

A supervoxel-based vegetation classification via decomposition and modelling of full-waveform airborne laser scanning data

Dong Chen, Jiju Peethambaran & Zhenxin Zhang

To cite this article: Dong Chen, Jiju Peethambaran & Zhenxin Zhang (2018) A supervoxel-based vegetation classification via decomposition and modelling of full-waveform airborne laser scanning data, International Journal of Remote Sensing, 39:9, 2937-2968, DOI: [10.1080/01431161.2018.1437293](https://doi.org/10.1080/01431161.2018.1437293)

To link to this article: <https://doi.org/10.1080/01431161.2018.1437293>



Published online: 07 Feb 2018.



Submit your article to this journal 



Article views: 315



View related articles 



View Crossmark data 



Citing articles: 4 View citing articles 



A supervoxel-based vegetation classification via decomposition and modelling of full-waveform airborne laser scanning data

Dong Chen *, Jiju Peethambaran^b and Zhenxin Zhang^{c,d}

^aCollege of Civil Engineering, Nanjing Forestry University, Nanjing, China; ^bSchulich School of Engineering, University of Calgary, Alberta, Canada; ^cBeijing Advanced Innovation Centre for Imaging Technology, Capital Normal University, Beijing, China; ^dBeijing Key Laboratory of Urban Spatial Information Engineering, Beijing Institute of Surveying and Mapping, Beijing, China

ABSTRACT

Vegetation classification is a fundamental task in several applications such as forest management, remote sensing-based crop monitoring, and mitigation of plant diseases, digital prototyping of plants, and plant phenotyping, among others. We propose a novel supervoxel-based methodology to accurately detect vegetation from small-footprint full-waveform airborne laser scanning data in urban and mountainous scenes. Mathematically, the full-waveform decomposition and fitting model based on multiple kernels is presented to generate high-density 3D point clouds and their relevant attributes. The homogeneous supervoxels are then generated by using an enhanced probability density clustering (PDC) algorithm. For each supervoxel, we employ latent Dirichlet allocation to obtain supervoxels features through generalisation of geometric and full-waveform features of point clouds. The Support Vector Machine (SVM) and ensemble classifier random forest (RF) are used to classify these supervoxels into vegetation and non-vegetation. Our experiments on urban and mountainous scenes demonstrate that our approach achieves an overall accuracy of 98.27% and 96.47% respectively by RF classifier and achieves an overall accuracy of 98.16% and 97.67% on the same data sets by SVM classifier. By integrating full-waveform information and more meaningful generalised features, our method outperforms state-of-the-art methods at preserving a trade-off between missing alarm rate and false alarm rate.

ARTICLE HISTORY

Received 5 May 2017
Accepted 30 January 2018

KEYWORDS

Vegetation classification;
airborne laser scanning;
full-waveform data;
supervoxels; latent
Dirichlet allocation

1. Introduction

Vegetation classification is an integral part of many high-level applications arising in forest management and planning. Several works focus on accurate detection of vegetation and estimation of forest inventory, such as tree height (Alexander et al. 2014; Puttonen et al. 2016), biomass (Ferraz et al. 2016), tree crown and stem delineation (Hu et al. 2014; Xia et al.

CONTACT Dong Chen  chendong@njfu.edu.cn  College of Civil Engineering, Nanjing Forestry University, No.159 Longpan Road, Nanjing, Jiangsu, China, 210037

*Co-first authors: Dr Dong Chen and Dr Jiju Peethambaran share the first authorship

© 2018 Informa UