness of outermost canopy (ROUGH), number of peaks (NP), return waveform energy (E) (Hermosilla et al., 2014a), waveform distance (WD), median energy height ratio (MEHR), front slope angle (FS) (Cao et al., 2016), height of 50% total waveform energy (HOHE), and half energy height ratio (HEHR). The average and standard deviation of these metrics were ultimately obtained as the model inputs. In addition, accumulative waveforms (AWF) based on the time bin and absolute height were generated as shown in Figure IV-3(c). The integral of these two AWFs, the integral of vegetation (VegI, 3 m above ground), the integral of ground (GI) and the ratio between the integral of vegetation and the additive integral of vegetation and ground parts (RvegT) (Allouis et al., 2013) were also investigated.

Previous studies have demonstrated that off-nadir angle of flights indirectly alters the absolute vertical distribution of reflected intensities and results in the stretchiness or tilt of waveforms (Hermosilla et al., 2014b). To reduce its possible impact on the characterization of internal structures of vegetation using FW LiDAR data, an approach similar to that of Hermosilla et al. (2014b) was adopted to convert raw waveforms to composite waveforms with the voxel resolution of 0.8×0.8×0.15 m. The reason for using this resolution was that the footprint of FW LiDAR is approximately 0.8 m and the vertical resolution of raw waveform is 0.15 m. Likewise, we applied the same procedures of retrieving metrics from raw waveforms to extract variables from composite waveforms. The main variables extracted from waveform and point cloud are summarized in Table IV-2 and a detailed description of all metrics from the three sources is demonstrated in Table IV-6.

Table IV-2. Description of main varibales extracted from waveform and point cloud.

Acronym (metrics)	Description
WD (waveform distance)	The distance from the waveform beginning to waveform ending.
WGD (waveform distance from ground)	The distance from the waveform beginning to assumed ground location.
HOHE (height of median energy)	The distance from waveform centroid to the assumed ground location.
MEHR (median energy height ratio)	HOHE/WGD
ROUGH (roughness of outermost canopy)	The distance from the waveform beginning to the first peak.
HOHE (height of half total energy)	The distance from half energy location to waveform ending.
HEHR (half energy height ratio)	HOHE/WD
FS (front slope angle)	The angle from waveform beginning to the first peak which is assumed to be canopy returns.
E (total return energy)	The total energy contained in the waveform from waveform beginning to ending.
VegI (integral of the vegetation part)	The integral of vegetation part which is 3 m above the assumed ground location.
GI (integral of the ground part)	The integral of ground part which is 3 m from the assumed ground location.
RvegT (the ratio between the integral of vegetation and the additive integral of vegetation and ground parts)	VegI / (VegI + GI)

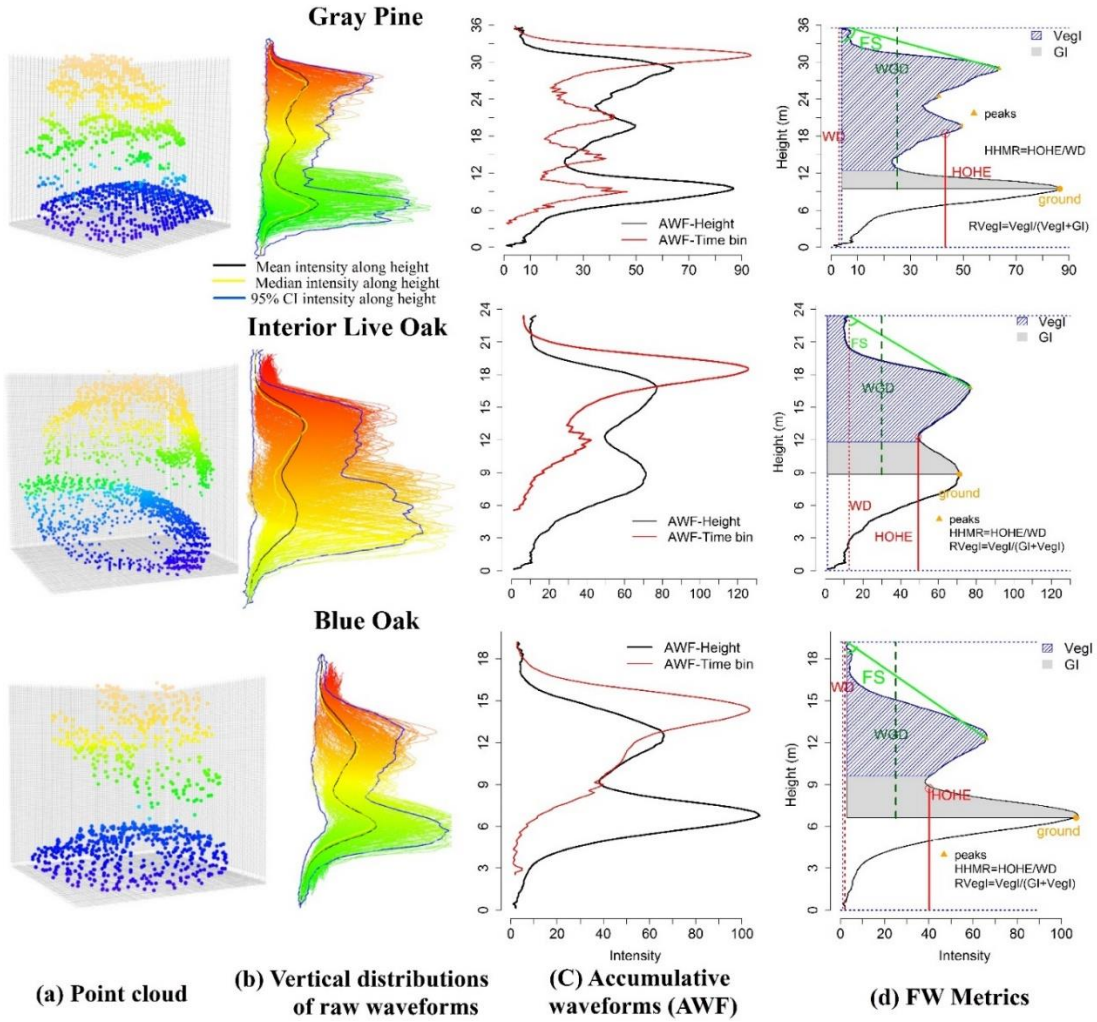


Figure IV-3. An example of waveform signals falling into the boundary of the gray pine, interior live oak and blue oak. (a) 3D visualization of three different tree species. (b) Vertical distribution of waveform signals along height. Blue, yellow and black lines are 95% confidence interval (CI) of intensity, median intensity and mean of intensity along the height, respectively. (c) Accumulative waveforms (AWF) along the time bin (red line) and height (black line). (d) An example of waveform metrics such as the waveform distance (WD), height of half energy (HOHE), the height of waveform beginning to the ground (WGD), crown integral (VegI), ground integral (GI) derived from raw AWF along the height (black).

4.2.3.4 Feature selection

Generally, dimension reduction techniques such as principal component analysis with statistical analysis are adopted to deal with a large number of variables. However, the effect of individual variables cannot be directly identified due to original variables are projected into a reduced set of components after an orthogonal transformation (Strobl et al., 2009b). Recently, a growing number of studies have highlighted the effectiveness of the RF algorithm to reduce the dimension of input variables for subsequent classification (Cao et al., 2016; Yu et al., 2014). RF (Breiman, 2001) is one of several popular ensemble learning methods based on the principle of recursive partitioning which could overcome the instability of individual classification trees induced by subtle changes of training samples. Originally, the variable importance is measured by merely counting the occurrence of each variable in all individual trees. As time evolves, the average splitting improvement of each variable in all individual trees is employed to assess the variable importance such as the Gini importance (Strobl et al., 2009a). More intuitive variable importance measure in the RF is the permutation accuracy importance which is calculated as the mean difference in prediction accuracy before and after permuting variable $\mathbf{X_j}$ over all trees. This measure has been shown to be a reliable criterion for quantifying the variable importance of uncorrelated predictors. However, the traditional permutation accuracy importance is prone to bring about potential bias for variable selection when predictors vary in scales or the number of categories (Strobl et al., 2009a). Additionally, the RF is likely to divide the importance amongst the correlated predictors. Consequently, none of these predictors may be significant for the model. In practice, the permutation results of the RF approach reflect independence of $\mathbf{X_j}$ from both Y and the remaining predictor variables $\mathbf{Z}$ ($\mathbf{X_1}$,… $\mathbf{X_{j-1}}$, $\mathbf{X_{j+1}}$,…, $\mathbf{X_p}$) instead of only the importance of $\mathbf{X_j}$ in predicting the response Y. Thus, we introduced a variant of permutation importance of the RF algorithm, conditional importance, to mitigate possible bias and avoid the impact of correlated variables for the feature selection in this study.

In contrast to the permutation importance in the RF, the conditional importance is achieved through unbiased classification tree generated with subsampling without

replacement method instead of the bootstrap sampling method. Additionally, the hypothesis of conditional importance measure is that $\mathbf{X_j}$ and **Y** are independent given the correlation structure between $\mathbf{X_j}$ and **Z** inherent in the data set. Within the framework of the conditional permutation, $\mathbf{X_j}$ is permuted only in a discretized permutation grid rather than the whole dataset. The grid is determined by the partitions derived from the model fitting or cut points of empirical correlations between $\mathbf{X_j}$ and **Z** (Strobl et al., 2009a). In the implementation, variables Z whose correlation with $\mathbf{X_j}$ meets the condition 1- p-value >0.2 are used for conditioning. For each grid, the out of bag (OOB) error prediction accuracy is measured in one tree after permuting the values of $\mathbf{X_j}$. Similar to the permutation importance, the conditional importance of $\mathbf{X_j}$ is the average prediction n accuracy of all classification trees.

In the present study, we employed both permutation importance and conditional importance to conduct the feature selection from 115 FW metrics. Prior to calculating the relative importance of variables, we excluded variables with importance (directly from the RF and CF methods) less than zero. The rationale behind this is that the importance of irrelevant variables varies randomly around zero (Strobl et al., 2009b). To minimize the overfitting of model and make it more generalized, we selected 9 candidate variables whose relative importance are larger than 1.5 with different characteristics as the input for subsequent Bayesian classification.

4.2.3.5 Tree species classification

To comprehensively explore the capability of FW LiDAR data on the tree species discrimination, three methods including the RF, CF and Bayesian inference were proposed to examine the usefulness of metrics and compare methods' prediction performances.

Random forests and Conditional inference forests

Not only can the RF measure the importance of variables, but it is also widely used for classification and regression. Briefly, the bootstrap method is adopted to resample part of the original dataset as training data to generate n classification or regression trees. For each tree, m variables out of p variables (total) were selected randomly to conduct

classification or regression. When a new data **X** is present, we aggregate predictions of the n trees' results using the new data. Suppose that the ensemble of trees is $\{T_i\}_1^n$, for regression, p(**X**) = $\frac{1}{n}\sum_1^n T_i(\boldsymbol{X})$; for classification, p(**X**) = majority votes $\{T_i(\boldsymbol{X})\}_1^n$. Within every tree, the OOB error is calculated using the remaining data from the original data as the OOB dataset.

Compared to the RF classifier, two main differences for the CF classifier are resampling method and permutation scheme as we have described in the *S*ection 2.3.4. The thrust for employing the CF method was to reduce the bias during variable selection in the RF method and provide a thorough understanding of the variable selection in a stringent statistical framework.

In this study, we set the number of trees to 10,000 for both methods and aimed to compare their performances with the same setting. Nine metrics were chosen for the final classification and other irrelevant variables were discarded. Through the experiment, incorporating more variables did not decrease the OOB error.

Bayesian inference

The motivation for introducing the Bayesian method in this study is to better understand the uncertainty of the classification and provide a new insight into the multinomial regression on tree species identification using FW LiDAR data. There are various models to conduct regression analysis for the multinomial data, such as the multinomial logit model, random utility model and multinomial probit model (Frühwirth-Schnatter and Frühwirth, 2016). Here, we employed the softmax regression model (Kruschke, 2014) to formulate the tree species prediction model to represent the categorical distribution.

We assume that **Y** follows a multinomial distribution with sample size n and parameter vector **p**. y_i is one of s unordered categories, labeled by $k = \{1, \cdots, s\}$. When the s is 2, the softmax model is reduced to the logistic regression model.

Generally, category 1 is adopted as a baseline category and we are interested in the model probability that belongs to categories from 1 to s. Thus, y_i takes value from 1 to s

according to the covariate information. For outcome k, the underlying linear propensity is denoted as:

$$\lambda_k = x_i \beta_k \quad \text{(IV-2)}$$

The softmax model can be formulated as follows:

$$\mathrm{P}(y_i = k|\beta_1, \cdots, \beta_s) = \frac{\exp(x_i \beta_k)}{\sum_{m=1}^{s} \exp(x_i \beta_m)} \quad \text{(IV-3)}$$

where $\beta_1, \cdots, \beta_s$ are category specific unknown regression coefficients and x_i is a row vector with covariates (1× d) which are not category specific and includes 1 for the intercept.

Subsequently, the Bayesian approach was pursued and a priori was assigned for all regression coefficients. There are two options to specify priors, the non-informative priors and empirical priors. The non-informative priors indicate that assigning equal probabilities to all possible values of parameter space which can reduce the effect of prior information on the posterior distribution of unknown parameters. For the empirical priors, a normal distribution N (b_0, S_0) or Dirichlet distributions can be assigned to these regression parameters. According to the Bayes' rule, the posterior distribution of β can be written as:

$$\mathrm{P}(\beta_1, \cdots, \beta_s|y) \propto p(\beta_1) \cdots p(\beta_s) \prod_{i=1}^{n} \frac{[\exp(x_i \beta_k)]^{y_i}}{\sum_{m=1}^{s} \exp(x_i \beta_m)} \quad \text{(IV-4)}$$

In our case, we have three tree species and one shrub (gray pine, interior live oak and blue oak), and either one can be a category response (y). In total, we used 127 delineated trees with field data as training data for the Bayesian model and each tree was characterized 9 variables that has been acknowledged as significant predictors by the CF and RF methods.

For the RF and CF methods, the OOB prediction accuracy with training data and confusion matrix derived from testing data were employed to quantify the prediction accuracy of tree species discrimination. Prior to deriving the Bayesian model inference, the model convergence was verified with the potential reduction factor, named Rhat ($\hat{R}$), which is a criterion to measure how well the Markov Chains are mixing and moving around the parameter space. Generally, $\hat{R}$ is expected to be close to 1; otherwise, a longer chain or more reasonable prior is necessary to be run to ensure that the chain reaches the

stationary state. As opposed to results of the RF and CF methods, the Bayesian methods could generate probability distributions belonging to four tree species for each delineated segments and the tree species of an individual tree segment was determined with the largest probability among these four estimated probabilities. Once the model was built, the left data (54 trees) as testing data were used to assess the accuracy of the tree species discrimination. An overview of the major steps for tree species classification using FW LiDAR data is demonstrated in Fig. 4.

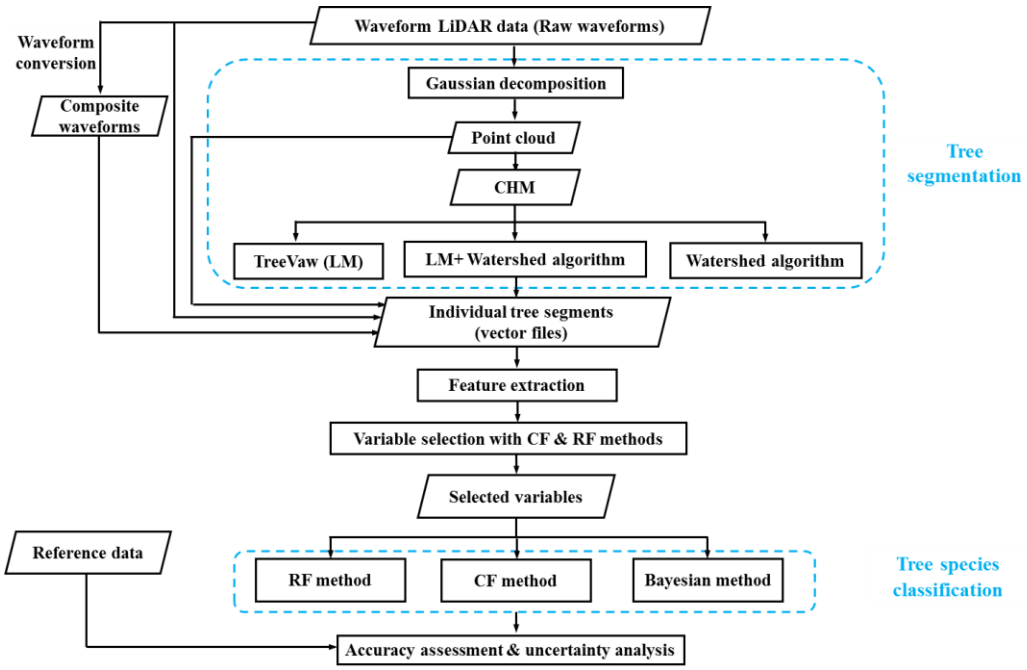


Figure IV-4. Flowchart for the tree species classification using waveform LiDAR data. **CHM**: Canopy Height Model. **RF**: Random forests. **CF**: Conditional inference forests.

4.3 Results

4.3.1 Tree segmentation

Three approaches including the TreeVaw (LM), watershed segmentation and LM + watershed segmentation were employed to conduct the tree segmentation in the SJER study site (Figure IV-2). The major difference of segmentation results between the

TreeVaw and the other two approaches was the shapes of individually delineated tree crowns. It was worthy to note that some identified tree tops with the TreeVaw had no delineated tree crowns in the border of the example region. The detected tree positions of these three approaches were indistinguishable while the watershed approach was prone to over segment individual tree crowns (blue circles).

To further illustrate segmentation results over the whole region, a quantitative evaluation of these approaches is presented in Table IV-3. Results of the tree segmentation demonstrated that the watershed approach outperformed other two approaches at tree detection, while its corresponding over-segmentation rate was the highest among these methods. In contrast, result of the TreeVaw showed the lowest over-segmentation rate with the highest false detection rate. These results, of course, were not surprising at all: it was natural to expect the higher accuracy of tree detection when more trees were delineated with a higher probability of over-segmentation. Through the LM + watershed approach, the bias caused by the over-segmentation can be reduced with lower over segmentation rate and decent tree detection rate (~ 90%). Thus, the segmentation results from the LM + watershed approach were adopted to conduct subsequent feature extraction.

Table IV-3. Results of different tree segmentation methods

Approaches	Tree detection rate (%)	False detection rate (%)	Over-segmentation rate (%)
TreeVaw	82.87	17.13	6.07
Watershed	92.82	7.18	15.58
LM + watershed	90.06	9.94	8.84

4.3.2 Feature extraction & selection

To visually display the pattern between tree species and waveform metrics from various sources, we generated waveform signatures for three representative trees (Figure IV-3). It is evident that the terrain and tree crown shapes are significantly different from

these three tree species for the waveform-based point cloud after the Gaussian decomposition (Figure IV-3(a)). As shown in Figure IV-3(b), the intensity distributions along the height for every individual waveform within an individual tree boundary are demonstrated. A common intensity distribution pattern for these three tree species emerged from Figure IV-3(b) showing that more energy was concentrated in the upper part and bottom part of three tree species. The 95% confidence intervals, median and mean of intensity were plotted against the height to display an overview of the intensity distribution of waveforms for different tree species. To present summary statistics of all waveforms within one tree segment, the AWF along the time bin (red line) and height (black line) for each tree species are generated as shown in Figure IV-3(c). As expected, the length of the time bin based AWF was shorter than the height based AWF. The length difference of these two AWFs was strongly correlated with the slope of the terrain where the tree located.

An interesting observation that emerged from the AWF comparison was that there was a subtle difference between interior live oak and blue oak while the latter had more energy distributed at the bottom part. However, their difference, with respect to the gray pine was evident in terms of the number of peaks and energy distribution along the height. These differences render a prospect for tree species classification with these waveform metrics. Thus, we demonstrated several waveform metrics extracted from the AWF along the height as an example in Figure IV-3(d). It was evident that the WD, VegI, FS and RVegT varied across different tree species from the visualization perspective.

To quantitatively identify influential waveform metrics for the tree species classification, the waveform metrics ranking and selection were conducted through the RF and CF methods. Figure IV-4 depicts the comparison of variable importance using the CF and RF methods with 10,000 random forest trees. For illustration purposes, we presented the largest 17 variables identified by the CF method with their relative importance and corresponding importance derived from the RF method to exhibit the comparison of variable importance results. Compared to the CF method, variables' importance of RF method was likely to be evenly distributed and difficult to recognize

significant variables, which undoubtedly mitigate the capacity of the RF method for measuring variables' importance. It was observed that the WDs and HOHE extracted from different sources were more significant than other variables for both approaches. Additionally, the WGD, RVegT, FS, VegI and NP were also highlighted as important variables for classifying tree species. In terms of sources of variables, composite waveforms were more representative than other sources. Interestingly, more variables derived from the AWF were found to be significant predictors than the average of variables from individual waveforms within one segment for the tree species discrimination.

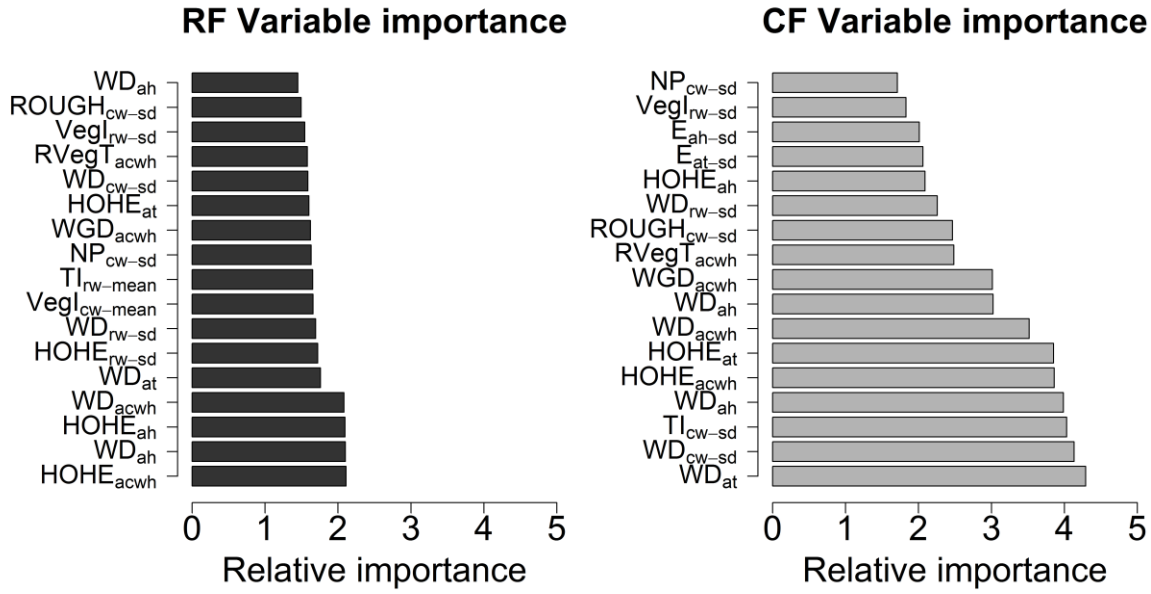


Figure IV-5. Comparison of variable importance using Conditional inference forests (CF) and Random forests (RF) methods.

To mitigate the overfitting caused by highly correlated variables and avoid the replication of the same variables from different sources, WD_{at}, $WD_{cw\text{-}sd}$, $TI_{cw\text{-}sd}$, $HOHE_{acwh}$, WGD_{acwh}, $RVegT_{acwh}$, $E_{at\text{-}sd}$, $VegI_{rw\text{-}sd}$ and $NP_{cw\text{-}sd}$ were used for subsequent Bayesian inference for the tree species classification. Here, the subscript represents the variables' source and statistical features (mean, max and standard deviation). For

example, WD_{at} denotes the WD from AWF of one tree segment along the time bins; $WD_{cw\text{-}sd}$ denotes that standard deviation of WDs derived from individual composite waveforms of one tree segment; $WD_{rw\text{-}sd}$ denotes that standard deviation of WDs derived from individual raw waveforms (rw) of one tree segment. Detailed descriptions of these variables can be found in Table IV-6.

4.3.3 Classification results

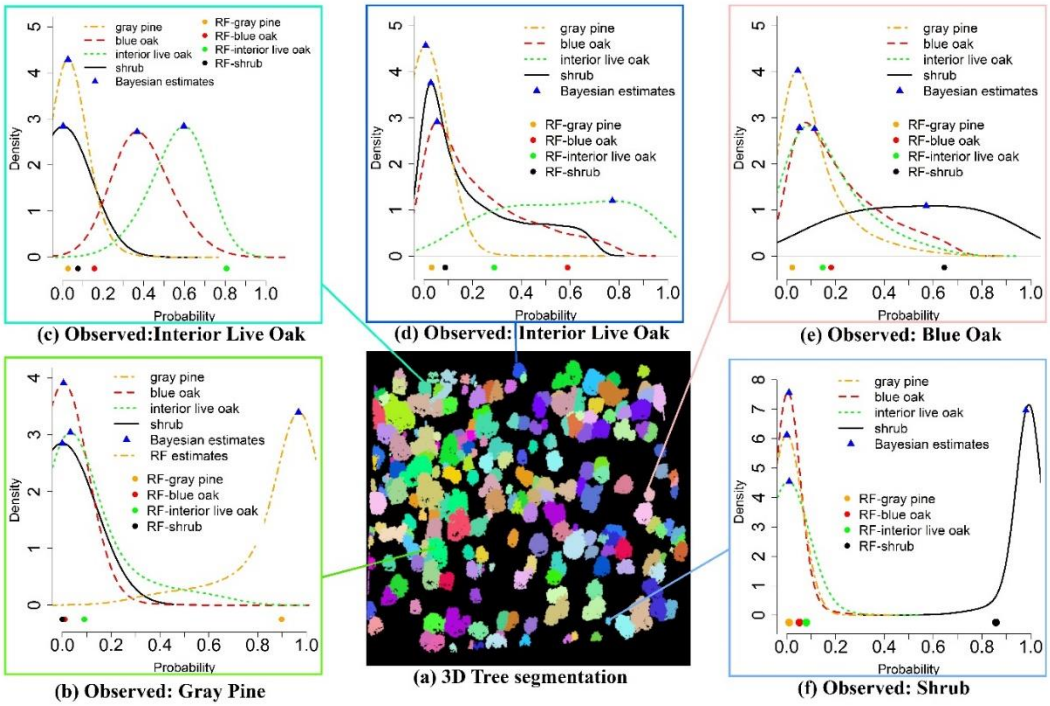


Figure IV-6. Examples of tree species classification using the RF and Bayesian methods with testing data.

Figure IV-6 exemplifies the classification results using the RF and Bayesian methods with five individual trees in a subset region. The CF method generated almost the same result as the RF method for these five trees, thus CF estimates were not plotted. In Figure IV-6(a), the clusters with different colors represent delineated individual tree segments with the LM + watershed method. The other five subplots were the five field measured trees with corresponding classification results using the RF (dots) and Bayesian

methods with probability distributions for possible tree species (blue triangles represents the mode of the probability). Compared to the field-measured data, Figure IV-6(b), (c) and (f) were correctly classified as gray pine, interior live oak and shrub, respectively, using both the RF and Bayesian methods. With the RF method, four estimated probabilities belonging to possible corresponding four tree species for an individual tree were generated and this tree is classified as gray pine with the largest probability (Figure IV-6(b)). As opposed to the RF method, the Bayesian method generated four probability distributions for these four tree species and the tree species of this tree was determined by comparing these probability distributions. For example, it is evident that trees are classified as gray pine and shrub with a compact probability distribution and high mode probability (> 0.92) in Fig 6(b) & (f), respectively, both of which are higher than corresponding probabilities obtained from the RF method. The RF method is also possible to correctly identify tree species with a higher probability than the Bayesian method as shown in Figure IV-6(c). In Figure IV-6(c), the tree is correctly identified as interior live oak for the RF method with a high probability, while the tree could be classified as blue oak or interior live oak in terms of probability distributions with the Bayesian method. Nevertheless, the tree is more probable to belong to interior live oak with a slightly higher probability when comparing the modes of probability distributions for both candidate tree species. Figure IV-6(d) illustrates that a tree is correctly classified as interior live oak using the Bayesian method according to its stretched probability distribution, but it is misclassified as blue oak with the RF method. It is also possible that the discrimination power of both methods fails when they are confronted with complex tree segments or wrong segmented tree crowns, such as in Figure IV-6(e).

To achieve a full comparison of classification performances, we summarized the classification results with the three methods in a confusion matrix. The prediction accuracy of the RF and CF methods can be measured with the training data (OOB error) and testing data, therefore we separated the results into Table IV-4 and Table IV-5. Overall, the most distinguishable species were the gray pine and shrub with a decent classification accuracy, both of which were differentiated from the other two tree species.

The least accurate tree species was the blue oak since it was easy to be misidentified as the interior live oak when taking the second column of tables into account. There was a subtle difference in the classification accuracy for the RF and CF methods in terms of the overall accuracy using the testing data, while the RF outperformed the CF from the perceptive of OOB error using the training data. A comparison of different methods using the testing data showed that the Bayesian method was superior to other two methods with higher overall accuracy and Kappa coefficient.

Table IV-4. Confusion matrix for vegetation classification using RF and CF methods with training data (OOB error).

Observed / Predicted	Gray Pine (%)	Blue Oak (%)	Interior Live Oak (%)	Shrub (%)
RF				
Gray Pine	94.59	5.00	6.67	0.00
Blue Oak	0.00	45.00	10.00	6.25
Interior Live Oak	5.41	40.00	80.00	6.25
shrub	0.00	10.00	3.34	87.50
Overall accuracy	79.61		Kappa	71.62
CF				
Gray Pine	93.75	0.00	20.00	0.00
Blue Oak	0.00	20.00	0.00	0.00
Interior Live Oak	6.25	80.00	66.67	0.00
shrub	0.00	0.00	13.33	100.00
Overall accuracy	68.18		Kappa	53.61

Furthermore, uncertainty of classification accuracy is also generated in a probabilistic sense with the Bayesian method rather than one estimate as shown in Figure IV-7. According to Figure IV-7, the 95% credible interval for the overall classification accuracy is approximately [0.52, 0.87]. Without doubt, this result is more informative than the counterpart such as the results from the RF and CF methods with only one estimate of classification accuracy.

Table IV-5. Confusion matrix for vegetation classification using RF and CF methods with test data.

Observed / Predicted	Gray Pine (%)	Blue Oak (%)	Interior Live Oak (%)	Shrub (%)
RF				
Gray Pine	93.75	0.00	13.33	0.00
Blue Oak	6.25	50.00	13.33	0.00
Interior Live Oak	0.00	50.00	60.01	0.00
shrub	0.00	0.00	13.33	100.00
Overall accuracy	72.73		Kappa	61.15
CF				
Gray Pine	94.59	5.00	10.00	0.00
Blue Oak	0.00	10.00	3.33	0.00
Interior Live Oak	5.41	80.00	83.34	12.50
shrub	0.00	5.00	3.33	87.50
Overall accuracy	73.79		Kappa	62.92
Bayesian				
Gray Pine	88.90	0.00	8.33	0.00
Blue Oak	5.55	27.27	8.33	10.00
Interior Live Oak	5.55	72.73	83.34	20.00
shrub	0.00	0.00	0.00	70.00
Overall accuracy	84.38		Kappa	77.56

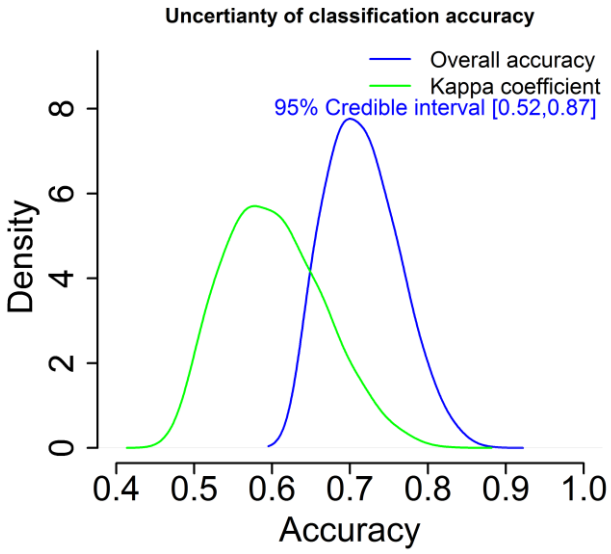


Figure IV-7. Uncertainty of classification accuracy using the Bayesian method.

4.4 Discussion

4.4.1 Tree segmentation

There are two major steps for the tree segmentation: the tree top detection and tree crown delineation. Based on experimental results, watershed algorithm alone is prone to face the over-segmentation problem at the tree top detection step. By contrast, the method implemented in the TreeVaw is more likely to give an unreasonable delineation of tree crowns. With the combination of the LM and watershed algorithm, these two flaws can be mitigated and consequently improve the accuracy of the tree crown delineation. The performance of the tree crown segmentation, on which subsequent accurate extraction of waveform metrics depends on, may be of greater significance. One salient reason is that our waveform metrics are derived from delineated individual tree crowns and inaccurate delineated tree crowns undoubtedly result in false waveform metrics and less precise tree species classification.

4.4.2 Individual trees' waveform signatures

It is evident that structures of the three tree species are different through visually comparing point clouds after the Gaussian decomposition (Figure IV-3(a)). Additionally, the vertical intensity distribution along the height for waveforms within one segment highlights the ground and canopy parts with concentrated energy distribution. These two intercepted surfaces with backscattering render significant insights into the tree height measurement and tree species characterization. In addition, higher amplitudes and intensities are prone to occur for the blue oak and interior live oak compared to the gray pine, while the gray pine has a larger number of echoes with longer pulse line (Figure IV-3(c)). What these trees share is that the first echo corresponding to the canopy part has a greater amplitude than those of subsequent echoes. A possible explanation for this maybe that transmission losses are occurring along the pulse penetrating the tree canopy: the deeper the pulse transmits, the less the energy is left. The internal structure of the blue oak (Figure IV-3(a)) is sparsely branched which may contribute to the more energy reaching the ground part and its second amplitude is almost the same as the first

amplitude. In addition, the comparisons of the AWF along the time bin and height bin can inform whether the topography of the tree's location is a slope or not, because the AWF along the time bin did not take the height information into account with the same starting bin and thus will make the length of its AWF shorter than the AWF along the height bin when a slope is present.

4.4.3 Waveform metrics and feature selection

This study expands on previous work in exploring waveform metrics for vegetation characterization using FW LiDAR data. We examine waveform metrics from three sources and conduct rank comparisons among these variables to identify which source of waveform metrics is more preferred for characterizing tree species. The raw waveform with the off-nadir angle could potentially provide more detailed structural information of vegetation, while it is possible to alter the vertical distribution of energy distribution and give rise to false information of representative objects along the height. Contrarily, composite waveforms overcome this problem by reassigning intensity into the absolute height bins. Results demonstrate that waveform metrics from composite waveforms are more informative than raw waveforms and point clouds. This may suggest that more attention should be paid to composite waveforms when the absolute vertical related information is used. Another important finding about variable importance is that the AWF for each individual tree segment contains enough information for characterizing features of different tree species. Therefore, the metrics derived from the AWF can be considered as a proxy for further vegetation characterization or biomass mapping without averaging extracted information from the individual waveforms within a segment.

The RF variable importance measures have been adopted as a sensible means for variable selection in many remote sensing applications (Cao et al., 2016; Chen et al., 2014; Karlson et al., 2015). Built upon the RF method, we introduced an alternative method, the CF, to explore their usefulness in variable selection. From a statistical perspective, both methods are capable of dealing with high dimensional datasets and solving "small n large p" problem with similar results (Figure IV-5). However, a closer examination of the variable importance suggests that the RF method is prone to be biased

for measuring the variable importance since the permutation process implemented in the RF method favors correlated predictor variables (Freeman et al., 2016) and divides importance among these predictors. Consequently, an evenly distributed importance for many variables are generated as shown in Figure IV-5. The main reason behind this is that the implicit assumption of independence between the variable (X_j) and remaining variables (Z) is commonly violated, and that degrades the capacity of the RF method for recognizing important variables (Strobl et al., 2009b). As expected, the CF method produces more reasonable variable ranking results and clearly distinguishes variables' importance by using the subsampling without replacement strategies and a conditional inference framework.

With regard to individual variables, the WD, TI, HOHE, WGD, RVegT, FS and ROUGH are identified as significant factors for tree species classification using both the RF and CF methods in this study. In accordance with previous studies (Allouis et al., 2013; Cao et al., 2016), the WD, HOHE and WGD have a stronger capacity for characterizing different tree species, which are related to the height or crown height of individual trees. Essentially, these results are also consistent with the height characteristics of these tree species, with significant differences in the tree heights among these tree species and shrub. More specifically, within our research area, the height of gray pine is generally the highest and the shrub has the lowest height. As demonstrated in Figure IV-3(d), the integral of the waveform (TI) and the integral ratio between vegetation part and the whole waveform (RvegT) vary among different tree species, and a higher percentage of the vegetation integral is expected in the gray pine since more energy is retained in the vegetation part with longer pulse lines and less energy reaches the ground.

It is worthy to note that variability of outcome for the variable selection process may exist using the RF and CF methods since the variables' importance is derived from a random permutation of predictor vectors (Strobl et al., 2009b). However, variables or the method proposed in this study may potentially serve as a practical guideline to identify significant metrics and infer characteristics of interest (e.g., volume, biomass, carbon) using FW LiDAR data.

4.4.4 Tree species classification and uncertainty

This study exemplifies the use of advanced statistical methods (the RF, CF and Bayesian) to explore a wealth of information contained in FW LiDAR data for discriminating tree species. In practice, various sources of randomness in the RF and CF could contribute to the discrepancy or instability of prediction accuracy: (1) subsampling process which randomly draws sub-data samples from the whole dataset, and (2) an individual tree building process which randomly draws part of variables from all variables (Strobl et al., 2009b). To obtain a reliable accuracy assessment, we set ntrees of 10,000 and mtry of 9 in both models. Table IV-4 and Table IV-5 demonstrate results with setting the same random seed at the beginning of running to ensure the replication of results. The results of OOB error using the training data demonstrate that the prediction accuracy of the RF method is higher than the CF method. However, the CF method slightly outperforms the RF method when employing the testing data to evaluate the prediction accuracy. Both methods can generate acceptable results with a decent accuracy. It appears that the RF is superior to the CF method when we combined results from Table IV-4 and Table IV-5, while it is still arbitrary to judge the performances of these two methods only relying on the accuracy assessment using the result of a study area.

The statistical basis of the analysis for these two methods is built upon the frequentist inference. As an alternative, we developed a method based on the Bayesian inference to conduct the tree species classification. As demonstrated in Figure IV-6, the Bayesian method not only provide one estimate or probability for possible tree species, also render users to interpret results in a probabilistic sense for each tree. This interpretation is more intuitive through updating beliefs of the tree species in response to new data, and further gives us more confidence in the prediction results with predictive uncertainty. These advantages are more evident in dealing with the sophisticated tree species classification than simple classification with one tree species' probability outstands among the other possible tree species. For instance, the tree species in Figure IV-6(d) is misclassified with the RF and CF methods while it has been correctly identified using the Bayesian method with a flat distribution. These three methods consistently

perform well in simple classification examples such as in Figure IV-6(b) & (f). In addition, the frequentist inference is too rigid to cope with complicated real-world modeling. In this study, the difference of the probability for one tree belonging to the blue oak and interior live oak is relatively small due to their similar physical characteristics such as the round tree crown and similar tree height. For the RF and CF methods, only one probability estimate is obtained for each possible tree species and the predicted tree species is identified with the largest probability. Any variance or uncertainty in variable extraction and model building steps undoubtedly contributes to misleading classification conclusions, especially for discriminating the blue oak and interior live oak. With the aid of the Bayesian method, the overall accuracy may not be enhanced significantly, but it can make the inference become more realistic with quantifiable predictive uncertainty (Figure IV-7). Simultaneously, the confidence about conclusion and uncertainty levels of results can be illustrated with the probability density functions. For example, compared with results of Figure IV-6(d), a stronger belief of the trees belonging to the gray pine and shrub in Figure IV-6(b) & (f), respectively, is guaranteed due to their more compact distributions.

The superior performance of the Bayesian method is further substantiated with a higher overall accuracy and Kappa coefficient than the other two methods. In essence, the RF and CF methods are more like "black-box" methods and they are developed in a less stringent statistical framework (Strobl et al., 2009b), both of which may lead to less predictive results than the Bayesian method. Assuredly, more studies and applications of these methods should be exploited to further compare their performances in various applications and provide a more thorough understanding of how these methods can be practically utilized for real-world problems and when the caution should be exercised for result interpretation.

Surprisingly, these three methods all suffer from discriminating the blue oak from the interior live oak with relatively low accuracy. Various factors possibly contribute to the failure of methods including an inaccurate tree segmentation, an absence of the accurate field data and an insufficient number of waveform metrics. Specifically, our

waveform LiDAR data were acquired in leaf-on season and crowns of neighboring trees are more likely to be overlaid. The blue oak and interior live oak share similar shape of tree crown and heights which undeniably reduces the success rate of tree segmentation and brings more challenges for subsequent tree species classification. The study of Yao et al. (2012) also confirmed that the segmentation is more precise when the leaf-off season LiDAR data are used. It is anticipated that higher accuracy is obtained and more transparent vegetation structure is expected to be characterized by FW LiDAR data in leaf-off season. We envision a concomitant expansion of adopting advanced statistical methods such as the Bayesian method for tackling complicated relationships between characteristics of interest (e.g. height, biomass, carbon) and remotely-sensed predictors, and assisting result interpretation and decision making in real-world applications with advances of computational capacity and handy operational tools.

4.5 Conclusions

This study integrates an exhaustive set of waveform metrics from the waveform-based point cloud, raw waveforms and composite waveforms with machine learning (the RF and CF) and Bayesian methods to discriminate tree species using FW LiDAR data alone. We combine the LM and watershed algorithms to derive more accurate individual tree segments, aiming to mitigate uncertainty and error brought into subsequent metrics extraction step. The machine learning methods such as the RF and CF demonstrate that they are powerful tools for coping with “small n large p” problems. Moreover, the CF method can overcome the possible bias of the RF method in measuring variable importance which violates the implicit null hypothesis and favors correlated waveform metrics. Results of the variable importance suggest that composite waveforms provide more informative metrics than other sources. Specifically, waveform metrics such as the WD, HOHE, WGD, RVegT, TI and ROUGH are highlighted as significant metrics to characterize tree species in our study. Tree species classification results show that the Bayesian method achieves higher overall accuracy and Kappa coefficient as compared to the RF and CF methods. Additionally, the prediction uncertainty of each tree is also generated with the Bayesian method which permits the users to interpret results in a

probabilistic sense without just relying on one estimated probability to construct decision rules to determine tree species. Furthermore, the uncertainty of accuracy for tree species classification provides a comprehensive overview of classification performance which can assist to determine whether the classification accuracy reaches a particular standard and further guarantee users with more confidence to apply results for real-world tasks such as forest inventory. Certainly, further testing of the methods or framework developed in this study is required and recommended for various forest or vegetation types. In addition, more efforts should be directed to exploit rich information contained in FW LiDAR for biomass and vegetation mapping.

4.6 Appendix

Table IV-6. Summary of FW metrics from FW LiDAR data

Waveform metrics	Definition (within an individual tree segments)
Individual raw waveforms	
Area	The area of an individual tree crown segment
$NP_{rw\text{-}mean}$	The average number of detected peaks for raw waveforms
$NP_{rw\text{-}sd}$	Standard deviation of NP
$MaxP_{rw}$	Maximum of NP
$WD_{rw\text{-}mean}$	The average distance from waveform beginning to waveform ending using raw waveform
$WD_{rw\text{-}sd}$	Standard deviation of WD_{mean}
$HOME_{rw\text{-}mean}$	The average distance from height of median energy in waveforms to the ground
$HOME_{rw\text{-}sd}$	Standard deviation of $HOME_{mean}$
$HTMR_{rw\text{-}mean}$	The average ration between HOHE and WD
$HTMR_{rw\text{-}sd}$	Standard deviation of HTMR
$HOHE_{rw\text{-}mean}$	The average distance from height of half energy in waveforms to the ground
$HOHE_{rw\text{-}sd}$	Standard deviation of HOHE
$HTHR_{rw\text{-}mean}$	The average ration between HOHE and WD
$HTHR_{rw\text{-}sd}$	Standard deviation of HTHR
$FS_{rw\text{-}mean}$	The average vertical angle from waveform beginning to the first peak
$FS_{rw\text{-}sd}$	Standard deviation of FS
$ROUGH_{rw\text{-}mean}$	The average distance form waveform beginning to the first peak
$ROUGH_{rw\text{-}sd}$	Standard deviation of ROUGH
$TE_{rw\text{-}mean}$	The average total energy of raw waveforms
$ME_{rw\text{-}mean}$	The average energy of each raw waveform
$MaxI_{rw\text{-}mean}$	Average maximum intensity of all raw waveforms
$TI_{rw\text{-}mean}$	The average integral of energy along height from waveform beginning to the ground
$VegI_{rw\text{-}mean}$	The average integral of energy along height from waveform beginning to 3-m above ground
$RVegT_{rw\text{-}mean}$	The average ratio between VegI and TI
$MaxI_{rw\text{-}sd}$	Standard deviation of maximum intensity

Table IV-6 Continued

$TI_{rw\text{-}sd}$	Standard deviation of the integral of energy along height from waveform beginning to the ground
$VegI_{rw\text{-}sd}$	Standard deviation of the integral of energy along height from waveform beginning to 3-m above ground
$RVegT_{rw\text{-}sd}$	Standard deviation of the ratio between VegI and TI for all waveforms in the individual tree crown segment
Accumulative waveform along the time bin	
NP_{at}	The number of peaks in the time bin based accumulative waveform
WD_{at}	The distance from waveform beginning to the ground in the time bin based accumulative waveform
$HOME_{at}$	The distance from height of median energy to ground in the time bin based accumulative waveform
$HOHE_{at}$	The distance from height of half energy to ground in the time bin based accumulative waveform
$HTMR_{at}$	The ration between $HOME_{at}$ and WD_{at}
$HTHR_{at}$	The ration between $HOHE_{at}$ and WD_{at}
FS_{at}	The front slope angle of the time bin based accumulative waveform
$ROUGH_{at}$	The average distance form waveform beginning to the first peak in the time bin based accumulative waveform
$E_{at\text{-}mean}$	The average energy of the time bin based accumulative waveform
$E_{at\text{-}sd}$	Standard deviation of energy of the time bin based accumulative waveform
Accumulative waveform along the height	
NP_{ah}	The number of peaks in the height based accumulative waveform
WD_{ah}	The distance from waveform beginning to the ground in the height based accumulative waveform
$HOME_{ah}$	The distance from height of median energy to ground in the height based accumulative waveform
$HOHE_{ah}$	The distance from height of half energy to ground in the height based accumulative waveform
$HTMR_{ah}$	The ration between $HOME_{ah}$ and WD_{ah}
$HTHR_{ah}$	The ration between $HOHE_{ah}$ and WD_{ah}
FS_{ah}	The front slope angle of the height based accumulative waveform

Table IV-6 Continued

$ROUGH_{ah}$	The average distance form waveform beginning to the first peak in the height based accumulative waveform
$E_{ah\text{-}mean}$	The average energy of the height based accumulative waveform
$E_{ah\text{-}sd}$	Standard deviation of energy of the height based accumulative waveform
$MaxI_{ah}$	Maximum intensity in height based accumulative waveform
WD_{ah}	The distance from waveform beginning to the ground in the height based accumulative waveform
TI_{ah}	The integral of energy along height from waveform beginning to the ground
$VegI_{ah}$	The integral of energy along height from waveform beginning to 3-m above ground
$RVegT_{ah}$	The ratio between $VegI_{ah}$ and TI_{ah}
Point cloud	
$A1_{p\text{-}mean}$	The average amplitude of detected first peak of all waveforms within an individual tree crown segments
$A1_{p\text{-}sd}$	Standard deviation of the amplitude of detected first peak for all waveforms within an individual tree crown segments
$TB1_{p\text{-}mean}$	The average time bin locations of detected first peak for all waveforms within an individual tree crown segments
$TB1_{p\text{-}sd}$	Standard deviation of the time bin locations of detected first peak for all waveforms within an individual tree crown segments
$EW1_{p\text{-}mean}$	The average echo width of detected first peak for all waveforms within an individual tree crown segments
$EW1_{p\text{-}sd}$	Standard deviation of the echo width of detected first peak for all waveforms within an individual tree crown segments
$A2_{p\text{-}mean}$	The average amplitude of detected second peak of all waveforms within an individual tree crown segments
$A2_{p\text{-}sd}$	Standard deviation of the amplitude of detected second peak for all waveforms within an individual tree crown segments
$TB2_{p\text{-}mean}$	The average time bin locations of detected second peak for all waveforms within an individual tree crown segments
$TB2_{p\text{-}sd}$	Standard deviation of the time bin locations of detected second peak for all waveforms within an individual tree crown segments

Table IV-6 Continued

$EW2_{p\text{-}mean}$	The average echo width of detected second peak for all waveforms within an individual tree crown segments
$EW2_{p\text{-}sd}$	Standard deviation of the echo width of detected second peak for all waveforms within an individual tree crown segments
$A_{p\text{-}mean}$	The average amplitude of all detected peaks of waveforms within an individual tree crown segments
$TB_{p\text{-}mean}$	The average time bin locations of all detected peaks for waveforms within an individual tree crown segments
$EW_{p\text{-}mean}$	The average echo width of all detected peak for waveforms within an individual tree crown segments
$A_{p\text{-}sd}$	Standard deviation of amplitude for all detected peaks for waveforms within an individual tree crown segments
$TB_{p\text{-}sd}$	Standard deviation of time bin locations of all detected peaks for waveforms within an individual tree crown segments
$EW_{p\text{-}sd}$	Standard deviation of echo width of all detected peak for waveforms within an individual tree crown segments
Individual composite waveforms	
$NP_{cw\text{-}mean}$	The average number of detected peaks for these composite waveforms
$NP_{cw\text{-}sd}$	Standard deviation of NP for the composite waveforms
$MaxP_{cw}$	Maximum of NP for the composite waveforms
$WD_{cw\text{-}mean}$	The average distance from waveform beginning to the ground for the composite waveforms
WDcw-sd	Standard deviation of WD for the composite waveforms
$HOME_{cw\text{-}mean}$	The average distance from height of median energy in waveforms to the ground for the composite waveforms
$HOME_{cw\text{-}sd}$	Standard deviation of HOME for the composite waveforms
$HTMR_{cw\text{-}mean}$	The average ration between HOHE and WD for the composite waveforms
$HTMR_{cw\text{-}sd}$	Standard deviation of HTMR for the composite waveforms
$HOHE_{cw\text{-}mean}$	The average distance from height of half energy in waveforms to the ground for the composite waveforms
$HOHE_{cw\text{-}sd}$	Standard deviation of HOHE for the composite waveforms
$HTHR_{cw\text{-}mean}$	The average ration between HOHE and WD for the composite waveforms
$HTHR_{cw\text{-}sd}$	Standard deviation of HTHR for the composite waveforms

Table IV-6 Continued

$FS_{cw\text{-}mean}$	The average vertical angle from waveform beginning to the first peak for the composite waveforms
$FS_{cw\text{-}sd}$	Standard deviation of FS for the composite waveforms
$ROUGH_{cw\text{-}mean}$	The average distance form waveform beginning to the first peak for the composite waveforms
$ROUGH_{cw\text{-}sd}$	Standard deviation of ROUGH for the composite waveforms
$TE_{cw\text{-}mean}$	The average total energy of composite waveforms
$ME_{cw\text{-}mean}$	The average energy of each composite waveform
Accumulative composite waveforms	
NP_{acwh}	The number of peaks in the height based accumulative composite waveform
WD_{acwh}	The distance from waveform beginning to waveform ending in the height based accumulative composite waveform
$HOME_{acwh}$	The distance from height of median energy to ground in the height based accumulative composite waveform
$HOHE_{acwh}$	The distance from height of half energy to ground in the time height based accumulative composite waveform
$HTMR_{acwh}$	The ration between $HOME_{acwh}$ and WD_{acwh}
$HTHR_{acwh}$	The ration between $HOHE_{acwh}$ and WD_{acwh}
FS_{acwh}	The front slope angle of the height based accumulative composite waveforms
$ROUGH_{acwh}$	The average distance form waveform beginning to the first peak in height based accumulative composite waveform
ME_{acwh}	The average energy of the height based accumulative composite waveforms
$MaxA_{acwh}$	Maximum amplitude of the height accumulative composite waveforms
WGD_{acwh}	The distance from waveform beginning to the ground in the height based accumulative composite waveform
TI_{acwh}	The integral of energy along height using the accumulative composite waveform
$VegI_{acwh}$	The integral of energy along height using the accumulative composite waveform
$RVegT_{acwh}$	The ratio between $VegI_{acwh}$ and TI_{acwh} using accumulative composite waveforms
$MaxI_{cw\text{-}mean}$	Average maximum intensity of composite waveforms

Table IV-6 Continued

$TI_{cw\text{-}mean}$	The average integral of energy along height from waveform beginning to the ground using composite waveforms
$VegI_{cw\text{-}mean}$	The average integral of energy along height from waveform beginning to 3 m above ground using composite waveforms
$GroI_{cw\text{-}mean}$	The average integral of energy along height from 3 m above ground to ground using composite waveforms
$RVegT_{cw\text{-}mean}$	The average ratio between VegI and TI using composite waveforms
$MaxI_{cw\text{-}sd}$	Standard deviation of maximum intensity using composite waveforms
$TI_{cw\text{-}sd}$	Standard deviation of the integral of energy along height from waveform beginning to the ground using composite waveforms
$VegI_{cw\text{-}sd}$	Standard deviation of the integral of energy along height from waveform beginning to 3 m above ground using composite waveforms
$GroI_{cw\text{-}sd}$	Standard deviation of the integral of energy along height from 3 m above ground to ground using composite waveforms
$RVegT_{cw\text{-}sd}$	Standard deviation of the ratio between VegI and TI using composite waveforms

CHAPTER V

PHOTON COUNTING LIDAR: AN ADAPTIVE GROUND AND CANOPY HEIGHT RETRIEVAL ALGORITHM FOR ICESAT-2 DATA

The upcoming Ice, Cloud and Land Elevation Satellite-2 (ICESat-2) mission will offer prospects for mapping and monitoring biomass and carbon of terrestrial ecosystems over large areas using photon counting LiDAR data. In this paper, we aim to develop a methodology to derive terrain elevation and vegetation canopy height from test-bed sensor data and further pre-validate the capacity of the mission to meet its science objectives for the ecosystem community. We investigated a novel methodological framework with two essential steps for characterizing terrain and canopy height using Multiple Altimeter Beam Experimental LiDAR (MABEL) data and simulated ICESat-2 data with various vegetation conditions. Our algorithm first implements a multi-level noise filtering approach to minimize noise photons and subsequently classifies the remaining photons into ground and top of canopy using an overlapping moving window method and cubic spline interpolation. Results of noise filtering show that the design of the multi-level filtering process is effective to identify background noise and preserve signal photons in the raw data. Moreover, calibration results using MABEL and simulated ICESat-2 data share similar trends with the retrieved terrain being more accurate than the retrieved canopy height, and the nighttime results being better than corresponding daytime results. Compared to the results of simulated ICESat-2 data, MABEL data achieve lower accuracy for ground and canopy heights in terms of root mean square error (RMSE), which may partly result from the inconsistency between MABEL and reference data. Specifically, simulated ICESat-2 data using 115 various nighttime and daytime scenarios, yield average RMSE values of 1.83 m and 2.80 m for estimated ground elevation, and 2.70 m and 3.59 m for estimated canopy height. Additionally, the accuracy assessment of percentile heights of simulated ICESat-2 data further substantiates the robustness of the methodology from different perspectives. The methodology developed in this study illustrates plausible ways of processing the data that are structurally similar

to expected ICESat-2 data and holds the potential to be a benchmark for further method adjustment once genuine ICESat-2 are available.

Keywords: ICESat-2, Photon classification, Photon counting LiDAR, ATLAS, MABEL, Canopy height, Terrain elevation

5.1 Introduction

Various types of Light Detection and Ranging (LiDAR) data such as discrete-return (DR) and full waveform (FW) LiDAR data are increasingly used to characterize the earth's topography, quantify vegetation structure and provide insightful solutions to natural resource inventory and carbon budget characterization (Allouis et al., 2013; Lefsky et al., 2005; Neigh et al., 2013; Popescu, 2007; Zhou et al., 2017; Zolkos et al., 2013). However, the utility of small-footprint LiDAR data over large spatial scales to accurately monitor forest ecosystems remains largely impractical due to their high acquisition cost (Gwenzi et al., 2016; Swatantran et al., 2016) and the limited spatial coverage caused by low operation altitude and high requirements of pulse energy (McGill et al., 2013). The advent of emerging technologies, such as photon-counting LiDAR (PCL), has offered prospects for future spaceborne laser altimeters. In contrast to analog LiDAR, the PCL is unique in that it employs low energy expenditure, increased measurement sensitivity, high repetition rate and space operational altitude. These properties enable PCL to overcome the restriction of spacecraft prime power by generating dense along-track sampling (Zhang and Kerekes, 2014) and ultimately resulting in the large spatial coverage (Wulder et al., 2012).

Due to these advantages of PCL systems, the Advanced Topographic Laser Altimeter System (ATLAS) sensor will be deployed on the upcoming Ice, Cloud and Land Elevation Satellite-2 (ICESat-2) (Markus et al., 2017). There are two notable features of ATLAS: (1) a multi-beam system that consists of six individual beams (split from a transmitted laser pulse by a diffractive optical element) with three pairs along the track designed to meet the science requirements of detecting the spatial variability of ice surface and monitoring ice dynamics (Herzfeld et al., 2014); for each pair, a weak and strong beam have an energy ratio of approximately 1:4 to compensate for varying surface

reflectance; and (2) the micro-pulse photon-counting technology that is capable of efficiently detecting photons reflected back from the earth surface. The ATLAS design allows for dense along-track sampling and large spatial coverage with low energy requirements at high flying altitude (Swatantran et al., 2016). For example, these configurations will generate overlapping footprints on the Earth surface with a diameter of 14 m, spaced at 0.7 m along track. By comparison, the Geoscience Laser Altimeter System (GLAS) aboard the Ice, Cloud and land Elevation Satellite (ICESat) illuminated spots (footprints) of 70 m in diameter, spaced at 170 m intervals. Compared to GLAS, the data to be provided by ATLAS consist of individual geo-located photons with profile configuration instead of waveforms (Markus et al., 2017). In addition, the ATLAS instrument will only operate at one single pulse (532 nm) with 10 kHz laser repetition rate. The dense sampling and extensive spatial coverage will be beneficial to large-scale applications such as sea level change monitoring, forest structural mapping and biomass estimation, improved estimation of Global Digital Terrain Models (GDTM), and reducing uncertainties associated with estimated forest biomass and carbon. Moreover, ICESat-2 will facilitate the production of gridded global products after the three-year mission anticipated lifespan (Neuenschwander and Magruder, 2016) with the potential for increased synergy with other existing remote sensing images, such as Landsat, to further complement ongoing biomass and vegetation mapping efforts.

Ambient noise is generated along the real signal photons since solar background photons can be simultaneously received by the detector. Consequently, an individual photon can be reflected back from targets within the footprint, but the exact corresponding origin location will be unknown (Gwenzi et al., 2016). Much of the noise can be avoided with the nighttime operation of the PCL system when there is less solar background noise. Furthermore, fewer signal photons are expected to reflect off from the vegetation than ice surfaces due to lower reflectance and higher aerosol densities over vegetated areas (Herzfeld et al., 2014).

The test-bed sensors for the upcoming ICESat-2 mission such as the Slope Imaging Multi-Polarization Photon-Counting Lidar (SIMPL) (Dabney et al., 2010) and Multiple

Altimeter Beam Experimental Laser (MABEL) (Rosette et al., 2011) have been developed to inform scientists about the potential of future spaceborne laser altimeters to meet various science objectives. Recent studies have demonstrated promising prospects of utilizing these data to discriminate ice and water (Kwok et al., 2016), to retrieve 3-D vegetation structural attributes in a savanna ecosystem (Gwenzi et al., 2016), and to estimate vegetation cover and biomass in conjunction with Landsat 8 data in a dryland ecosystem (Glenn et al., 2016).

A critical task for the ecosystem community is to identify the ground and canopy surface from these photons to meet science objective of determining global canopy heights which hinges upon the ability to detect both the canopy surface and the underlying topography (Neuenschwander and Magruder, 2016). Generally, there are two major steps to derive terrain and canopy height from PCL data: (1) noise filtering of raw photons, and (2) canopy and terrain classification of possible signal photons. The performance of noise filtering, on which canopy and terrain classifications depend on, may be of greater significance. A few methods have been developed for noise filtering, such as histogram-based filtering algorithms (Gwenzi et al., 2016; Moussavi et al., 2014), the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) approach (Zhang and Kerekes, 2015) and the Bayesian approach (Wang et al., 2016). Prior to these studies, Magruder et al. (2012) proposed three filtering techniques including canny edge detection, probability distribution and local angle mapping to process MABEL data. Tang et al. (2016) developed a voxel-based spatial filtering method to generate noise-free dataset using the data from the High-Resolution Quantum Lidar System. All of them have been proven effective to some extent while also suffering from some concerns. For example, the histogram-based and DBSCAN methods are prone to error over complex terrain and can potentially lose useful information for subsequent signal and noise photons classification (Tang et al., 2016; Wang et al., 2016). The voxel-based spatial filtering method is mainly oriented for the point cloud dataset with high signal-to-noise ratio (SNR), which may be not suitable for the profile data with low SNR of ICESat-2 data. In addition, these methods were mainly tested on a limited number of MABEL data with accuracy assessment

conducted with co-registered reference data (Gwenzi et al., 2016). Undoubtedly, an additional co-registration processes or adjustment will complicate the validation processes and performance evaluation. Furthermore, the efficiency of algorithms for retrieving the canopy height using PCL data over different forest types and noise levels have not been adequately explored.

The overall goal of this paper is to develop a methodological framework to filter noise and retrieve the terrain and canopy height in various vegetation conditions and noise levels using PCL data such as MABEL and simulated ICESat-2 data, both of which share similar characteristics of the expected data from the ICESat-2 mission. The study is anticipated to advance understanding of ATLAS data, provide insights into the challenges to be expected from data processing and pre-validate the upcoming ICESat-2 mission. The innovative aspects of this study consist of (1) introducing an adaptive methodology to cluster the signal photons and refine parameters for ground and top of canopy interpolation over diverse vegetation conditions, and (2) building a framework to conduct noise filtering suitable for different possible data scenarios of the upcoming ICESat-2 mission which could render a valuable basis for processing genuine ATLAS data. Ultimately, the methodology and results of this study may potentially allow a more rapid adoption of ICESat-2 data once available by a large scientific community, for a range of ecosystem studies through the probable incorporation of existing remote sensing data.

5.2 Methods and materials

5.2.1 Study sites

To prepare the automatic ground and TOC detection algorithms over vegetation areas for the ICESat-2 mission, two test-bed sensor data (MABEL and simulated ICESat-2 data) with multiple vegetation conditions were investigated. For MABEL data, we used two data sets acquired on September 14, 2012, near Hinsdale and Chester in Vermont (VT), and one data set collected on September 21, 2012, near Jacksonville in North Carolina (NC), shown in Figure V-1. The vegetation in Vermont is mainly comprised of sugar maple (*Acer saccharum*), American beech (*Fagus grandifolia*) and hemlock (*Tsuga canadensis*), which belong to the eastern mixed forest type. In contrast, the study site near

Jacksonville is mostly deciduous forest covered with Black Willow (*Salix nigra*), Loblolly Pine (*Pinus taeda*) and Eastern Cottonwood (*Populus deltoides*).

Simulated ICESat-2 data were generated from DR LiDAR or FW LiDAR data using the simulator provided by the ICESat-2's NASA Science Definition Team (SDT). 118 datasets with multiple ecosystem types and different acquisition times were investigated. We selected five representative study sites shown in Figure V-1 to demonstrate the performance of the methodology. More specifically, we selected the woodland savanna on the Freeman Ranch in Texas, the boreal forest in Fairbanks, Alaska, the temperate hardwood in the Bays Mountain of Tennessee, the tropical forest in the Mondah forest of Gabon and the temperate pine forest near Huntsville, Texas. These study sites were selected because they represent different forest and vegetation types that were anticipated to produce different SNR data from the ATLAS system.

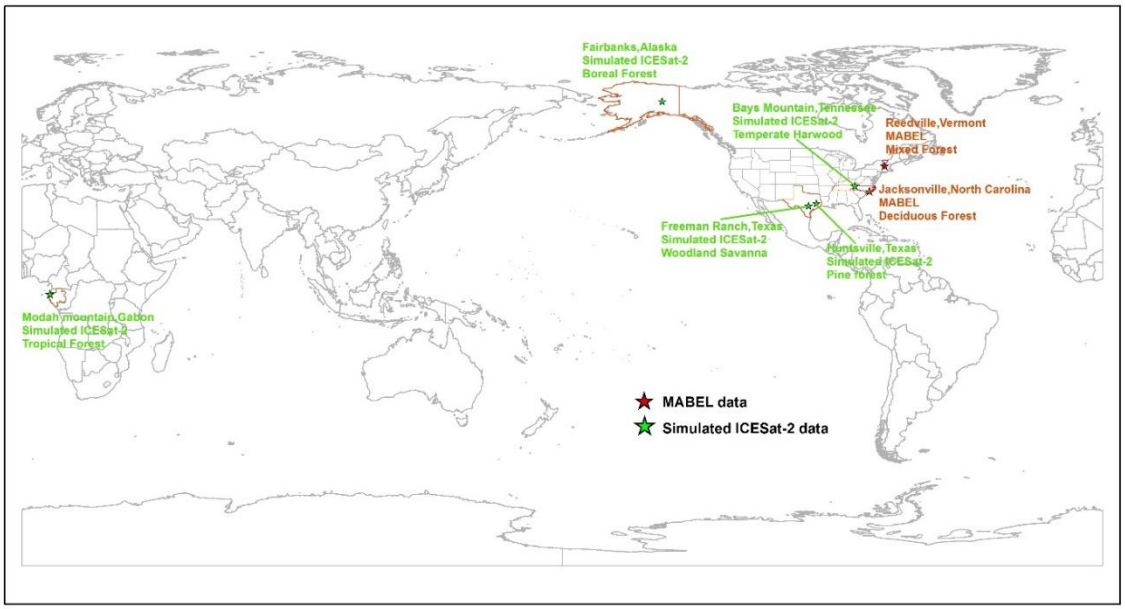


Figure V-1. Overview of representative study sites using MABEL (red) and simulated ICESat-2 data (green) with different ecosystem types.

5.2.2 Data

5.2.2.1 MABEL data

In preparation for constructing and launching of NASA's second spaceborne LiDAR system ICESat-2, NASA developed MABEL as a technology demonstrator for the ATLAS instrument. MABEL is typically operated at 20 km altitude using a NASA ER-2 aircraft to simulate data collection conditions close to those expected for the ICESat-2 mission using both green (532 nm) and near infrared (NIR, 1064 nm) laser wavelengths, with a pulse repetition rate that can be varied between 5 and 25 kHz. Beam splitters are employed to generate a total of 8 near infrared beams and 16 green beams from a single laser (McGill et al., 2013). The geometry of MABEL's beam configuration is designed to be reconfigurable to allow data collection and evaluation for multiple design cases. Data from the NIR channels were used in this vegetation study since they have higher reflectance and are resulting in higher SNR than the green wavelength (Gwenzi et al., 2016). Additionally, these returned NIR photons are more representative of the photons from the ATLAS instrument aboard ICESat-2 (Glenn et al., 2016).

5.2.2.2 Simulated ICESat-2/ATLAS data

The comparisons of the laser detector modalities for FW LiDAR and PCL systems demonstrate that the accumulated vertical distribution of the signal photons reflected back from the PCL is similar to the profile of FW LiDAR data (Neuenschwander and Magruder, 2016; Yin et al., 2016). Thus, the probability distribution of the detected signal photons along the vertical distribution can be represented by a normalized waveform (FW LiDAR) or pseudo-waveform (high-density DR LiDAR, >30 points/m^2). Based on this principle, the ICESat-2 SDT developed a simulator to generate simulated PCL data from existing DR and FW LiDAR data with 14 m diameter footprint and 0.7 m along-track spacing, the same as the planned ICESat-2 measurement scenario.

In reality, it is expected that 0 to 3 signal photons per transmitted laser pulse will be received by the ATLAS detector based on the ratio between the reflectance of highly reflective surfaces such as ice sheets and the canopy or terrain surface (Neuenschwander and Magruder, 2016). To ensure that the simulation is realistic, three design cases for

vegetation were developed to represent different vegetation conditions. Specifically, the mean number of signal photons in each space interval (0.7 m) expected by the ATLAS instrument is 1.0, 1.9 and 0.6 for the boreal forest, temperate forest and tropical forest, respectively. These design cases are consistent with the previous study's conclusion that more humid atmospheres could reduce the number of received signal photons and observations of previous studies conducted in tropical forest regions (Herzfeld et al., 2014). Furthermore, these settings could ensure that the model performance and accuracy are comparable to the science requirements of the ICESat-2 mission. There are four major steps to obtain simulated PCL data: (1) aggregating the frequency of discrete-returns along the elevation or intensities of small-footprint waveforms within each ICESat-2 footprint into pseudo-waveform; (2) constructing a vector along the elevation or height; (3) randomly determining the number of photons (x) per outgoing pulse according to design cases which was developed by the ICESat-2 SDT; and (4) randomly sampling the height vector weighted by pseudo-waveform x times to generate x photon(s) in any given footprint. The detailed simulation steps, design cases for generating simulated PCL data and their accuracy compared to airborne LiDAR data are presented in Neuenschwander and Magruder (2016). After the first step of generating simulated PCL data, different levels of noise representing the impact of the solar background and atmosphere were added into these data to obtain simulated ICESat-2 data.

The motivation for generating simulated ICESat-2 data is multifold. First, the simulated data from existing DR or FW data are automatically co-registered with reference data used as the basis of the simulation and no additional error will bring in the subsequent accuracy assessment. Second, noise levels for the simulated data can be adjusted flexibly to simulate effects of different solar elevation angles, atmospheric conditions and operation time (daytime and nighttime). This is unlike conditions during MABEL data collection which are typically static and only available for certain campaign dates (Gwenzi et al., 2016). Third, the data can simulate possible scenarios in different worldwide vegetation conditions and forest types that allow us to test the robustness of processing algorithms and fine-tune algorithm parameters in preparation for using

genuine ATLAS data once available. In this study, 115 scenarios were analyzed to explore the potential challenges of data processing and develop the algorithms to derive the ground and canopy height from the raw PCL data. Specifically, 13 representative scenarios within five study sites were selected to fully demonstrate the results. Among five study sites, three of them have two noise levels for daytime (Day low for the clear sky vs. Day high for high humid) and one level for the nighttime (Night). The remaining two study sites consisted of one daytime scenario and one nighttime scenario (Table V-1). The remaining 102 scenarios at 51 study locations with daytime and nighttime scenarios were also investigated.

Table V-1. Overview of the representative scenarios for simulate ICESat-2 datasets.

Vegetation types	**Locations**	**Noise level scenarios**
Woodland Savanna	Freeman Ranch, Texas	Day high, Day low, Night
Boreal Forest	Fairbanks, Alaska	Day high, Night
Temperate Hardwood	Bays Mountain, Tennessee	Day high, Day low, Night
Tropical Forest	Mondah Forest, Gabon	Day high, Day low, Night
Temperate forest	Huntsville, Texas	Day high, Night

5.2.2.3 Ancillary data and reference data

The existing Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) GDEM (version 9) and the Global Canopy Height Map (GCHM) produced by Simard et al. (2011) were used as ancillary data in the noise filtering step to narrow down the possible signal photons. There are two types of reference data to evaluate the algorithms' performances. The reference data for MABEL were DR LiDAR data and their corresponding products such as DEM and CHM from the Goddard's LiDAR, Hyperspectral & Thermal Imager (G-LiHT) system (Cook et al., 2013) available for download from NASA's G-LiHT website (https://gliht.gsfc.nasa.gov/). The reference data used in the present study were collected

on August 2011 with a point density of 10.3/m^2. For simulated ICESat-2 data, the signal photons generated from DR LiDAR data before adding noise serve as reference data, as well.

5.2.3 Methods

The methodological framework for processing MABEL and simulated ICESat-2 data is divided into two major sections: minimizing noise in raw PCL data (MABEL and simulated ICESat-2 data), and classifying the filtered PCL data as ground photons (GPs) and top of canopy photons (named canopy top photons in subsequent sections) to obtain the estimated ground and TOC surface. Prior to noise removal, the geolocation (latitude and longitude) were transformed to the along track distance (ATD) between the beginning location to the desired locations by calculating their great circle distances.

5.2.3.1 Minimizing noise in raw PCL data

In this step, we proposed a multi-level filtering approach to reduce noise from raw PCL data. The basic assumption of noise filtering was that signal photons were clustering around targets such as ground and canopy with higher density, and noise photons were more likely to be randomly distributed. Data from different regions acquired over various topography and vegetation conditions were characterized by different SNRs, which required us to develop an adaptive methodology suitable for various conditions. The filtering framework consisted of three steps with different objectives to minimize as much noise as possible from raw PCL data.

We proposed two approaches for the first step to mitigate atmospheric noise or random noise photons from raw PCL data (black points in Figure V-2(1) and Figure V-4(1)). The first approach was intended to create an envelope of signal and noise photons above and below the probable ground elevation known from ancillary data, such as the ASTER GDEM and GCHM, as illustrated in Figure V-2. These two elevation products were not intended to precisely filter all possible noise photons caused by the atmosphere or solar background but were meant to narrow down possible signal photons which could realistically be returned from the vegetation canopy and ground. The first step towards extracting possible signal photons was to utilize user defined thresholds ($minr$ & $maxr$,

red lines in Figure V-2(1)) above and below the GCHM and GDEM based on Eqs. (1) and (2).

$$minr = elevation_{DEM} - (pmin \times height_{GCHM}) \quad \text{(V-1)}$$

$$maxr = elevation_{DEM} + (pmax \times height_{GCHM}) \quad \text{(V-2)}$$

where $elevation_{DEM}$ and $height_{GCHM}$ were the DEM and CHM values of the given location for raw PCL data. The *pmin = 3* and *pmax = 3* were determined to be appropriate to mitigate potential errors in the two ancillary datasets and to avoid deleting possible signal photons (Figure V-2(1)).

To avoid reliance on ancillary data, a second alternative approach, named the grid-based statistical (GBS) filtering method, was proposed. This approach divided the horizontal and vertical space along ATD of the raw data into grids cells and then identified possible signal grid cells based on their statistical characteristics such as the number of photons, standard deviation (SD) and the p-value of the photon-fitting line for each cell. The assumption of the second approach was that more points were clustering around targets and a general trend could be observed from the specified grid cell without the aid of ancillary data. The grid spacing was chosen based on a trial-and-error approach that worked efficiently for all datasets to be 20 m for the vertical or elevation axis and 200 m along the ATD axis, as shown in Figure V-4(5) along the green and blue arrows, respectively. The statistical metrics such as the total number of the photons, SD and p-value of the fitting line through all photons in each grid cell were generated. The fitting line through all photons was produced mainly to check whether there is a possible linear trend of photons in the cell. Cells with a larger number of photons could indicate the presence of signal photons based on the assumption that noise may be randomly distributed throughout the grid space. In each grid interval along the ATD, the grid cell with the largest total number of photons was identified as possible signal cell. Additionally, two neighboring grid cells above and below the cell with the highest number of photons were selected. At this stage, 5 grids with 100 m elevation range were kept for each distance interval (Figure V-4(2)). Through comparing the adjacent grids' elevation index (green), dramatic changes of elevation could be detected as shown in Figure V-4(5).

This might suggest that our approach was able to cope with real world conditions where topography or canopy height was changing dramatically. To enable subsequent self-adaptive cluster analysis filtering, the photon density of the possible signal photons after first filter was calculated as *pd1* using the total number of photons divided by the maximum ATD of these photons.

The second step using cluster analysis was intended to remove noise photons while keeping signal photons within the envelope established through the first step, as shown in Figure V-2(2) and Figure V-4(3). The core of the cluster analysis was to trim the data by removing a proportion α of the "most outlying" photons (the most probable noise photons). In each ATD interval, one photon was either assigned to a cluster labeled with the index number or marked as outlier (noise) (Figure V-3).

The cluster filter was implemented in an R package named *tclust* (Fritz et al., 2012) which adopted a "crisp" clustering approach, meaning that each observation is either deleted or labeled as a cluster, using a robust mathematical probability framework to maximize the trimmed log-likelihood objective function (Eq. (V-3)). For a photon dataset $D = \{x_1, \cdots, x_n\}$, these photons were divided into k clusters labeled 1,…, k and the remaining data were labeled with zero which were regarded as noise photons. A k-variate normal distribution with mean μ and covariance matrix $\sum$ was proposed as the probability density function $f(x_i; \mu_j, \sum_j)$. Then the "spurious-outlier model" is defined via "likelihoods" as follows:

$$\left[\prod_{j=1}^{k} \prod_{i \in R_j} f(x_i; \mu_j, \sum_j)\right] \left[\prod_{i \in R_o} g_i(x_i)\right] \quad \text{(V-3)}$$

where R_j are photons belonging to the j[th] cluster, i is the indices of the data set D. R_o is the "outlier" cluster with the other probability functions $g_i(x)$. To estimate the parameters such as $\mu_j, \sum_j$, $g_i(x)$ and $f(x_i; \mu_j, \sum_j)$, the maximization of Eq. (V-3) was conducted. According to the study of Gallegos and Ritter (2005), maximizing the products in the first bracket in Eq. (V-3) also maximized the second one under certain sensible assumptions (see details in Gallegos and Ritter (2005)). Thus, Eq (3) can be simplified as

$$\prod_{j=1}^{k} \prod_{i \in R_j} f(x_i; \mu_j, \sum_j) \quad \text{(V-4)}$$

When k > 1, Eq. (V-4) is not a well-defined problem that requires the constraints on the cluster scatter matrix $\sum_j$ such as constraints on its eigenvalue and determinants (Fritz et al., 2012). The strength of the constraint can control the level of heterogeneity among clusters. With the aid of the constraint, the maximization of Eq. (V-4) can determine the elements belong to R_j. For k clusters, the total r elements of n were classified as signals and thus the trimming proportion will be 1 - r/n.

The *tclust* package had two functions that could be used to cluster the data, tkmeans and tclust, both of which worked well in all data sets for trimming possible outliers considered to be noise photons. Determining the suitable number of clusters k and the trimming proportion α are two key decisions for the cluster analysis. We tried different values for k in the range of 2 to 10, and our analysis indicated that trimming results were fairly insensitive to the choice of the number of clusters k in each moving window as long as k was not too small or too large. Thus, k=4 was used in each moving window for processing all data sets. However, the choice of α significantly affected the classification of signal and noise photons and more photons were removed with higher α (Figure V-3).

To make the clustering process become automated and self-adaptive for different terrain and vegetation conditions, the photon density was adopted as a criterion to determine the value of α. Specifically, we gradually increased α by 0.02 from its starting α value 0.45 and 0.05 for daytime and nighttime scenarios respectively, until the photon density of the current data was less than 60% and 95% of *pd1* for the corresponding daytime and nighttime scenarios. The two proportions (60% and 95%) were determined by comparing expected reference photon density to *pd1* using multiple datasets, which proved to be effective for the datasets used in this study. The parameters here are not universally valid for global terrain and vegetation conditions, however, the concept or method proposed here serves as a practical approach to determine these parameters.

To further reduce the distinct outliers that indicate noise photons, the third step was implemented to run a series of 95% confidence interval (CI) filters within each ATD distance interval to keep photons within 95% CI until the relative change of the total number of signal photons between two adjacent iterations was less than 5%. This step

was mainly intended to remove extreme outliers around possible ground and TOC after implementing the cluster filter and to minimize the amount of noise in the data before the classification procedure was initiated. The results using the third filter were illustrated in Figure V-2(3) and Figure V-4(4).

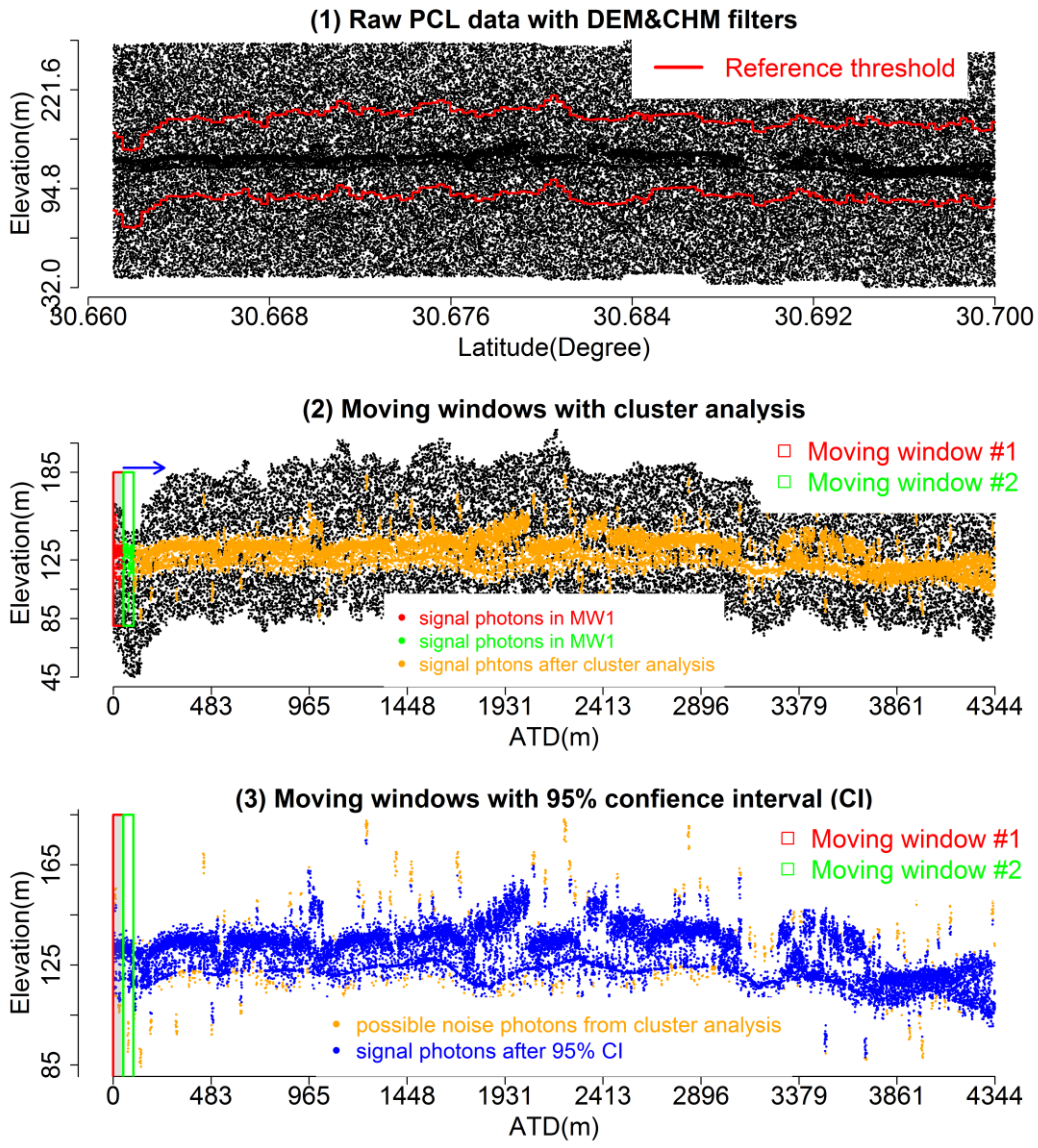


Figure V-2. An illustration of noise photons filtering using ancillary data (GDEM and GCHM) with moving windows: (1) Reference thresholds (red) generated from GDEM and GCHM using Eq. (V-1) and Eq. (V-2) with photon counting LiDAR (PCL) data; (2)

Moving windows example with cluster analysis; (3) The moving windows with 95% confidence CI filter.

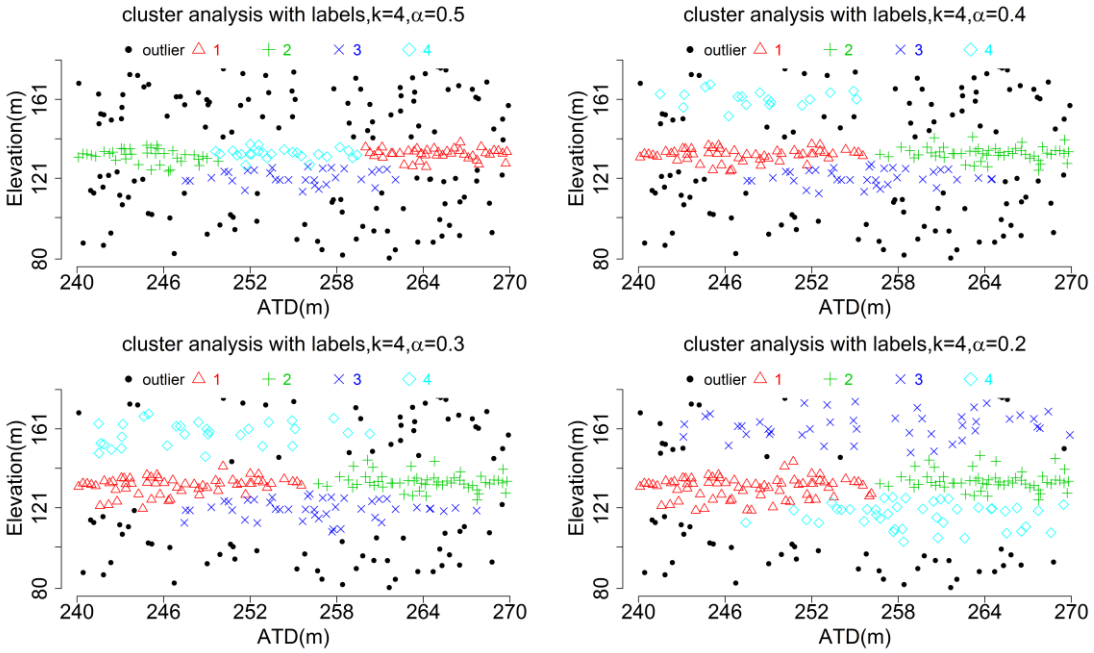


Figure V-3. Clustering results for photon counting LiDAR data with the same number of cluster (k) and different trimming proportions α.

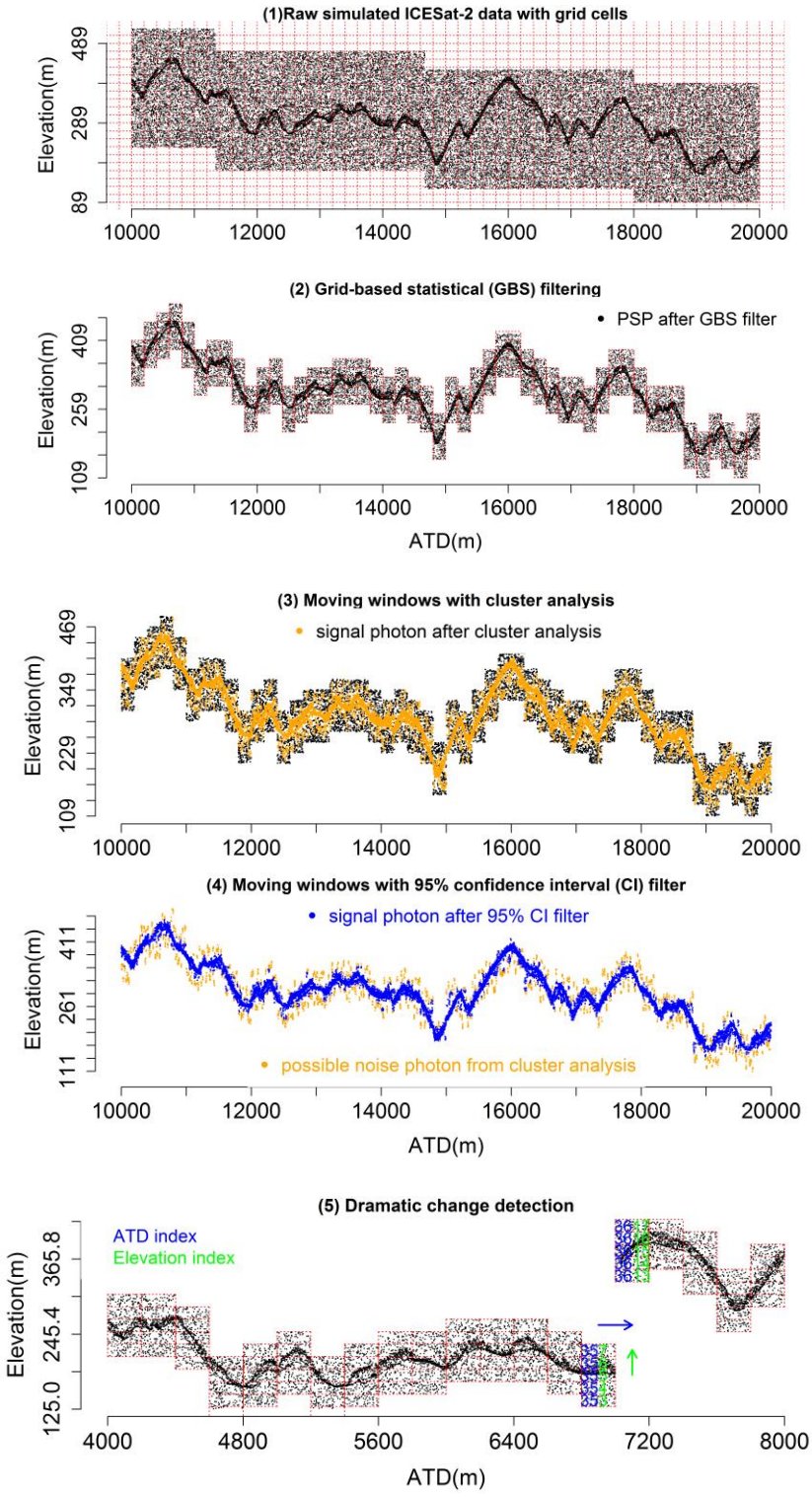


Figure V-4. An illustration of noise filtering using the grid-based statistical method, cluster analysis filter and 95% CI filter with simulated ICESat-2 data. (1) The grid cells

(red dash line) generated along ATD and elevation with raw simulated ICESat-2 data; (2) The possible signal photons (PSP) (black) after grid-based statistical filtering. (3) The PSP (orange) after the cluster analysis filter. (4) The PSP (blue) after 95% CI filter with possible noise photons from the cluster analysis (orange). (5) An example of dramatic change of signal photons with ATD and elevation indexes.

5.2.3.2 Classification of filtered PCL Data

Since noise photons around ground and TOC were still mixed with signal photons and difficult to be correctly classified after implementing the above steps, we introduced additional steps to classify some filtered photons as noise photons prior to classifying them into the GPs and canopy top photons. Next, an overlapping moving window was designed to process remaining photons to identify possible GPs and canopy top photons. Following this identification, the cubic spline function was utilized to generate continuous TOC and ground surface using the GPs and canopy top photons from the above step. Detailed steps can be found in the following Section *5.2.3.3* and *5.2.3.4*.

The fundamental idea of the cubic spline is to fit a piecewise function with third-degree polynomials over various specified intervals (Desquilbet and Mariotti, 2010). The cubic spline function requires continuity and slope constraint at each knot (the junctions of adjacent intervals) which represents continuous first and second derivatives. These constraints allow consistency and efficiency of the spline and avoid dramatic bends at the junction of intervals. Both characteristics make the cubic spline interpolation suitable for capturing the realistic canopy height and ground along the ATD. Specifically, the number of knots (NK) in the cubic interpolation can control the smoothness of interpolation and significantly impact the performances of the cubic spline interpolation. The ground surface is expected to be smoother than the TOC surface in most conditions, which in turn makes the cubic spline interpolation with adjusting NK appropriate to capture the ground and top of canopy estimates. In Figure V-5, we randomly selected a 2000 m interval to demonstrate the variation of the cubic spline interpolation with a different number of knots used.

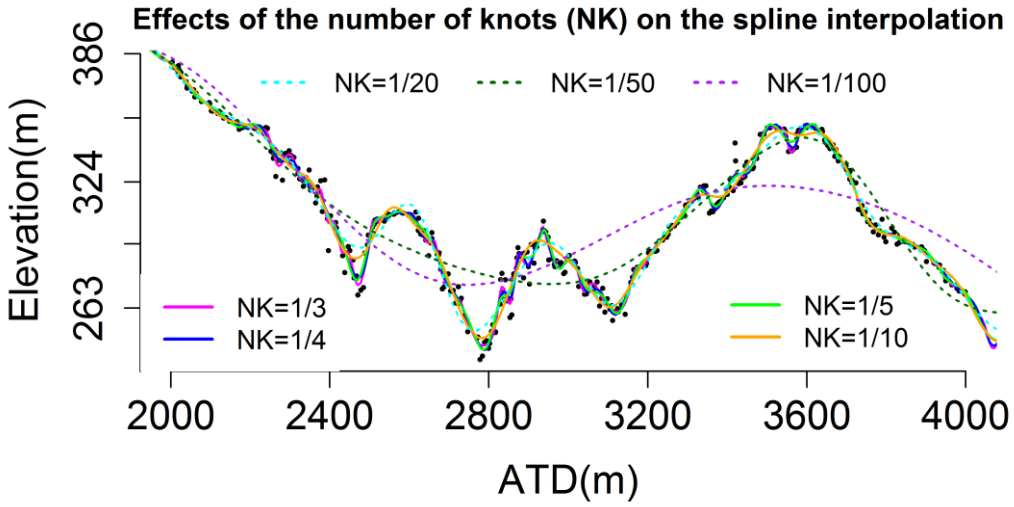


Figure V-5. An example of the effects of a different number of knots on the cubic spline interpolation. NK=1/3 represents the number of knots (NK) in the cubic spline interpolation equal to a third of the total number of estimated canopy top photons.

5.2.3.3 Canopy top identification

More specifically, three essential steps were used to generate the TOC from the signal photons after noise filtering. (1) An empirically derived width of the moving window for selecting canopy photons was 50 m and the moving step was 10 m, both of which had been found to be effective distance intervals in various conditions. We exemplified three moving overlapping windows to demonstrate this process in Figure V-6. The window moved one-fifth of its width to allow for the overlap between two adjacent windows. For each moving window, the photons (the red square, green triangle and blue plus sign photons) within specified elevation quantile range [0.99, 1] for the nighttime scenario and [0.96, 1] for the daytime scenario were assumed to be the possible noise photons in the TOC. The duplicated photons (orange circle) among these possible noise photons were identified by the adjacent windows and were assumed to be noise or false canopy top photons.

(2) To mitigate the effect of outliers (red circle in Figure V-7(2)) among these possible canopy top photons, we used a series of boxplot filters in each moving window to remove possible outliers and identify signal photons until the relative change of the

total number of signal photons between two adjacent iterations was less than 20%. The photons within the elevation quantile range [0.95, 0.99] were classified as canopy top photons from these signal photons. (3) Next, a cubic smoothed spline was adopted to fit these canopy top photons (green) after utilizing the *smooth.spline* function in R (2013) to generate the interpolated TOC line (blue) (Figure V-7(3)). To further illustrate, we displayed an example of identifying outliers (Figure V-8(1)) and refining possible canopy top photons with the boxplot filter (Figure V-8 (2) & (3)). The boxplot filter was based on the Tukey's method which used interquartile range to identify the outliers with less depending on the distribution of data (Figure V-8). For the sparse nighttime scenario, the step (2) was not necessary due to less noise after implementing the first two steps. Through various experiments, 1/5 of the total number of canopy top photons was found be an effective number of knots for the TOC interpolation.

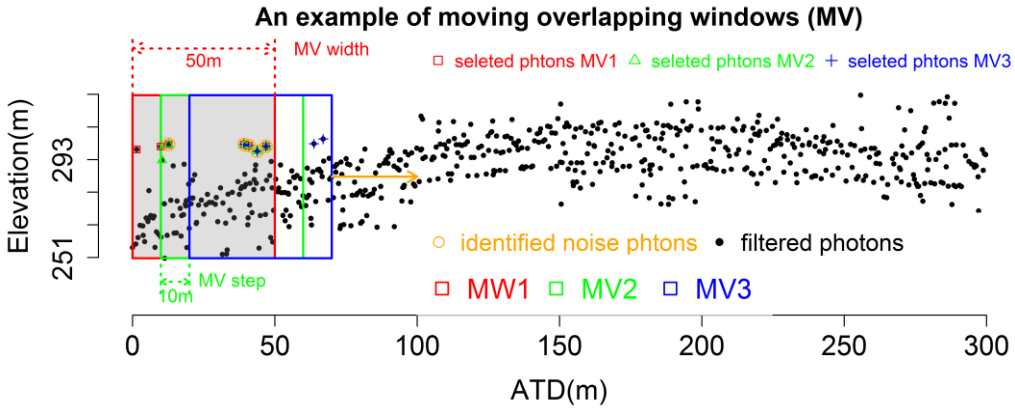


Figure V-6. An example of overlapping moving windows for identifying noise photons around the top of canopy (TOC).

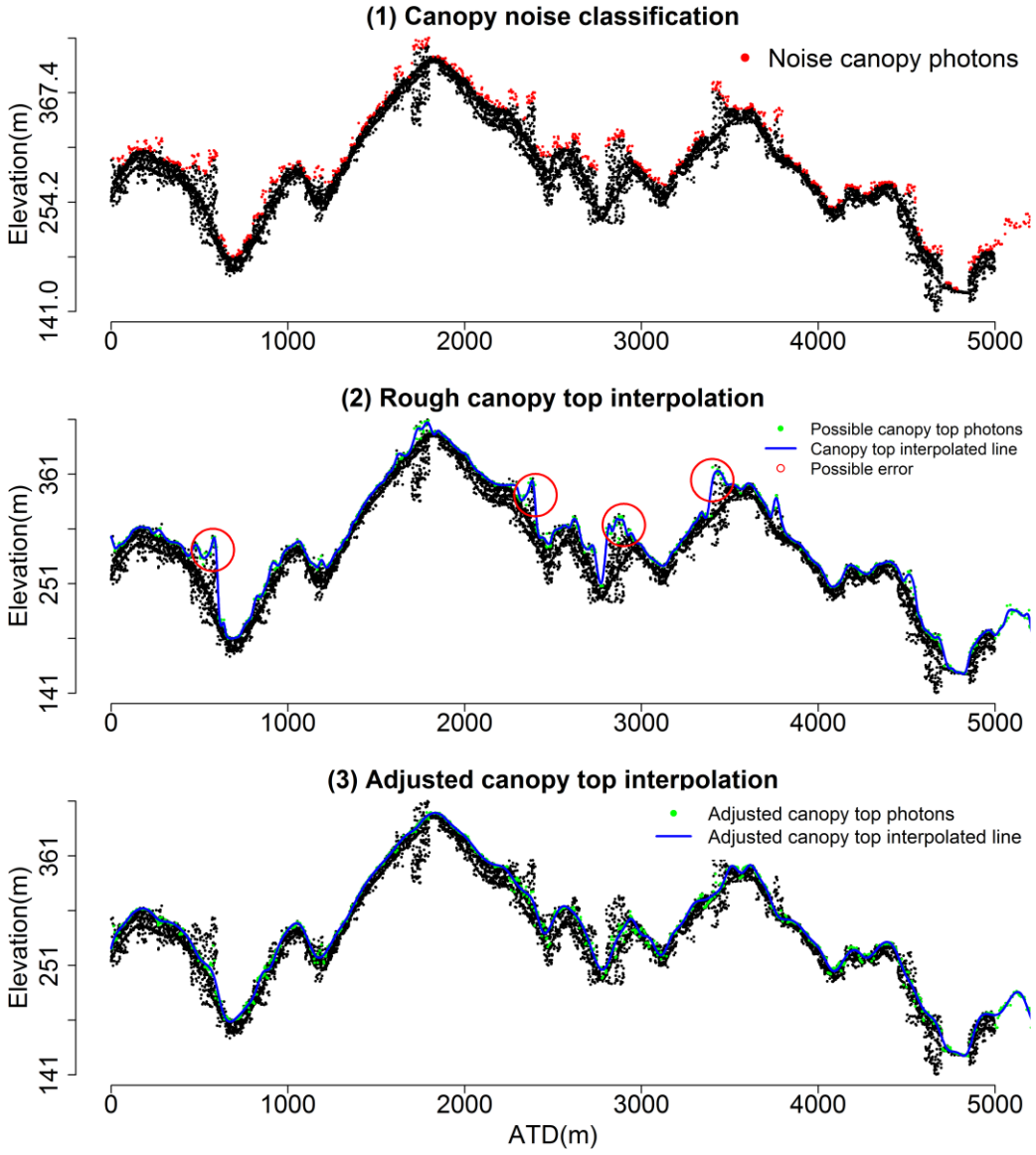


Figure V-7. Canopy top photons identification the continuous surface generation processes using simulated ICESat-2 data.

5.2.3.4 Ground identification

Compared to finding the possible canopy photons, identifying possible GPs was much more challenging especially under dense vegetation conditions. The possible reason was that the original transmitted energy has difficulties penetrating the atmosphere and

heavy tree cover, which consequently lead to fewer photons detected by the receiver aboard the satellite (Herzfeld et al., 2014; Moussavi et al., 2014). The brief steps of classifying GPs were summarized as follows: (1) as we did with the TOC identification, the first step of the ground identification was to identify false GPs (noise). The main difference was in the window width and moving step which for the sparse photon density scenario became 100 m and 20 m, respectively, to ensure that a sufficient number of photons could be used for terrain height estimation (Figure V-9(1)). Additionally, the elevation quantile range for each interval became [0.00, 0.05]. (2) Prior to selecting the possible GPs, a boxplot filter was used to further avoid some extreme outliers. For dense photon density scenario, the remaining photons within the elevation quantile [0.00, 0.05] were selected as the possible GPs (green dots) in each moving window to generate the rough ground interpolated line (blue line) as shown in Figure V-9(2). With regard to the sparse density scenario, we selected photons within the elevation quantile [0.00, 0.15] in each interval, and then derived the kernel density of these photons. The photons within 3 m of the maximum density elevation were assumed as the possible GPs. Here, the large range of elevation quantile [0.00, 0.15] was mainly chosen to get more possible GPs since fewer possible GPs were selected when we still used the elevation range [0.00, 0.05] in the sparse photon density scenario. In Figure V-9(2), there were several obvious incorrect interpolations marked with red circles because of the wrong choice of possible GPs in the dense photon density scenarios. Thus, (3) the automated refined process was conducted to adjust incorrect possible GPs after generating the rough interpolated ground line. It aimed to adjust the interpolated line close to the "real" GPs. Similar to the TOC adjustment, we used a series of boxplot filters in each moving window to remove possible outliers until the relative change of the total number of signal photons between two adjacent iterations was less than 20%. (4) Next, the cubic spline interpolation was used to fit the left GPs, the number of knots for the ground was 1/10 of the total number of GPs. Once ground and TOC surfaces were generated, the photons between ground and TOC surfaces were assigned as canopy photons.

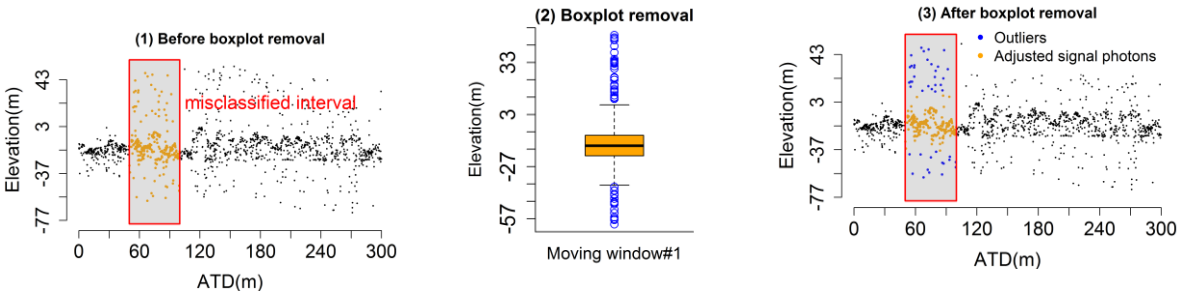


Figure V-8. An example of box plot removal filter for adjusting possible signal photons.

5.2.3.5 Accuracy assessment

To quantitatively calibrate the methods' performances, the accuracy assessment of the estimated ground, percentile height metrics (50^{th}, 60^{th}, 75^{th}, 80^{th}, 85^{th}, 90^{th}, 95^{th} percentile heights) and TOC from possible signal photons were conducted. Since DR LiDAR data from G-LiHT and MABEL data used different vertical coordinate systems, we conducted the height conversion between two coordinate systems according to the geoid height of the corresponding locations prior to implementing the accuracy assessment of MABEL data. Additionally, we further adjusted the elevation of MABEL data by adding the median difference between the converted MABEL and DR LiDAR data as described in the Gwenzi et al. study (2016) to reduce the geolocation error. For terrain and TOC, we extracted the validation elevations and heights with MABEL's individual points and their corresponding 2 m buffers from the DEM and CHM provided by G-LiHT to assess the accuracy of terrain and TOC detection. The 2 m buffer was utilized mainly because the approximated MABEL footprint diameter was 2m. Subsequently, the mean, SD and root mean square error (RMSE) were calculated based on these differences to quantify the performances of our algorithm.

Similar accuracy assessment procedures were applied to simulated ICESat-2 data. Unlike the reference data for MABEL, the reference metrics of simulated ICESat-2 were extracted from simulated PCL data with accurate geolocation information that did not need further conversion or co-registration. To validate the accuracy of the terrain and TOC of simulated ICESat-2 data, simulated PCL data (reference data) within 0 to 5^{th}

percentile and 95th to 100th percentile were first assigned to be the GPs and canopy top photons. Next, we employed the cubic spline interpolation to generate the continuous reference terrain and TOC line using the previously selected photons. The LiDAR metrics such as percentile heights have been commonly used to estimate the biomass of forest and reflect the canopy structures (García et al., 2010; Glenn et al., 2016; Wulder et al., 2008). In the present study, we conducted a validation process for these metrics for simulated ICESat-2 data to further test the performances of the methodology. The height metrics such as 50th, 60th, 75th, 80th, 85th, 90th, 95th percentile heights were directly extracted from the reference data and then compared with their corresponding metrics of simulated ICESat-2 data. Simultaneously, the descriptive statistics such as mean, SD and RMSE of the differences were derived to further assess the accuracy of the methodology. To clearly illustrate the proposed methodology, an overview of the major steps is exhibited in Figure V-10.

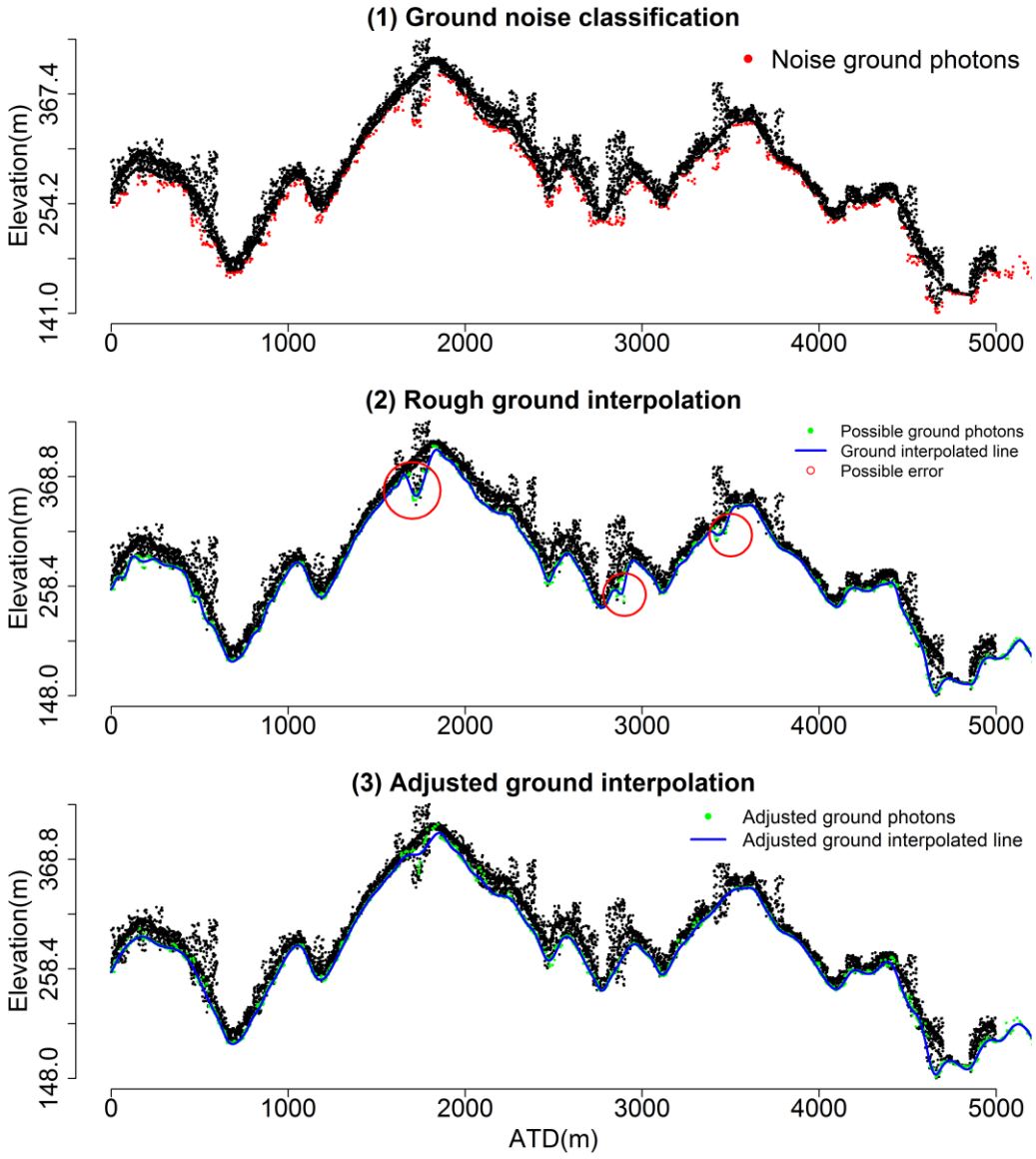


Figure V-9. Ground photons identification and the continuous surface generation processes using simulated ICESat-2 data.

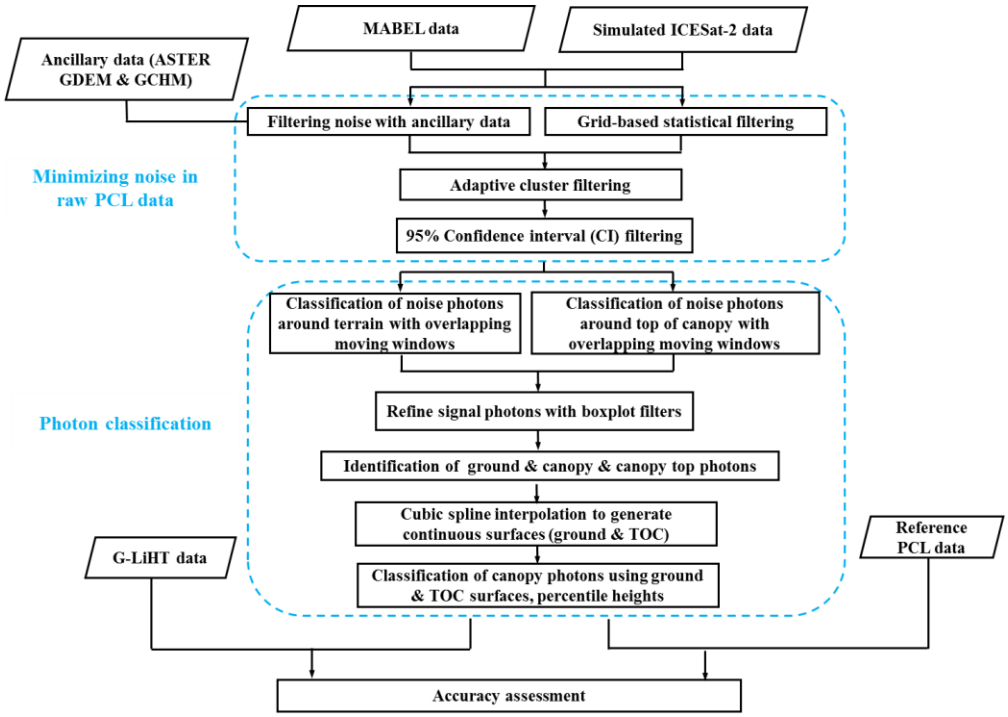


Figure V-10. Flowchart for the classification of photon counting LiDAR data.

5.3 Results

5.3.1 Noise filtering

As mentioned in the methods section, we tested two approaches for the first application of the noise filtering, using the ancillary elevation datasets and the grid based method. Results demonstrated that both approaches were effective in filtering out random noise photons and narrowing down the possible signal photons. Specifically, the first filter removed on average 65% and 5% of the total number of photons for the daytime and nighttime scenarios, respectively. Due to fewer noise photons existing in the nighttime scenarios, the proportion of the filtered photons was significantly smaller compared to the daytime scenarios (Table V-2). With the aid of the cluster analysis filter implemented in the second step of the filtering sequence, the majority of noise photons around ground and TOC were deleted as shown in Figure V-2(2) and Figure V-4(3). The third filter using

95% CI was useful for removing most of the extreme outliers which would significantly affect the performance of the classification algorithm. It should be noted that some noise photons still existed after implementing the three-step noise filtering sequence, our approach tried to mitigate the risk of removing possible signal photons. A summary of all scenarios indicate that the second and third filters remove on average 10% and 4% of the total number of photons for daytime scenarios, respectively. As expected, the second and third filters removed a lower percentage of the total of photons for the nighttime scenario, with approximately 5% and almost 0%, respectively. A closer examination showed that nighttime scenarios with less noise could generate satisfactory input datasets for subsequent classification after implementing the first two filters.

5.3.2 Ground and canopy top classification

5.3.2.1 MABEL

Figure V-11 presents the retrieved ground and TOC from one MABEL transect against the adjusted reference data. The original MABEL data was not co-registered accurately with the reference data (green) derived from G-LiHT (Figure V-11(1)), however, there was a consistent trend that could be observed between MABEL data and reference data. After alignment, MABEL data overlaid closely the adjusted reference data (red) used for subsequent accuracy assessment. Figure V-11(2) demonstrates an example of retrieved ground and TOC with corresponding descriptive statistics. The retrieved ground was potentially closer to the adjusted reference ground compared to the retrieved TOC. This observation was further confirmed by a smaller RMSE of the retrieved ground (5.25 m) than the TOC (6.98 m). Two other MABEL datasets were also investigated and the validation results were summarized in Table V-2. As expected, the validation results using individual points were less accurate than the result using the 2 m buffer for all datasets.

Table V-2. Validation results of MABEL retrieved ground and top of canopy.

Study sites	Criteria	Retrieved ground (m)			Retrieved top of canopy (m)		
		mean	SD	RMSE	mean	SD	RMSE
Chester, VT	Individual points	-0.84	5.25	5.62	0.32	6.45	6.98
	Buffers	0.21	4.16	4.20	0.54	5.04	5.05
Hinsdale, VT	Individual points	0.22	5.41	5.41	1.20	5.73	5.86
	Buffers	0.27	5.14	5.15	0.58	4.40	4.41
Jacksonville, NC	Individual points	0.07	0.78	0.85	-0.24	6.15	6.47
	Buffers	0.06	0.67	0.75	-0.19	4.26	4.57

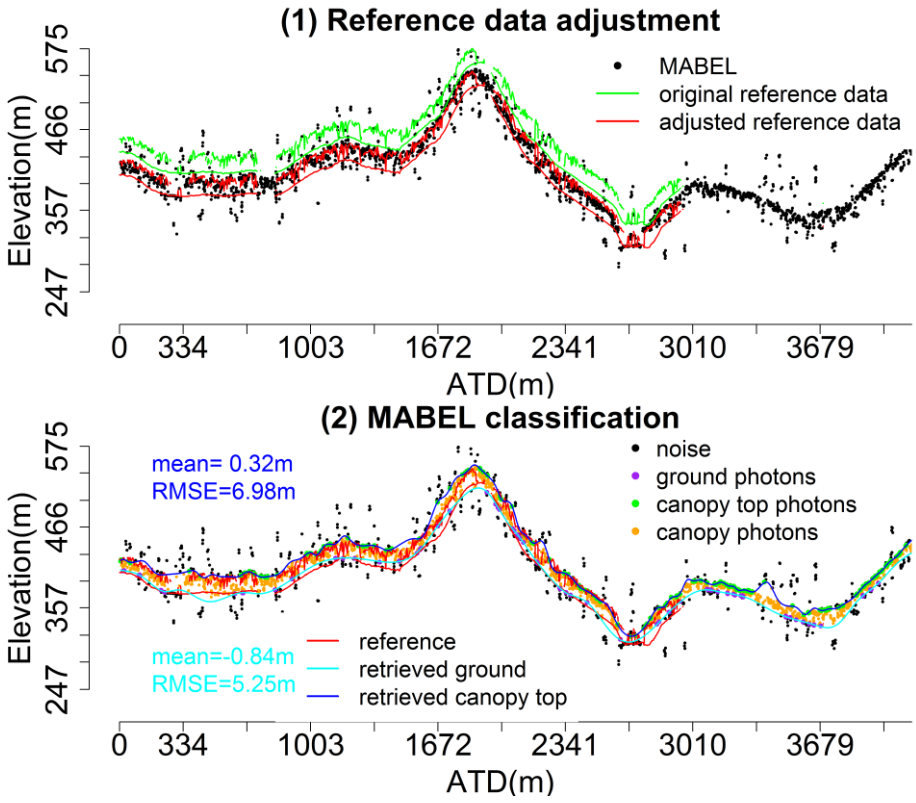


Figure V-11. Retrieved ground and top of canopy (TOC) results using MABEL data near Chester of Vermont (VT). (1) Original reference ground and canopy top (green) derived from G-LiHT data and adjusted reference ground and TOC data (red) against the filtered

MABEL data. (2) Filtered MABEL data, identified noise (black), ground (purple), canopy (orange), TOC photons (green), and reference ground and TOC (red) with interpolated ground (cyan) and TOC (blue) surfaces

5.3.2.2 Simulated ICESat-2 data

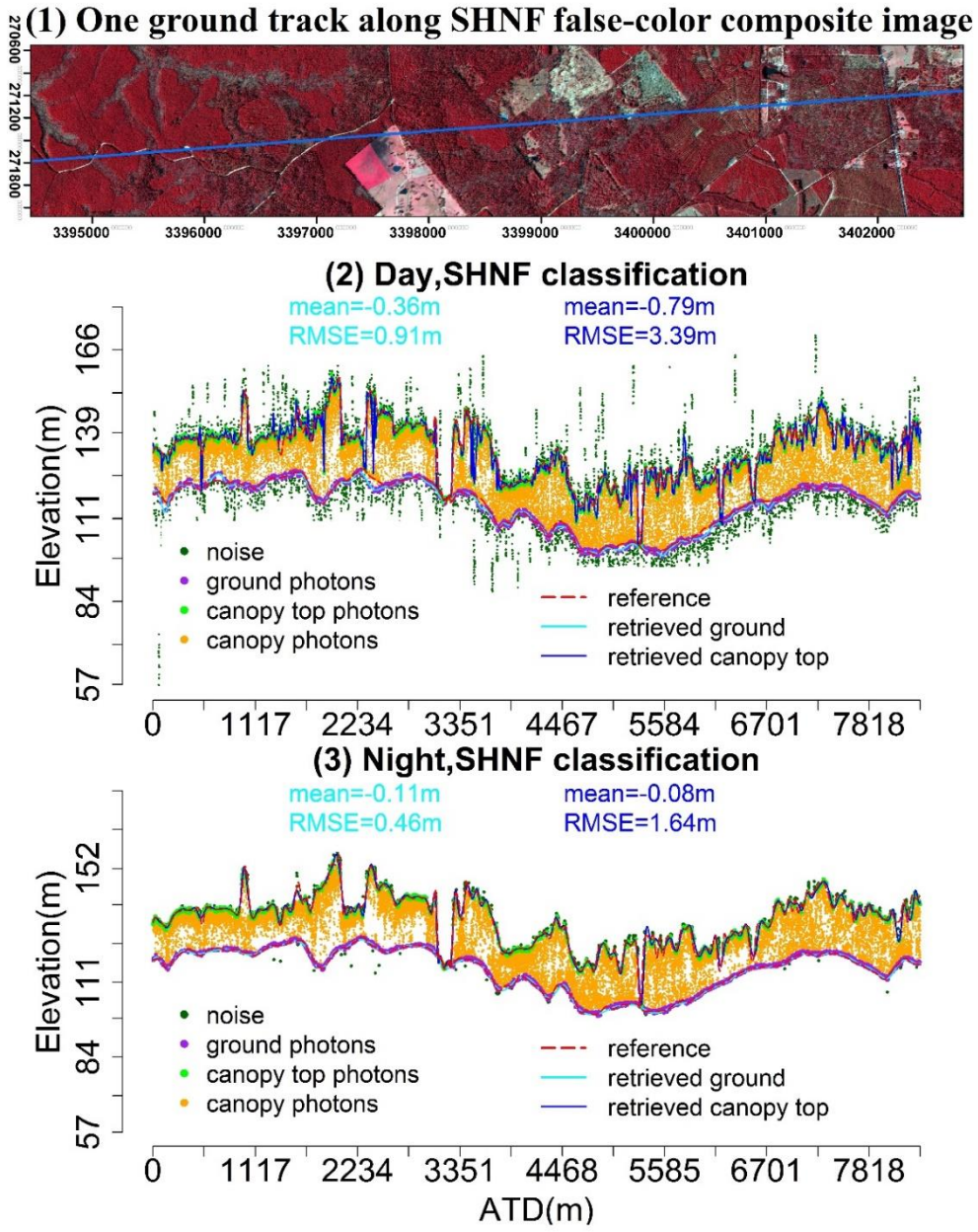


Figure V-12. Retrieved ground and canopy top results using one simulated ICESat-2 profile data with different acquisition time scenarios in the Sam Houston National Forest (SHNF). (1) False-color composite image with a ground track (blue line). (2) Filtered

simulated ICESat-2 data, identified noise (dark green), ground (purple), canopy (orange), canopy top photons (green), and reference ground and canopy top (red dash lines) with interpolated ground (cyan) and canopy top (blue) surfaces for the daytime scenario. (3) Filtered simulated ICESat-2 data, identified noise (dark green), ground (purple), canopy (orange), canopy top photons (green), and reference ground and canopy top (red dash lines) with interpolated ground (cyan) and canopy top (blue) surfaces for the nighttime scenario.

The capacity of simulated ICESat-2 data to identify ground and vegetation were explored in different scenarios. For illustration purposes, here we only selected two representative study sites to visually display the classification results. Simultaneously, a complete statistical performance evaluation for different vegetation conditions was summarized in Table V-3 and Table V-5.

As shown in Figure V-12, the photons left after noise filtering were classified as noise (dark green), ground (purple), canopy (orange) and TOC (green) photons for the daytime and nighttime scenarios at the SHNF region. As expected, more photons were assigned to the noise photons (green) in the daytime scenario. Compared to the reference ground and canopy top (red dash lines), mean bias and RMSE of interpolated ground and canopy top were smaller for the nighttime scenario than the daytime scenario. For the nighttime scenario, the identified ground and canopy top photons closely matched the reference data. Additionally, the RMSEs of interpolated canopy top were larger than the interpolated ground for both scenarios (3.39 vs. 0.91 m for daytime scenario and 1.64 vs. 0.46 m for nighttime scenario). Moreover, it was worth noting that more canopy photons (orange) were identified in the daytime scenario than the nighttime scenario. A closer examination revealed that the noise photons were mixed with the real signal photons within the canopy part and they were difficult to be detected and filtered out.

To further demonstrate the algorithms' performances in different vegetation conditions, we selected another study site representing sparse vegetation to visually display the classification result with three different scenarios. Figure V-13 depicts classification results of high noise level scenario with low returned photon density obtained from tropical forests of Gabon. It was observed that the amount of noise was around the reference data (green) (Figure V-13(1)) after GBS filtering of raw simulated

ICESat-2 data. Fewer photons reflected from the understory were kept after the cluster filtering step, especially for the ATD between 2700 and 4000 m (Figure V-13(2)), which matches the photon distribution of the reference data with low density at understory heights of the same interval (green) (Figure V-13(1)). Consequently, more variation between the retrieved ground (cyan line) and reference ground (red dash line) occurred in the region within this ATD interval. Regarding the canopy top, larger differences were more likely to happen in the same region where the large bias of the retrieved ground occurred.

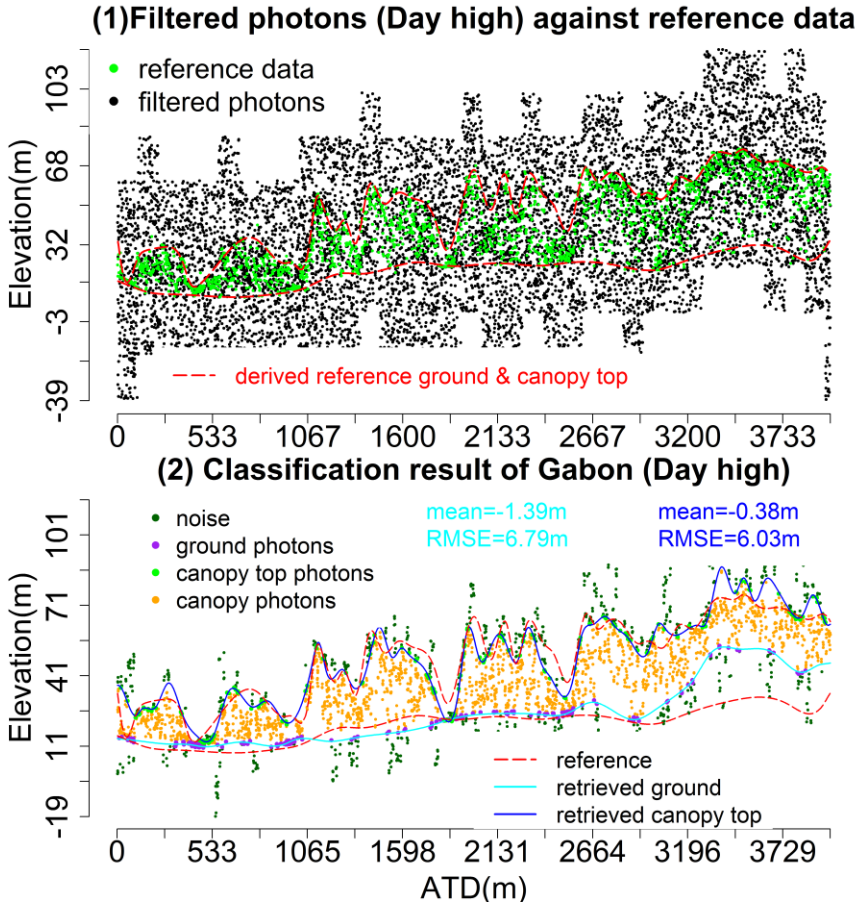


Figure V-13. Retrieved ground and canopy top results using simulated ICESat-2 data representing scenario acquired during the daytime with high noise level (Day high) in the Modah forest of Gabon.

Classification results of two other scenarios at the same study site are presented in Figure V-14 and Figure V-15. Results of filtered photons after the GBS filter showed that the Day low scenario of tropical forest in Gabon was indistinguishable from the Day high scenario through visually comparing Figure V-13(1) and Figure V-14(1). However, the subtle difference could be observed from Table V-3 with the photon density after the GBS filter of 2.65 and 2.60 photons/m, respectively. Additionally, the retrieved ground from the Day low scenario (Figure V-14) was more close to the reference ground with RMSE of 5.05 m, while the empty space in the understory of the region between 2700 and 4000 m still existed. For the accuracy of canopy top identification, it was just slightly better than the Day high scenario with RMSE of 5.69 m. In contrast, the nighttime scenario (Figure V-15) with higher SNR generated more accurate estimated ground and canopy top by comparing their corresponding reference data. As a consequence, more accurate retrieved ground and canopy top were achieved in terms of the RMSEs (3.05 m for ground vs. 3.77 m for the canopy top).

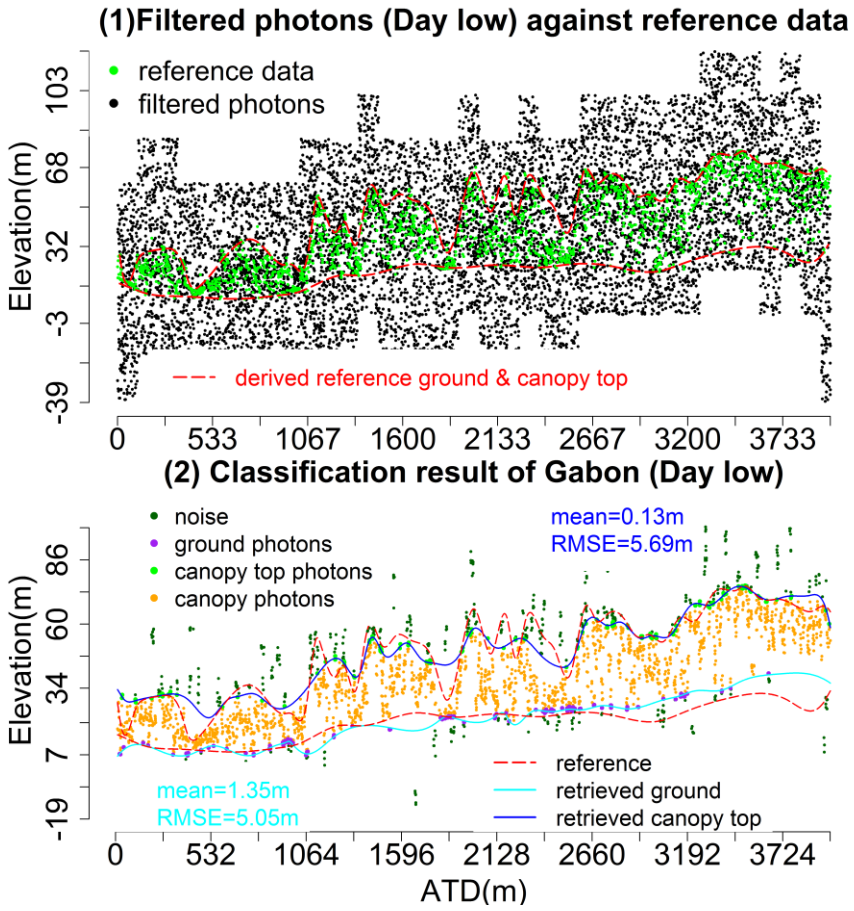


Figure V-14. Retrieved ground and canopy top results using simulated ICESat-2 data representing scenario acquired during the daytime with low noise level (Day low) in the Modah forest of Gabon.

A quantitative accuracy assessment of representative scenarios is summarized in Table V-3. The nighttime scenario outperformed the daytime scenario from the perspective of the average bias and RMSE. As expected, in all scenarios the RMSE for the retrieved ground and canopy decreased as the SNR of simulated ICESat-2 data increased. Table V-3 also reveals better performance results are expected when the reference photon density is high and terrain is mostly flat, in study sites such as the SHNF in Texas. Results for the Tennessee study site (temperate hardwood) were the least accurate which was consistent throughout all scenarios when comparing to other study sites. The main reason was that the raw photon density and photon density after the GBS

filter of this study site were the largest among all scenarios, especially when we just took the daytime scenarios into account (Table V-3).

Based on results of the remaining scenarios (102 scenarios) presented in Table S1, the nighttime conditions generates better results than the daytime scenario, and higher SNR level scenario gives smaller RMSE. The average RMSE of retrieved ground and canopy top for these nighttime scenarios was about 1.83 m and 2.70 m, respectively. By contrast, the accuracy for the estimated ground and canopy top using daytime datasets were lower with RMSE of 2.80 m and 3.59 m, respectively. Our results highlighted that large RMSE (> 5 m) were mainly occurring for daytime scenarios with sparse vegetation (reference density <1 photon/m).

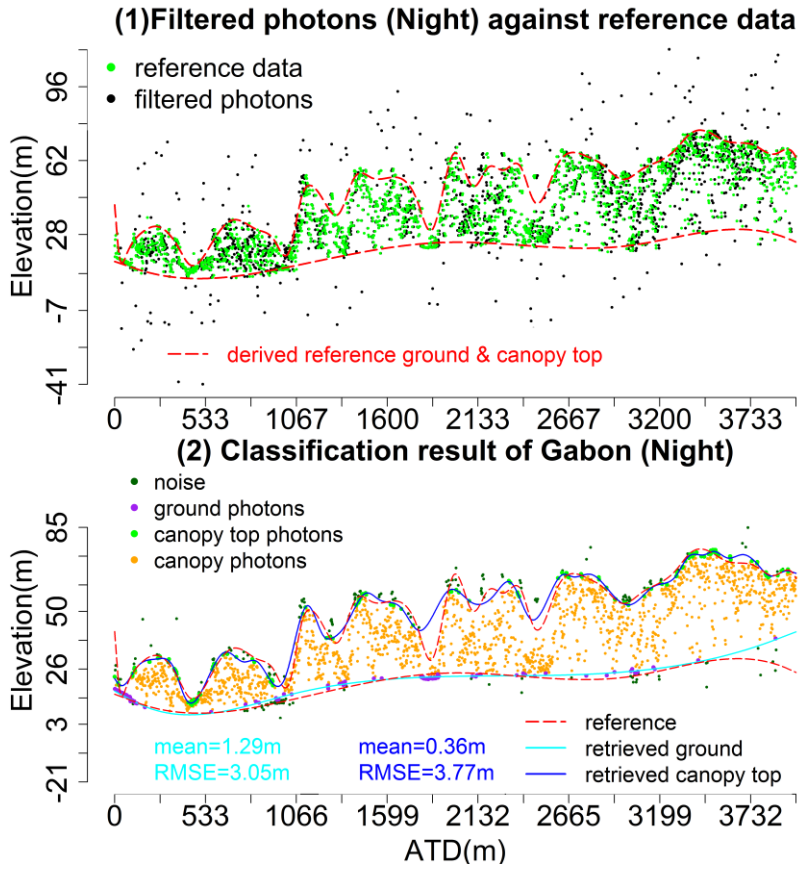


Figure V-15. Retrieved ground and canopy top results using simulated ICESat-2 data representing scenario acquired during the nighttime (Night) in the Modah forest of Gabon.

Descriptive statistics of the percentile heights' calibration results for representative vegetation conditions are summarized in Table V-4. We did not provide results for Alaska because there was no available reference percentile height information that could be used for the validation. As shown in Table V-4, nighttime scenarios were more prone to achieve smaller mean bias and RMSE than corresponding daytime scenarios. Compared to validation results of retrieved ground or canopy top (Table V-3), the percentile heights yielded better results with smaller RMSEs except for the results of the nighttime scenario at the SHNF. Interestingly, almost all RMSEs of percentile heights were likely to be inside the range of RMSE values of corresponding retrieved ground and TOC. Additionally, the higher the percentile height, the larger RMSEs were observed in most scenarios.

In summary, algorithm performance in different vegetation conditions showed that our methodology was capable of deriving ground, percentile heights and TOC from simulated ICESat-2 data with reasonable accuracies.

Table V-3. Validation results of retrieved ground and canopy top with representative vegetation conditions using simulated ICESat-2 data.

Study sites	Scenarios	Retrieved ground (m)			Retrieved canopy top (m)			Reference photon density	Raw photon density/SNR	Photon density after GBF/Cluster
		mean	SD	RMSE	mean	SD	RMSE			
Gabon	Day high	-1.39	6.65	6.79	-0.38	6.02	6.03	0.53	6.79 / 0.08	2.65 / 1.28
	Day low	1.35	4.86	5.05	0.13	5.69	5.69	0.53	6.75 / 0.09	2.60 / 1.21
	Night	1.29	2.75	3.05	0.36	3.75	3.77	0.53	0.70 / 3.13	0.60 / 0.53
TN	Day high	0.31	6.27	6.28	-2.41	9.27	9.58	2.18	34.09 / 0.07	12.87 / 5.10
	Day low	1.54	5.46	5.68	0.8	5.96	6.01	2.18	10.22 / 0.27	4.84 / 2.45
	Night	0.27	4.47	4.48	-0.2	5.17	5.18	2.18	2.34 / 13.62	2.22 / 2.05
FR, TX	Day high	0.09	0.62	0.64	-0.63	2.23	2.31	2.19	18.41 / 0.14	6.05 / 3.62
	Day low	0.11	0.55	0.57	-0.63	1.61	1.73	2.19	9.50 / 0.30	3.76 / 2.25
	Night	0.07	0.6	0.61	-1.15	2.1	2.4	2.19	2.51 / 6.84	2.39 / 2.18
SHNF, TX	Day 2004	-0.36	0.84	0.91	-0.79	3.3	3.39	3.37	20.25 / 0.2	8.79 / 5.26
	Night 2004	-0.11	0.45	0.46	-0.08	1.63	1.64	3.37	3.55 / 19.75	3.42 / 3.35
	Day 2010	-0.32	1.48	1.51	-0.61	2.19	2.27	3.37	19.08 / 0.22	8.31 /4.14
	Night 2010	-0.03	0.51	0.52	0.01	1.48	1.48	3.37	3.54 / 21.60	3.40 / 3.18
Alaska	Day high	-0.26	1.44	1.47	-1.03	2.24	2.46	1.22	1.86 / 1.91	1.39 / 1.19
	Night	-0.07	1.31	1.31	-0.87	1.63	1.85	1.22	1.24 / 61.35	1.23 / 1.15

Table V-4. Validation results of percentile heights (PHs) with representative condition using simulated ICESat-2 data.

Study sites	Scenarios	Descriptive Statistics	50th PH	60th PH	75th PH	80th PH	85th PH	90th PH	95th PH
Gabon	Day high	Mean bias	-0.78	-0.99	-0.58	-0.71	-0.87	-1.11	-1.02
		SD	4.37	3.92	4.72	4.86	5.03	5.35	5.54
		RMSE	4.41	4.02	4.72	4.88	5.07	5.43	5.59
	Day low	Mean bias	-1.10	-1.39	-1.60	-1.91	-2.26	-2.44	-2.56
		SD	4.37	4.15	4.53	5.04	4.98	5.36	5.23
		RMSE	4.48	4.35	4.77	5.35	5.44	5.85	5.79
	Night	Mean bias	0.34	-0.16	-0.73	-0.83	-0.90	-1.46	-1.64
		SD	3.11	3.04	3.54	3.34	3.45	4.05	4.19
		RMSE	3.11	3.03	3.59	3.42	3.54	4.28	4.48
TN	Day high	Mean bias	1.84	1.53	0.46	0.31	0.08	-0.23	0.20
		SD	6.80	7.33	7.04	7.46	7.46	8.01	8.01
		RMSE	6.96	7.39	6.97	7.37	7.37	7.92	7.91
	Day low	Mean bias	1.74	1.14	0.92	0.43	0.10	-0.16	-0.17
		SD	4.42	4.78	3.83	4.29	5.32	6.09	5.87
		RMSE	4.70	4.86	3.89	4.26	5.25	6.02	5.80
	Night	Mean bias	0.52	0.39	-0.51	-0.59	-0.83	-1.09	-0.85
		SD	5.03	4.67	4.15	4.52	4.83	5.71	6.16
		RMSE	5.00	4.62	4.13	4.50	4.84	5.74	6.14
FR,TX	Day high	Mean bias	-0.14	-0.21	-0.37	-0.39	-0.49	-0.58	-0.72
		SD	1.30	1.43	1.56	1.55	1.58	1.64	1.87
		RMSE	1.30	1.44	1.60	1.59	1.65	1.74	1.99
	Day low	Mean bias	-0.15	-0.12	-0.28	-0.30	-0.37	-0.43	-0.54
		SD	1.05	1.30	1.45	1.48	1.53	1.48	1.68
		RMSE	1.06	1.30	1.47	1.50	1.56	1.54	1.76
	Night	Mean bias	-0.15	-0.22	-0.32	-0.35	-0.40	-0.44	-0.63
		SD	1.29	1.45	1.56	1.51	1.57	1.67	1.91
		RMSE	1.29	1.46	1.58	1.54	1.61	1.72	2.00
SHNF, TX	2004 Day	Mean bias	-0.60	-0.64	-0.72	-0.71	-0.78	-0.82	-0.87
		SD	2.98	2.90	2.74	2.56	2.43	2.33	2.61
		RMSE	3.04	2.97	2.83	2.66	2.55	2.47	2.75
	2004 Night	Mean bias	-0.59	-0.59	-0.64	-0.76	-0.84	-1.01	-1.19
		SD	2.67	2.67	2.52	2.56	2.35	2.16	2.30
		RMSE	2.73	2.73	2.60	2.67	2.49	2.38	2.58
	2010 Day	Mean bias	-0.43	-0.64	-0.83	-0.81	-0.83	-1.04	-1.25
		SD	2.80	3.01	2.88	2.65	2.63	2.46	2.23
		RMSE	2.83	3.07	2.99	2.77	2.76	2.67	2.55
	2010 Night	Mean bias	-0.49	-0.56	-0.72	-0.65	-0.65	-0.85	-1.02
		SD	2.66	2.70	2.52	1.98	2.02	1.82	1.80
		RMSE	2.70	2.76	2.61	2.08	2.12	2.00	2.06

5.4 Discussion

The methodology proposed in the present paper provided encouraging forest structure and terrain mapping results from test-bed sensor data for the ICESat-2 mission. Our results are in agreement with findings of earlier studies (Gwenzi et al., 2016; Moussavi et al., 2014; Wang et al., 2016). Moreover, we built upon the findings of previous studies by investigating different SNR scenarios across various vegetation conditions to comprehensively assess the performance of our algorithm.

5.4.1 Noise filtering

Effective noise filtering is a pivotal step toward generating useful terrain and vegetation metrics from further ICESat-2 data. The main objective of noise filtering is to improve the classification results of photons by removing noise photons that may distort the analysis and hinder subsequent classification. For this study, we developed a framework with three multi-level noise filters and attempted to find the appropriate filtering parameters and self-adaptive approach to allow classification of ground and canopy photons under diverse terrain and vegetation conditions. One salient reason to employ the multi-level filtering approach is to preserve as much information contained in raw data and gradually reduce the effect of possible noise photons. In addition, two approaches for the first filtering step were implemented and tested. Specifically, the approach using the ancillary elevation data requires accurate geolocation of the original data, such as simulated ICESat-2 or MABLE data. Accurate geolocation is essential to ensure the extracted ancillary information (GCHM & GDEM) provides a realistic envelope to locate signal photons (Figure V-2(1)). Moreover, a competitive or even superior alternative to the ancillary data approach, named the GBS filter, has been developed as the second approach. This filter has been proved to work efficiently in all scenarios and is preferred for preliminary noise filtering to process ICESat-2 data. The second filter, named the cluster filter, is effective to remove the noise around the targets in most scenarios. However, it may be problematic when the returned photon density is relatively low, as is the case with the Gabon study site. A closer examination of these scenarios reveals that signal photons and noise photons are difficult to be discerned in the low returned photon density condition. We suspect that this problem is caused by the lack of an obvious clustering pattern in the low density of return photons scenario, which violates the assumption of the cluster filter that signal photons are more prone to cluster around the targets.

5.4.2 MABEL

Compared to the results of simulated ICESat-2 data, MABEL data generated less accurate results (Table V-2). Many factors contribute to the large bias of MABEL validation such as low geolocation precision of MABEL data (RMSE: ~30 m) (Hancock and Lee, 2014) and interpolation error (Moussavi et al., 2014). Undoubtedly, better validation results are anticipated when more accurate reference data are available as stated in a previous study (Gwenzi et al., 2016). The accuracy of MABEL data for retrieving the ground and canopy top is comparable to other studies, but it is hard to judge the performances of algorithms solely based on their mean bias and RMSEs. Due to inconsistencies with both reference data and MABEL data, the manual adjustment of the co-registration could introduce bias. However, we consider that using MABEL data in our analysis is useful as a test-bed sensor data for the ICESat-2 mission to explore the possible challenges of ATLAS data.

Along with the results of earlier studies, the retrieved ground is more accurate than the retrieved canopy top (Gwenzi et al., 2016; Moussavi et al., 2014; Tang et al., 2016). Several potential reasons may partly explain this discrepancy. For example, the reference data for MABEL come from high-resolution surface products (1 m) such as the CHM of G-LiHT, and a slight shift of geolocation could bring large discrepancies of extracted reference value, especially for the CHM. In addition, the temporal difference in acquisition dates of G-LiHT data and MABEL data, i.e., G-LiHT data was collected in August of 2011 while MABEL data used in this study was obtained in September of 2012 unarguably brings additional variability to validation results, especially for the validation of canopy height.

5.4.3 Simulated ICESat-2 data

Results of processing simulated ICESat-2 data suggest that our methodology is applicable for vegetation structure characterization and terrain mapping. Although we cannot extend with full certainty the performance of the methodology to genuine ATLAS data, results of different exploratory scenarios offer a good understanding of expectations in terms of accuracies obtainable for ground and vegetation canopy metrics. It is fully

expected that our methodology will be adjusted post-launch to reflect the on-orbit performance of ATLAS and serve as a baseline for further preprocessing ICESat-2 data for vegetation and terrain characterization or comparison with other algorithms. Moreover, exploring these scenarios can provide insights into the possible challenging issues of ICESat-2 data for vegetation studies and prepare for the ICESat-2 mission applications (Moussavi et al., 2014).

Comparing results of the daytime scenario with nighttime scenario (Figure V-11), more noise photons are present within the canopy mixed with real signal photons, which are difficult to be detected and may further lead to inaccurate estimated metrics of vegetation structure, such as percentile heights. In reality, the photons caused by solar background could also occur in the middle part of the vegetation canopy space. Therefore, identifying noise photons from the mid-story of vegetation canopy may be one aspect of further research investigations.

Among all examined scenarios, the tropical forest with lowest return photon density has the least accurate validation results. Both returned photon density and noise level influence the algorithms' performances to a significant degree. In particular, the appropriate trimming proportion (α) of the second filter that deletes noise photons around targets is adjusted based on the noise level of the dataset; the optimal value for α is 0.10 and 0.65 for nighttime and daytime scenarios in our case, respectively. Assuredly, finding the suitable range of α requires further testing of our algorithm over various vegetation conditions. According to the currently limited data sets available to us (115 scenarios), the reasonable range of α was found to be [0.05, 0.15] and [0.6, 0.7] for nighttime and daytime scenarios, respectively.

With regard to the return photon density, previous studies have demonstrated its importance and usefulness in ground and canopy top identification (Gwenzi et al., 2016; Zhang and Kerekes, 2015). Our study also confirms that with higher returned photon density, more accurate results are anticipated according to Table V-3 and Table V-5. Challenging situations could occur when the dense vegetation cover is present as demonstrated in Figure V-13. The accuracy of the ground for low return photon density

is heavily affected by the width of the moving window. In the present study, we employ larger window width (100 m vs. 50 m) for the tropical forest scenarios to preserve the spatial integrity of topography under dense vegetation. In addition, the kernel density method implemented for retrieving ground in low returned photon density is useful for optimizing the process of GP identification. Both designs further enhance the performance of our algorithm. As suggested by Gwenzi et al. (2016), a high repetition rate sensor possibly extenuates the impact of low return photon density on the retrieval of ground at dense vegetation and canopy height at the sparse vegetation cover. This aspect is of particular concern for the ecosystem community, since tropical forests concentrate high biomass values that need to be estimated with high accuracy and precision.

Percentile heights are key parameters for characterizing canopy structure and monitoring biomass and carbon dynamics (Falkowski et al., 2009; Popescu, 2007; Swatantran et al., 2016; Zhao et al., 2011). Thus, we conducted accuracy assessment of percentile heights to further evaluate the performance of the methodology. Validation of percentile heights showed similar accuracies as the retrieved ground and canopy top, indicating that our methodology can achieve high precision results. More accurate percentile height results are yielded in the study site with low return photon density or high-level noise such as in Gabon and TN study sites, which mainly result from that the photons in the middle part have higher chance to be correctly identified as signals when comparing to the identified ground and canopy top photons.

To our best knowledge, this study expands on early similar work on developing and testing algorithms of vegetation characterization for the ICESat-2 mission, with encouraging calibration results. Given the availability of test bed data, only limited scenarios are investigated in this study. It is recommended that more simulated ICESat-2 data scenarios should be explored in future studies before the genuine ATLAS data are available. A crucially important characteristic of a successful methodology is its ability to self-regulate processing parameters based on conditions exhibited within datasets. Minimizing the customized parameters to enable automatic and adaptive processing for various terrain and vegetation conditions is thus preferred. With this study, we presented

a methodology for optimizing processing parameters, such as the trimming proportion, the window width and quantile threshold, to enable an adaptive and efficient ICESat-2 type of data processing. Certainly, the current methodology is not guaranteed to reliably extract terrain and characterize vegetation structure for all global dataset conditions. Future research investigations could look into the possibility of increasing the synergy between ICESat-2 data and other satellite missions, such as Landsat, to further refine processing parameters and the self-adaptive capability to meet the needs of the ecosystem science community for characterizing vegetation structure, biomass, and carbon with high accuracy.

5.5 Conclusion

ICESat-2 holds great potential to enhance estimates of forest biomass, carbon and volume through acquiring synoptic measurements of vegetation canopy height, density, the vertical distribution of photosynthetically active material over vast expanses e.g., nations, continents and globe. The present study proposes a useful framework to process ICESat-2 like data, e.g., MABEL data and simulated ICESat-2 data, to derive the terrain elevation and vegetation structure metrics. In addition, a realistic understanding of possible challenges related to processing upcoming ICESat-2 data over different vegetation conditions is provided. The multi-level noise filters show their effectiveness in identifying the possible signal photons and noise photons under different scenarios. Our validation results of terrain and vegetation structure metrics demonstrate that the proposed methodology is capable of meeting the science objectives of the ICESat-2 mission on measuring vegetation canopy height and supporting ongoing biomass mapping over a large scale. Despite that validation results obtained over densely vegetated conditions are not impressive, this study facilitated our understating of possible issues related to data processing. The current methodology can provide a valuable basis to characterize terrain and vegetation structure that can be adjusted after launching ICESat-2 to derive useful vegetation metrics over large areas. Due to the fuzzy margin between noise photons and ground or canopy top photons of ATLAS data, adopting the

uncertainty concept to classify the photons of ground and canopy top could be more intuitive and possibly have potential for further research investigations. Certainly, more efforts and scenarios of simulated ICESat-2 data are needed to further enhance their utility for canopy structure and terrain characterization. Given the profile configuration of ICESat-2 data, incorporating them with other data such as optical images, airborne LiDAR, radar and upcoming GEDI data can expand their potential applications and increase the accuracy of estimated parameters. In addition, future studies should also place an emphasis on generating gridded products using the dense orbit pattern of ICESat-2 data.

5.6 Appendix

Table V-5. Complete validation results of retrieved ground and canopy top using simulated ICESat-2 data with detailed background information. (density unit: photons/m)

Scenarios	Raw density	After GBS density	After cluster density	Reference Density	Ground (m)			Canopy top (m)		
					Mean	SD	RMSE	Mean	SD	RMSE
day	39.75	11.46	5.61	3.97	-0.33	1.25	1.30	-0.17	1.25	1.26
night	5.20	4.23	3.99	3.94	0.37	0.81	0.89	-0.79	0.96	1.24
day	20.31	8.74	4.80	0.81	-0.97	2.97	3.13	-0.07	6.81	6.81
night	1.38	1.09	1.01	0.88	-1.89	1.89	2.67	-0.44	4.63	4.65
day	58.63	14.06	6.88	4.82	-0.56	2.32	2.39	-0.86	2.00	2.18
night	6.58	5.13	4.85	4.80	0.09	0.79	0.79	-1.00	1.84	2.10
day	23.33	7.08	3.90	0.63	0.39	7.08	7.09	-1.00	6.47	6.54
night	1.44	0.93	0.87	0.72	-1.27	3.54	3.76	-0.82	3.37	3.47
day	42.01	17.39	8.52	6.01	-0.92	1.05	1.39	-3.52	1.80	3.95
night	7.16	6.39	6.05	6.00	-0.24	0.36	0.44	-2.52	1.41	2.88
day	23.06	7.09	3.89	0.63	1.72	7.22	7.42	0.24	5.36	5.37
night	1.43	0.93	0.85	0.72	-0.92	3.06	3.20	-0.11	4.13	4.13
day	51.43	10.85	5.29	3.62	-1.17	2.47	2.74	-1.60	2.31	2.81
night	5.29	3.97	3.79	3.72	0.10	1.45	1.45	-2.40	1.53	2.85
day	23.06	7.01	3.85	0.63	1.72	6.08	6.32	2.77	7.27	7.78
night	1.42	0.92	0.84	0.71	-0.75	3.70	3.77	-0.74	4.65	4.70
day	22.21	9.79	4.79	3.42	-0.71	0.82	1.09	-3.14	2.09	3.77
night	4.04	3.64	3.44	3.40	-0.20	0.32	0.38	-2.46	1.50	2.88
day	25.34	8.02	4.40	0.72	1.30	6.47	6.60	1.10	6.61	6.70
night	1.60	1.05	0.97	0.81	-1.39	3.95	4.19	0.70	3.20	3.27
day	26.34	14.86	7.28	5.09	0.10	0.47	0.48	-0.49	1.14	1.24
night	5.60	5.33	5.05	5.02	0.20	0.65	0.68	-0.60	0.89	1.07
day	26.53	7.93	4.33	0.71	2.60	8.12	8.53	1.36	7.25	7.38
night	1.58	1.03	0.94	0.80	-0.97	4.23	4.34	0.12	3.35	3.35
day	24.13	14.17	6.92	4.93	0.00	0.13	0.13	-0.49	0.67	0.84
night	5.40	5.16	4.88	4.85	0.00	0.12	0.12	-0.41	0.48	0.63
day	23.95	12.41	6.05	4.13	-0.31	1.53	1.56	-0.92	2.45	2.62
night	4.74	4.47	4.24	4.20	0.13	1.29	1.30	-1.85	1.80	2.58
day	17.57	10.29	5.04	3.59	0.02	0.09	0.09	-0.57	0.59	0.83
night	3.91	3.74	3.54	3.53	0.01	0.08	0.08	-0.43	0.44	0.62
day	18.85	10.08	4.92	3.37	-0.84	0.95	1.27	-2.00	2.90	3.53
night	3.84	3.62	3.43	3.40	-0.21	0.30	0.36	-1.87	1.97	2.72
day	16.36	9.48	4.62	3.22	-0.14	0.22	0.26	-0.66	0.82	1.05
night	3.52	3.36	3.16	3.15	-0.09	0.10	0.13	-0.66	0.70	0.96
day	24.30	12.54	6.14	4.16	0.15	2.41	2.42	-0.87	2.50	2.65
night	4.73	4.45	4.22	4.19	0.29	2.15	2.17	-0.84	2.05	2.21
day	19.68	11.56	5.65	3.99	-0.11	0.15	0.18	-1.44	0.96	1.73
night	4.34	4.16	3.92	3.91	-0.09	0.12	0.15	-0.87	0.74	1.14
day	17.58	9.66	4.73	3.33	-0.11	0.53	0.55	-0.81	2.53	2.66
night	3.74	3.54	3.34	3.32	-0.01	0.23	0.23	-0.79	1.98	2.13
day	21.40	12.26	5.97	4.11	-0.11	0.28	0.30	-0.67	1.28	1.45
night	4.53	4.31	4.07	4.05	-0.03	0.14	0.14	-0.84	0.98	1.29
day	15.82	8.41	4.62	2.85	-0.79	2.27	2.40	0.66	3.46	3.52

Table V-5 Continued

night	3.17	3.00	2.84	2.82	0.19	1.34	1.36	-0.43	1.37	1.44
day	27.86	16.00	7.83	5.48	-0.15	0.26	0.30	-1.28	1.22	1.77
night	5.98	5.72	5.41	5.38	-0.08	0.10	0.12	-1.15	0.98	1.51
day	18.29	10.37	5.06	3.51	0.38	1.15	1.21	-0.31	1.27	1.31
night	3.89	3.70	3.51	3.49	0.69	1.19	1.37	-0.97	0.98	1.38
day	39.82	12.60	6.12	4.38	1.01	4.26	4.38	-1.02	2.30	2.51
night	6.11	5.24	4.96	4.88	1.67	3.04	3.47	-1.32	2.23	2.60
day	16.24	8.17	4.48	2.79	-2.41	2.58	3.53	1.58	3.36	3.72
night	3.23	3.01	2.86	2.82	-0.10	1.12	1.12	-1.77	1.59	2.38
day	35.23	13.38	6.53	5.13	-0.20	1.21	1.23	-1.62	3.81	4.14
night	6.30	5.59	5.30	5.25	-0.25	1.05	1.08	-0.72	3.34	3.41
day	22.69	11.29	5.53	3.77	-0.01	2.10	2.10	-1.00	3.52	3.66
night	4.48	4.17	3.94	3.89	0.31	1.36	1.40	-1.21	2.55	2.82
day	33.23	11.07	5.42	4.00	1.23	3.32	3.54	-1.25	2.08	2.43
night	5.34	4.62	4.39	4.32	1.38	2.39	2.76	-1.32	1.94	2.34
day	22.04	10.93	5.34	3.72	-0.18	1.59	1.60	-1.06	2.70	2.90
night	4.36	4.06	3.83	3.80	0.10	1.29	1.29	-1.31	2.06	2.44
day	36.25	10.03	4.91	3.48	1.73	5.10	5.39	-0.56	2.02	2.09
night	4.91	4.09	3.86	3.80	2.63	4.65	5.34	-1.35	1.73	2.19
day	19.54	9.50	4.63	3.08	1.54	3.03	3.40	-0.78	2.77	2.88
night	3.76	3.49	3.30	3.26	1.43	3.00	3.33	-1.60	1.96	2.53
day	34.43	14.59	7.14	5.42	1.29	4.02	4.22	-1.20	2.68	2.94
night	6.67	6.05	5.73	5.67	0.88	2.92	3.05	-1.12	1.97	2.27
day	33.51	11.31	5.54	3.57	-1.58	2.27	2.77	-2.50	4.36	5.02
night	4.75	4.04	3.82	3.74	-0.05	1.14	1.14	-2.31	3.80	4.45
day	16.94	9.59	4.67	3.67	0.33	1.37	1.41	-0.88	1.73	1.95
night	4.10	3.92	3.73	3.71	0.16	0.92	0.93	-0.85	1.34	1.59
day	35.30	10.62	5.17	3.43	-1.77	2.36	2.95	-4.69	5.01	6.86
night	4.66	3.87	3.67	3.60	-0.45	1.31	1.39	-2.29	4.45	5.00
day	24.07	13.16	6.45	5.20	-0.15	0.39	0.42	-0.61	1.63	1.74
night	5.70	5.43	5.14	5.12	-0.04	0.16	0.17	-0.40	1.32	1.38
day	59.02	13.04	6.38	4.11	-2.00	3.30	3.86	-4.22	5.96	7.30
night	6.32	4.80	4.54	4.45	-0.69	1.94	2.06	-2.02	4.50	4.93
day	20.60	11.17	5.47	4.31	-0.54	0.77	0.94	-1.51	3.19	3.53
night	4.84	4.60	4.36	4.33	-0.37	0.54	0.66	-0.83	3.17	3.28
day	57.44	13.02	6.37	4.14	-1.55	3.05	3.42	-3.46	5.21	6.25
night	6.33	4.84	4.58	4.48	-0.39	2.30	2.33	-1.82	4.80	5.14
day	16.60	9.29	4.55	3.51	0.18	1.29	1.30	-1.84	2.62	3.20
night	3.96	3.78	3.58	3.56	0.22	1.26	1.28	-1.37	1.82	2.28
day	39.26	11.06	5.38	3.50	-1.21	2.88	3.13	-3.58	3.54	5.04
night	5.01	4.05	3.82	3.74	-0.24	1.94	1.95	-2.74	3.01	4.07
day	29.74	16.36	8.01	6.36	0.10	0.96	0.97	-0.91	2.23	2.41
night	7.00	6.69	6.34	6.31	0.05	0.81	0.81	-0.77	2.01	2.16
day	23.85	12.57	6.13	4.27	-0.24	0.84	0.87	-0.34	1.87	1.90
night	4.77	4.50	4.27	4.24	-0.11	0.57	0.58	-0.41	1.43	1.49
day	15.35	8.53	4.69	3.25	1.48	2.39	2.81	0.44	2.23	2.27
night	3.71	3.52	3.32	3.28	2.29	2.34	3.28	-1.66	1.15	2.02
day	28.07	14.70	7.17	5.01	-0.15	1.32	1.33	-0.34	1.89	1.92
night	5.66	5.32	5.04	5.01	-0.10	0.86	0.87	-0.27	1.77	1.79
day	46.62	16.06	7.83	5.95	0.06	2.94	2.94	-0.93	2.22	2.40
night	7.54	6.59	6.25	6.18	-0.09	2.69	2.69	-1.05	1.86	2.14

Table V-5 Continued

day	22.87	9.44	4.61	0.78	-1.05	7.83	7.89	1.64	4.83	5.10
night	1.54	1.19	1.10	0.96	-1.88	3.08	3.61	0.68	3.59	3.65
day	35.58	16.80	8.22	6.43	-0.26	2.12	2.14	-0.90	3.54	3.65
night	7.43	6.91	6.56	6.50	-0.22	1.96	1.98	-0.01	3.45	3.45
day	19.45	8.17	4.51	0.79	-1.57	3.17	3.54	-2.41	4.95	5.50
night	1.31	1.03	0.95	0.83	-1.05	1.77	2.06	-1.69	3.23	3.65
day	32.35	13.93	6.81	5.19	0.74	3.48	3.56	-1.46	3.13	3.46
night	6.31	5.72	5.42	5.35	0.17	3.18	3.18	-1.35	2.52	2.86

CHAPTER VI
CONCLUSIONS AND FURTHER WORK

FW LiDAR bears great potential to provide more information for vegetation characterization than DR LiDAR due to the complete digitization of the returned signal that can be achieved instead of simple time measurements for individual pulses within the returned signal. In this dissertation, we first developed a novel algorithm, named the Gold deconvolution, for processing FW LiDAR data to extract information such as amplitudes and echo widths for terrain and vegetation characterization. The algorithms developed in this study is of both scientific and practical significance in that it provides dedicated and non-proprietary alternative tools with broad ecosystem researchers for extensive applications of FW LiDAR data. In addition, we conducted comparisons of the new method with existing FW LiDAR processing methods such as the Gaussian decomposition and RL deconvolution to comprehensively explore advantages and limitations of various waveform processing techniques to derive topography and canopy height information.

This study also applied the Bayesian non-linear modeling concept to process small-footprint FW LiDAR data and provided a new insight into the waveform decomposition and uncertainty estimation. Through the Bayesian decomposition, uncertainty at the parameter estimates, point cloud and surface model generation steps were quantified in a probabilistic sense. Moreover, uncertainty estimates from the Bayesian method enhanced the credibility of decomposition results to capture the true error of estimates and traced the uncertainty propagation along the processing steps.

A plethora of information contained in FW LiDAR data also offers prospects for real-world applications such as characterizing vegetation structures and tree species classification. We successfully integrated machine learning methods (the RF and CF) and Bayesian method with FW metrics to identify tree species with FW LiDAR data alone. The CF method introduced in this study rendered a new perceptive for variable selection

by overcoming waveform metrics selection bias caused by the RF method which violates the implicit null hypothesis and favors correlated metrics, and enhanced the accuracy of subsequent classification. Both machine learning methods (the RF and CF) and the Bayesian method generated satisfactory overall accuracy and the Bayesian method slightly outperformed the other two methods. Moreover, uncertainty estimates from the Bayesian method rendered users with more confidence for interpreting and applying classification results to real-world tasks such as forest inventory. It is expected that a concomitant expansion of adopting advanced statistical methods such as the Bayesian method for tackling complicated relationships between characteristics of interest (e.g. height, biomass, carbon) and remotely-sensed predictors, and assisting result interpretation and decision making in real-world applications with advances in computational capacity and handy operational tools.

It is envisioned that a continuing interest is arising in employing upcoming ICESat-2 data for biomass estimation and vegetation mapping over large areas. The mythological framework developed in this study explored the possible challenges of ICESat-2 data processing and provided a basis for further method adjustment for processing genuine ICESat-2 data.

Built on the outcomes of this dissertation, more attention should be paid to the following two main research directions: (1) exploiting FW LiDAR data for real-world applications such as biomass estimation and vegetation mapping with advanced statistical models (machine learning and Bayesian methods) and the development of more handy tools and software implementations; (2) incorporating ICESat-2 data with existing remote sensing data such as LiDAR and optical imagery for improved global canopy height and terrain elevation estimations, and forest structure and biomass mapping over large areas.

REFERENCES

Acevedo, M.A., Corrada-Bravo, C.J., Corrada-Bravo, H., Villanueva-Rivera, L.J., & Aide, T.M. (2009). Automated classification of bird and amphibian calls using machine learning: A comparison of methods. *Ecological Informatics, 4*, 206-214

Allouis, T., Durrieu, S., Véga, C., & Couteron, P. (2013). Stem volume and above-ground biomass estimation of individual pine trees from LiDAR data: Contribution of full-waveform signals. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6*, 924-934

Azadbakht, M., Fraser, C., & Khoshelham, K. (2016). A Sparsity-Based Regularization Approach for Deconvolution of Full-Waveform Airborne Lidar Data. *Remote Sensing, 8*, 648

Babcock, C., Finley, A.O., Cook, B.D., Weiskittel, A., & Woodall, C.W. (2016). Modeling forest biomass and growth: Coupling long-term inventory and LiDAR data. *Remote Sensing of Environment, 182*, 1-12

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing, 114*, 24-31

Beucher, S., & Meyer, F. (1992). The morphological approach to segmentation: the watershed transformation. *Optical Engineering-New York-Marcel Dekker Incorporated-*, *34*, 433-433

Blair, J.B., Rabine, D.L., & Hofton, M.A. (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*, 115-122

Boudreau, J., Nelson, R., Margolis, H., Beaudoin, A., Guindon, L., & Kimes, D. (2008). Regional aboveground forest biomass using airborne and spaceborne LiDAR in Québec. *Remote Sensing of Environment, 112*, 3876-3890

Briese, C., Höfle, B., Lehner, H., Wagner, W., Pfennigbauer, M., & Ullrich, A. (2008). Calibration of full-waveform airborne laser scanning data for object classification. In, *SPIE Defense and Security Symposium* (pp. 69500H-69500H-69508): International Society for Optics and Photonics

Buerkner, P.-C. (2016). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*, 1-28

Cao, L., Coops, N.C., Innes, J.L., Dai, J., Ruan, H., & She, G. (2016). Tree species classification in subtropical forests using small-footprint full-waveform LiDAR data. *International Journal of Applied Earth Observation and Geoinformation, 49*, 39-51

Carlsson, T., Steinvall, O., & Letalick, D. (2001). Signature simulation and signal analysis for 3-D laser radar. *Month, 4*, C4ISR

Cawse-Nicholson, K., van Aardt, J., Hagstrom, S., Romanczyk, P., Schaaf, C., Strahler, A., Li, Z., & Krause, K. (2014). Improving waveform lidar processing toward robust deconvolution of signals for improved structural assessments. In, *SPIE Defense+ Security* (pp. 90800I-90800I-90806): International Society for Optics and Photonics

Chauve, A., Mallet, C., Bretar, F., Durrieu, S., Deseilligny, M.P., & Puech, W. (2007). Processing full-waveform lidar data: modelling raw signals. In, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 2007* (pp. 102-107)

Chauve, A., Vega, C., Durrieu, S., Bretar, F., Allouis, T., Pierrot Deseilligny, M., & Puech, W. (2009). Advanced full-waveform lidar data echo detection: Assessing quality of derived terrain and tree height models in an alpine coniferous forest. *International Journal of Remote Sensing, 30*, 5211-5228

Chen, Q. (2007). Airborne lidar data processing and information extraction. *Photogrammetric engineering and remote sensing, 73*, 109

Chen, Q., Baldocchi, D., Gong, P., & Kelly, M. (2006). Isolating individual trees in a savanna woodland using small footprint lidar data. *Photogrammetric Engineering & Remote Sensing, 72*, 923-932

Chen, Q., Gong, P., Baldocchi, D., & Tian, Y.Q. (2007). Estimating basal area and stem volume for individual trees from lidar data. *Photogrammetric Engineering & Remote Sensing, 73*, 1355-1365

Chen, Q., Vaglio Laurin, G., & Valentini, R. (2015). Uncertainty of remotely sensed aboveground biomass over an African tropical forest: Propagating errors from trees to plots to pixels. *Remote Sensing of Environment, 160*, 134-143

Chen, W., Li, X., Wang, Y., Chen, G., & Liu, S. (2014). Forested landslide detection using LiDAR data and the random forest algorithm: A case study of the Three Gorges, China. *Remote Sensing of Environment, 152*, 291-301

Chhatkuli, S., Mano, K., Kogure, T., Tachibana, K., & Shimamura, H. (2012). Full waveform lidar exploitation technique and its evaluation in the mixed forest hilly

region. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 505-509

Cook, B., Corp, L., Nelson, R., Middleton, E., Morton, D., McCorkel, J., Masek, J., Ranson, K., Ly, V., & Montesano, P. (2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. *Remote Sensing, 5*, 4045-4066

Dabney, P., Harding, D., Abshire, J., Huss, T., Jodor, G., Machan, R., Marzouk, J., Rush, K., Seas, A., & Shuman, C. (2010). The slope imaging multi-polarization photon-counting lidar: Development and performance results. In, *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE international* (pp. 653-656): IEEE

De Lannoy, G.J.M., Reichle, R.H., & Vrugt, J.A. (2014). Uncertainty quantification of GEOS-5 L-band radiative transfer model parameters using Bayesian inference and SMOS observations. *Remote Sensing of Environment, 148*, 146-157

Denham, R., Mengersen, K., & Witte, C. (2009). Bayesian analysis of thematic map accuracy data. *Remote Sensing of Environment, 113*, 371-379

Desquilbet, L., & Mariotti, F. (2010). Dose-response analyses using restricted cubic spline functions in public health research. *Stat Med, 29*, 1037-1057

Doneus, M., Briese, C., Fera, M., & Janner, M. (2008). Archaeological prospection of forested areas using full-waveform airborne laser scanning. *Journal of Archaeological Science, 35*, 882-893

Drake, J.B., Dubayah, R.O., Clark, D.B., Knox, R.G., Blair, J.B., Hofton, M.A., Chazdon, R.L., Weishampel, J.F., & Prince, S. (2002). Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment, 79*, 305-319

Edwards, M., Huet, S., Goreaud, F., & Deffuant, G. (2003). Comparing an individual-based model of behaviour diffusion with its mean field aggregate approximation. *Journal of Artificial Societies and Social Simulation, 6*

Ellison, A.M. (2004). Bayesian inference in ecology. *Ecology Letters, 7*, 509-520

Elzhov, T.V., Mullen, K.M., Spiess, A.-N., & Bolker, B. (2013). minpack. lm: R interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds. R package version 1.1-8. In

Falkowski, M.J., Evans, J.S., Martinuzzi, S., Gessler, P.E., & Hudak, A.T. (2009). Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. *Remote Sensing of Environment, 113*, 946-956

Fieber, K.D., Davenport, I.J., Tanase, M.A., Ferryman, J.M., Gurney, R.J., Becerra, V.M., Walker, J.P., & Hacker, J.M. (2015). Validation of Canopy Height Profile methodology for small-footprint full-waveform airborne LiDAR data in a discontinuous canopy environment. *ISPRS Journal of Photogrammetry and Remote Sensing, 104*, 144-157

Finley, A.O., Banerjee, S., Cook, B.D., & Bradford, J.B. (2013). Hierarchical Bayesian spatial models for predicting multiple forest variables using waveform LiDAR, hyperspectral imagery, and large inventory datasets. *International Journal of Applied Earth Observation and Geoinformation, 22*, 147-160

Fish, D., Brinicombe, A., Pike, E., & Walker, J. (1995). Blind deconvolution by means of the Richardson–Lucy algorithm. *JOSA A, 12*, 58-65

Frazer, G.W., Magnussen, S., Wulder, M.A., & Niemann, K.O. (2011). Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sensing of Environment, 115*, 636-649

Freeman, E.A., Frescino, T.S., & Moisen, G.G. (2016). Pick Your Flavor of Random Forest.

Fritz, H., Garcıa-Escudero, L.A., & Mayo-Iscar, A. (2012). tclust: An r package for a trimming approach to cluster analysis. *Journal of Statistical Software, 47*, 1-26

Frühwirth-Schnatter, S., & Frühwirth, R. (2016). Bayesian inference in the multinomial logit model. *Austrian Journal of Statistics, 41*, 27-43

Gallegos, M.T., & Ritter, G. (2005). A robust method for cluster analysis. *Annals of statistics*, 347-380

Gao, S., Niu, Z., Sun, G., Zhao, D., Jia, K., & Qin, Y. (2015). Height Extraction of Maize Using Airborne Full-Waveform LIDAR Data and a Deconvolution Algorithm. *IEEE Geoscience and Remote Sensing Letters, 12*, 1978-1982

García, M., Riaño, D., Chuvieco, E., & Danson, F.M. (2010). Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sensing of Environment, 114*, 816-830

Gelfand, A.E., & Smith, A.F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association, 85*, 398-409

Gelman, A., Carlin, J.B., Stern, H.S., & Rubin, D.B. (2014). *Bayesian data analysis*. Chapman & Hall/CRC Boca Raton, FL, USA

Gelman, A., Lee, D., & Guo, J. (2015). Stan: A Probabilistic Programming Language for Bayesian Inference and Optimization. *Journal of Educational and Behavioral Statistics, 40*, 530-543

Girardeau-Montaut, D. (2015). Cloud Compare—3D Point Cloud and Mesh Processing Software. *Open Source Project*

Gleason, C.J., & Im, J. (2012). Forest biomass estimation from airborne LiDAR data using machine learning approaches. *Remote Sensing of Environment, 125*, 80-91

Glenn, N.F., Neuenschwander, A., Vierling, L.A., Spaete, L., Li, A., Shinneman, D.J., Pilliod, D.S., Arkle, R.S., & McIlroy, S.K. (2016). Landsat 8 and ICESat-2: Performance and potential synergies for quantifying dryland ecosystem vegetation cover and biomass. *Remote Sensing of Environment, 185*, 233-242

Gouveia, W.P., & Scales, J.A. (1998). Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis. *Journal of Geophysical Research: Solid Earth, 103*, 2759-2779

Griewank, A., & Walther, A. (2008). *Evaluating derivatives: principles and techniques of algorithmic differentiation*. Siam

Gwenzi, D., & Lefsky, M.A. (2014). Modeling canopy height in a savanna ecosystem using spaceborne lidar waveforms. *Remote Sensing of Environment, 154*, 338-344

Gwenzi, D., Lefsky, M.A., Suchdeo, V.P., & Harding, D.J. (2016). Prospects of the ICESat-2 laser altimetry mission for savanna ecosystem structural studies based on airborne simulation data. *ISPRS Journal of Photogrammetry and Remote Sensing, 118*, 68-82

Hancock, D.W., & Lee, J.E. (2014). Mabel Release 010 Software Change and Release Note. In. http://icesat.gsfc.nasa.gov/icesat2/data/mabel/docs/MABEL_Release_010_Note.pdf

Hancock, S., Anderson, K., Disney, M., & Gaston, K.J. (2017). Measurement of fine-spatial-resolution 3D vegetation structure with airborne waveform lidar: Calibration and validation with voxelised terrestrial lidar. *Remote Sensing of Environment, 188*, 37-50

Hancock, S., Armston, J., Li, Z., Gaulton, R., Lewis, P., Disney, M., Mark Danson, F., Strahler, A., Schaaf, C., Anderson, K., & Gaston, K.J. (2015). Waveform lidar over vegetation: An evaluation of inversion methods for estimating return energy. *Remote Sensing of Environment, 164*, 208-224

Hancock, S., Lewis, P., Disney, M., Foster, M., & Muller, J. (2008). Assessing the accuracy of forest height estimation with long pulse waveform lidar through Monte-Carlo ray tracing. *Edinburgh: Silvilaser. September, 17, 18*

Harding, D.J. (2005). ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure. *Geophysical Research Letters, 32*

Harding, D.J., & Carabajal, C.C. (2005). ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure. *Geophysical Research Letters, 32*

Harsdorf, S., & Reuter, R. (2000). Stable deconvolution of noisy lidar signals. *tc*, *10*, 1

Heinzel, J., & Koch, B. (2011). Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation, 13*, 152-160

Heinzel, J.N., Weinacker, H., & Koch, B. (2008). Full automatic detection of tree species based on delineated single tree crowns–a data fusion approach for airborne laser scanning data and aerial photographs. *Proceedings of SilviLaser, 2008*, 8th

Hermosilla, T., Coops, N.C., Ruiz, L.A., & Moskal, L.M. (2014a). Deriving pseudo-vertical waveforms from small-footprint full-waveform LiDAR data. *Remote Sensing Letters, 5*, 332-341

Hermosilla, T., Ruiz, L.A., Kazakova, A.N., Coops, N.C., & Moskal, L.M. (2014b). Estimation of forest structure and canopy fuel parameters from small-footprint full-waveform LiDAR data. *International Journal of Wildland Fire, 23*, 224

Hernandez-Marin, S., Wallace, A.M., & Gibson, G.J. (2008). Multilayered 3D LiDAR image construction using spatial models in a Bayesian framework. *IEEE transactions on pattern analysis and machine intelligence, 30*, 1028-1040

Herzfeld, U.C., McDonald, B.W., Wallin, B.F., Neumann, T.A., Markus, T., Brenner, A., & Field, C. (2014). Algorithm for detection of ground and canopy cover in micropulse photon-counting lidar altimeter data in preparation for the ICESat-2 mission. *IEEE Transactions on Geoscience and Remote Sensing, 52*, 2109-2125

Hoff, P.D. (2009). *A first course in Bayesian statistical methods*. Springer Science & Business Media

Hoffman, M.D., & Gelman, A. (2014). The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research, 15*, 1593-1623

Hofton, M.A., Minster, J.B., & Blair, J.B. (2000). Decomposition of laser altimeter waveforms. *IEEE Transactions on Geoscience and Remote Sensing,, 38*, 1989-1996

Hollaus, M., Mücke, W., Höfle, B., Dorigo, W., Pfeifer, N., Wagner, W., Bauerhansl, C., & Regner, B. (2009a). Tree species classification based on full-waveform airborne laser scanning data. *Proceedings of SilviLaser*, 54-62

Hollaus, M., Mücke, W., Höfle, B., Dorigo, W., Pfeifer, N., Wagner, W., Bauerhansl, C., & Regner, B. (2009b). Tree species classification based on full-waveform airborne laser scanning data. *Proceedings of Silvilaser 2009*, 54-62

Holmgren, J., & Persson, Å. (2004). Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment, 90*, 415-423

Holmgren, J., Persson, Å., & Söderman, U. (2008). Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. *International Journal of Remote Sensing, 29*, 1537-1552

Hong, T., & Sen, M.K. (2009). A new MCMC algorithm for seismic waveform inversion and corresponding uncertainty analysis. *Geophysical Journal International, 177*, 14-32

Hovi, A. (2015). Towards an enhanced understanding of airborne LiDAR measurements of forest vegetation. *Dissertationes Forestales, 2015*

Ioannides, M., Arnold, D., Niccolucci, F., & Mania, K. (2006). Digital terrain modelling for archaeological interpretation within forested areas using full-waveform laserscanning

Isenburg, M. (2012). LAStools-Efficient tools for LiDAR processing. *Available at: http: http://www. cs. unc. edu/~ isenburg/lastools/[Accessed October 9, 2012]*

Jalobeanu, A., & Gonçalves, G. (2014). Robust ground peak extraction with range error estimation using full-waveform LiDAR. *IEEE Geoscience and Remote Sensing Letters, 11*, 1190-1194

Jones, T.G., Coops, N.C., & Sharma, T. (2010). Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. *Remote Sensing of Environment, 114*, 2841-2852

Jutzi, B., & Stilla, U. (2006). Range determination with waveform recording laser systems using a Wiener Filter. *ISPRS Journal of Photogrammetry and Remote Sensing, 61*, 95-107

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B.M., & Wu, J.-C. (2012). An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. *Remote Sensing, 4*, 950-974

Kampe, T.U. (2010). NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. *Journal of Applied Remote Sensing, 4*, 043510

Karlson, M., Ostwald, M., Reese, H., Sanou, J., Tankoano, B., & Mattsson, E. (2015). Mapping Tree Canopy Cover and Aboveground Biomass in Sudano-Sahelian Woodlands Using Landsat 8 and Random Forest. *Remote Sensing, 7*, 10017-10041

Ke, Y., Quackenbush, L.J., & Im, J. (2010). Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. *Remote Sensing of Environment, 114*, 1141-1154

Keith, K., & Tristan, G. (2015). NEON L0-TO-L1 Discrete-return LiDAR Algorithm Theoretical Basis Document (ATBD). In: The National Ecological Observatory Network

Keller, M. (2007). Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms. *Journal of Applied Remote Sensing, 1*, 013537

Khosravipour, A., Skidmore, A.K., Isenburg, M., Wang, T., & Hussin, Y.A. (2014). Generating Pit-free Canopy Height Models from Airborne Lidar. *Photogrammetric Engineering & Remote Sensing, 80*, 863-872

Kim, S., McGaughey, R.J., Andersen, H.-E., & Schreuder, G. (2009). Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. *Remote Sensing of Environment, 113*, 1575-1586

Kindermann, G.E., McCallum, I., Fritz, S., & Obersteiner, M. (2008). A global forest growing stock, biomass and carbon map based on FAO statistics. *Silva Fennica, 42*, 387

Koch, B., Heyder, U., & Weinacker, H. (2006). Detection of individual tree crowns in airborne lidar data. *Photogrammetric Engineering & Remote Sensing, 72*, 357-363

Kraus, K., & Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing, 53*, 193-203

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press

Kwok, R., Cunningham, G.F., Hoffmann, J., & Markus, T. (2016). Testing the ice-water discrimination and freeboard retrieval algorithms for the ICESat-2 mission. *Remote Sensing of Environment, 183*, 13-25

Leckie, D., Gougeon, F., Hill, D., Quinn, R., Armstrong, L., & Shreenan, R. (2003). Combined high-density lidar and multispectral imagery for individual tree crown analysis. *Canadian Journal of Remote Sensing, 29*, 633-649

Lee, A.C., & Lucas, R.M. (2007). A LiDAR-derived canopy density model for tree stem and crown mapping in Australian forests. *Remote Sensing of Environment, 111*, 493-518

Lefsky, M.A. (2010). A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophysical Research Letters, 37*, n/a-n/a

Lefsky, M.A., Cohen, W.B., Parker, G.G., & Harding, D.J. (2002). Lidar Remote Sensing for Ecosystem Studies Lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular interest to forest, landscape, and global ecologists. *BioScience, 52*, 19-30

Lefsky, M.A., Harding, D.J., Keller, M., Cohen, W.B., Carabajal, C.C., Del Bom Espirito-Santo, F., Hunter, M.O., & de Oliveira, R. (2005). Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters, 32*

Li, W., Guo, Q., Jakubowski, M.K., & Kelly, M. (2012). A new method for segmenting individual trees from the lidar point cloud. *Photogrammetric Engineering & Remote Sensing, 78*, 75-84

Lucy, L.B. (1974). An iterative technique for the rectification of observed distributions. *The astronomical journal, 79*, 745

Magruder, L.A., Wharton, M.E., Stout, K.D., & Neuenschwander, A.L. (2012). Noise filtering techniques for photon-counting ladar data. In, *SPIE Defense, Security, and Sensing* (pp. 83790Q-83790Q-83799): International Society for Optics and Photonics

Mallet, C., & Bretar, F. (2009). Full-waveform topographic lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing, 64*, 1-16

Mallet, C., Lafarge, F., Bretar, F., Roux, M., Soergel, U., & Heipke, C. (2009). A stochastic approach for modelling airborne lidar waveforms. *Laserscanning*, 201-206

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., Jasinski, M., Kwok, R., Magruder, L., Lubin, D., Luthcke, S., Morison, J., Nelson, R., Neuenschwander, A., Palm, S., Popescu, S., Shum, C.K., Schutz, B.E., Smith, B., Yang, Y., & Zwally, J. (2017). The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation. *Remote Sensing of Environment, 190*, 260-273

Marvin, D.C., Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E., Sinca, F., & Tupayachi, R. (2014). Amazonian landscapes and the bias in field studies of forest structure and biomass. *Proceedings of the National Academy of Sciences, 111*, E5224-E5232

McGill, M., Markus, T., Scott, V.S., & Neumann, T. (2013). The Multiple Altimeter Beam Experimental Lidar (MABEL): An Airborne Simulator for theICESat-2Mission. *Journal of Atmospheric and Oceanic Technology, 30*, 345-352

McGlinchy, J., van Aardt, J.A.N., Erasmus, B., Asner, G.P., Mathieu, R., Wessels, K., Knapp, D., Kennedy-Bowdoin, T., Rhody, H., Kerekes, J.P., Ientilucci, E.J., Wu, J., Sarrazin, D., & Cawse-Nicholson, K. (2014). Extracting Structural Vegetation Components From Small-Footprint Waveform Lidar for Biomass Estimation in Savanna Ecosystems. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7*, 480-490

McRoberts, R.E., Cohen, W.B., Næsset, E., Stehman, S.V., & Tomppo, E.O. (2010). Using remotely sensed data to construct and assess forest attribute maps and related spatial products. *Scandinavian Journal of Forest Research, 25*, 340-367

Mémoli, F., & Sapiro, G. (2004). Comparing point clouds. In, *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing* (pp. 32-40): ACM

Morhac, M. (2012). Peaks: Peaks. R package version 0.2. In. https://CRAN.R-project.org/package=Peaks

Morhac, M., Kliman, J., Matousek, V., Veselsky, M., & Turzo, I. (1997). Efficient one- and two-dimensional gold deconvolution and its application to gamma-ray spectra decomposition. *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment, 401*, 385-408

Morháč, M., Matoušek, V., & Kliman, J. (2003). Efficient algorithm of multidimensional deconvolution and its application to nuclear data processing. *Digital Signal Processing, 13*, 144-171

Morsdorf, F., Meier, E., Allgöwer, B., & Nüesch, D. (2003). Clustering in airborne laser scanning raw data for segmentation of single trees. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 34*, W13

Moussavi, M.S., Abdalati, W., Scambos, T., & Neuenschwander, A. (2014). Applicability of an automatic surface detection approach to micro-pulse photon-counting lidar altimetry data: implications for canopy height retrieval from future ICESat-2 data. *International Journal of Remote Sensing, 35*, 5263-5279

Neal, R.M. (2003). Slice sampling. *Annals of statistics*, 705-741

Neal, R.M. (2011). MCMC using Hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo, 2*, 113-162

Neigh, C.S.R., Nelson, R.F., Ranson, K.J., Margolis, H.A., Montesano, P.M., Sun, G., Kharuk, V., Næsset, E., Wulder, M.A., & Andersen, H.-E. (2013). Taking stock of circumboreal forest carbon with ground measurements, airborne and spaceborne LiDAR. *Remote Sensing of Environment, 137*, 274-287

Nesterov, Y. (2009). Primal-dual subgradient methods for convex problems. *Mathematical programming, 120*, 221-259

Neuenschwander, A., & Magruder, L. (2016). The Potential Impact of Vertical Sampling Uncertainty on ICESat-2/ATLAS Terrain and Canopy Height Retrievals for Multiple Ecosystems. *Remote Sensing, 8*, 1039

Neuenschwander, A.L. (2008). Evaluation of waveform deconvolution and decomposition retrieval algorithms for ICESat/GLAS data. *Canadian Journal of Remote Sensing, 34*, S240-S246

Nordin, L. (2006). Analysis of waveform data from airborne laser scanner systems. In: MS thesis, Lulea Univ. Technol., Lulea, Sweden

Oh, S.-H., & Kwon, B.-D. (2001). Geostatistical approach to Bayesian inversion of geophysical data: Markov chain Monte Carlo method. *Earth, planets and space, 53*, 777-791

Patenaude, G., Milne, R., Van Oijen, M., Rowland, C.S., & Hill, R.A. (2008). Integrating remote sensing datasets into ecological modelling: a Bayesian approach. *International Journal of Remote Sensing, 29*, 1295-1315

Persson, Å., Söderman, U., Töpel, J., & Ahlberg, S. (2005). Visualization and analysis of full-waveform airborne laser scanner data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36*, W19

Pirotti, F. (2011). Analysis of full-waveform LiDAR data for forestry applications: a review of investigations and methods. *iForest - Biogeosciences and Forestry, 4*, 100-106

Plowright, A. (2017). ForestTools: Analyzing Remotely Sensed Forest Data. R package version 0.1.5.

Popescu, S.C. (2007). Estimating biomass of individual pine trees using airborne lidar. *Biomass and Bioenergy, 31*, 646-655

Popescu, S.C., & Wynne, R.H. (2004). Seeing the trees in the forest. *Photogrammetric Engineering & Remote Sensing, 70*, 589-604

Popescu, S.C., Wynne, R.H., & Nelson, R.F. (2002). Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. *Computers and electronics in agriculture, 37*, 71-95

Popescu, S.C., Wynne, R.H., & Nelson, R.F. (2003). Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing, 29*, 564-577

Qin, H., Xie, X., Vrugt, J.A., Zeng, K., & Hong, G. (2016). Underground structure defect detection and reconstruction using crosshole GPR and Bayesian waveform inversion. *Automation in Construction, 68*, 156-169

Ray, A., Alumbaugh, D.L., Hoversten, G.M., & Key, K. (2013). Robust and accelerated Bayesian inversion of marine controlled-source electromagnetic data using parallel tempering. *Geophysics, 78*, E271-E280

Reitberger, J., Krzystek, P., & Stilla, U. (2008). Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. *International Journal of Remote Sensing, 29*, 1407-1431

Reitberger, J., Schnörr, C., Krzystek, P., & Stilla, U. (2009). 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing, 64*, 561-574

Roberts, G.O., Gelman, A., & Gilks, W.R. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms. *The annals of applied probability, 7*, 110-120

Roncat, A., Bergauer, G., & Pfeifer, N. (2010). Retrieval of the backscatter cross-section in full-waveform LIDAR data using B-splines. *Proc. Int. Archives Photogramm. Remote Sens. Spatial Inf. Sci*, 137-142

Roncat, A., Bergauer, G., & Pfeifer, N. (2011). B-spline deconvolution for differential target cross-section determination in full-waveform laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing, 66*, 418-428

Roonizi, E.K., & Sassi, R. (2016). A signal decomposition model-based bayesian framework for ECG components separation. *IEEE Transactions on Signal Processing, 64*, 665-674

Rosette, J., Field, C., Nelson, R., DeCola, P., & Cook, B. (2011). A new photon-counting lidar system for vegetation analysis. *Proceedings of SilviLaser*, 16-19

Rowe, J. (2013). Ground Classification and Below Ground Response Assessment of Forested Regions using Full-Waveform LiDAR

Schlerf, M., Atzberger, C., & Hill, J. (2005). Remote sensing of forest biophysical variables using HyMap imaging spectrometer data. *Remote Sensing of Environment, 95*, 177-194

Sen, M.K., & Stoffa, P.L. (1996). Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion. *Geophysical Prospecting, 44*, 313-350

Simard, M., Pinto, N., Fisher, J.B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research, 116*

Strobl, C., Boulesteix, A.L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics, 8*, 25

Strobl, C., Hothorn, T., & Zeileis, A. (2009a). Party on! A new, conditional variable importance measure for random forests available in the party Package. University of Munich Department of Statistics. In: Technical Report

Strobl, C., Malley, J., & Tutz, G. (2009b). An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods, 14*, 323-348

Swatantran, A., Tang, H., Barrett, T., DeCola, P., & Dubayah, R. (2016). Rapid, High-Resolution Forest Structure and Terrain Mapping over Large Areas using Single Photon Lidar. *Sci Rep, 6*, 28277

Tang, H., Swatantran, A., Barrett, T., DeCola, P., & Dubayah, R. (2016). Voxel-Based Spatial Filtering Method for Canopy Height Retrieval from Airborne Single-Photon Lidar. *Remote Sensing, 8*, 771

Team, R.C. (2013). R: A language and environment for statistical computing

Team., R.C. (2013). R: A Language and Environment for Statistical Computing. In. Vienna, Austria: R Foundation for Statistical Computing

Tison, C., Nicolas, J.-M., Tupin, F., & Maître, H. (2004). A new statistical model for Markovian classification of urban areas in high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing, 42*, 2046-2057

Treitz, P., & Howarth, P. (2000). High spatial resolution remote sensing data for forest ecosystem classification: an examination of spatial scale. *Remote Sensing of Environment, 72*, 268-289

Ulrych, T.J., Sacchi, M.D., & Woodbury, A. (2001). A Bayes tour of inversion: A tutorial. *Geophysics, 66*, 55-69

Vaughn, N.R., Moskal, L.M., & Turnblom, E.C. (2012). Tree Species Detection Accuracies Using Discrete Point Lidar and Airborne Waveform Lidar. *Remote Sensing, 4*, 377-403

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., & Hauglin, K.M. (2011). Comparative testing of single-tree detection algorithms under different types of forest. *Forestry*, cpr051

Vehtari, A., Gelman, A., & Gabry, J. (2016). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *arXiv preprint arXiv:1507.04544*

Vincent, G., Sabatier, D., Blanc, L., Chave, J., Weissenbacher, E., Pélissier, R., Fonty, E., Molino, J.F., & Couteron, P. (2012). Accuracy of small footprint airborne LiDAR in its predictions of tropical moist forest stand structure. *Remote Sensing of Environment, 125*, 23-33

Wagner, W., Ullrich, A., Ducic, V., Melzer, T., & Studnicka, N. (2006). Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. *ISPRS Journal of Photogrammetry and Remote Sensing, 60*, 100-112

Wang, G., & Weng, Q. (2013). *Remote sensing of natural resources*. CRC Press

Wang, H., & Glennie, C. (2015). Fusion of waveform LiDAR data and hyperspectral imagery for land cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing, 108*, 1-11

Wang, X., Pan, Z., & Glennie, C. (2016). A novel noise filtering model for photon-counting laser altimeter data. *IEEE Geoscience and Remote Sensing Letters, 13*, 947-951

Wang, Y., Weinacker, H., Koch, B., & Sterenczak, K. (2008). Lidar point cloud based fully automatic 3D single tree modelling in forest and evaluations of the procedure. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci, 37*, 45-51

Wu, J., van Aardt, J., & Asner, G.P. (2011). A comparison of signal deconvolution algorithms based on small-footprint LiDAR waveform simulation. *IEEE Transactions on Geoscience and Remote Sensing,, 49*, 2402-2414

Wulder, M.A., Bater, C.W., Coops, N.C., Hilker, T., & White, J.C. (2008). The role of LiDAR in sustainable forest management. *The Forestry Chronicle, 84*, 807-826

Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W., & Gobakken, T. (2012). Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment, 121*, 196-209

Yao, W., Krzystek, P., & Heurich, M. (2012). Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sensing of Environment, 123*, 368-380

Yin, T., Lauret, N., & Gastellu-Etchegorry, J.-P. (2016). Simulation of satellite, airborne and terrestrial LiDAR with DART (II): ALS and TLS multi-pulse acquisitions, photon counting, and solar noise. *Remote Sensing of Environment, 184*, 454-468

Yu, X., Litkey, P., Hyyppä, J., Holopainen, M., & Vastaranta, M. (2014). Assessment of Low Density Full-Waveform Airborne Laser Scanning for Individual Tree Detection and Tree Species Classification. *Forests, 5*, 1011-1031

Zhang, C., & Qiu, F. (2012). Mapping individual tree species in an urban forest using airborne lidar data and hyperspectral imagery. *Photogrammetric Engineering & Remote Sensing, 78*, 1079-1087

Zhang, C., Zhou, Y., & Qiu, F. (2015). Individual Tree Segmentation from LiDAR Point Clouds for Urban Forest Inventory. *Remote Sensing, 7*, 7892-7913

Zhang, J., & Kerekes, J. (2015). An adaptive density-based model for extracting surface returns from photon-counting laser altimeter data. *IEEE Geoscience and Remote Sensing Letters, 12*, 726-730

Zhang, J., & Kerekes, J.P. (2014). First-principle simulation of spaceborne micropulse photon-counting lidar performance on complex surfaces. *IEEE Transactions on Geoscience and Remote Sensing, 52*, 6488-6496

Zhao, K., & Popescu, S. (2009). Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA. *Remote Sensing of Environment, 113*, 1628-1645

Zhao, K., Popescu, S., Meng, X., Pang, Y., & Agca, M. (2011). Characterizing forest canopy structure with lidar composite metrics and machine learning. *Remote Sensing of Environment, 115*, 1978-1996

Zhao, K., Popescu, S., & Nelson, R. (2009). Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Remote Sensing of Environment, 113*, 182-196

Zhou, T., & Popescu, S.C. (2017). Bayesian decomposition of full waveform LiDAR data with uncertainty analysis. *Remote Sensing of Environment, 200*, 43-62

Zhou, T., Popescu, S.C., Krause, K., Sheridan, R.D., & Putman, E. (2017). Gold – A novel deconvolution algorithm with optimization for waveform LiDAR processing. *ISPRS Journal of Photogrammetry and Remote Sensing, 129*, 131-150

Zhu, J., Zhang, Z., Hu, X., & Li, Z. (2011). Analysis and application of LiDAR waveform data using a progressive waveform decomposition method. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 38*, 5/W12-31-36

Zhuang, W., & Mountrakis, G. (2014). An accurate and computationally efficient algorithm for ground peak identification in large footprint waveform LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing, 95*, 81-92

Zolkos, S.G., Goetz, S.J., & Dubayah, R. (2013). A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sensing of Environment, 128*, 289-298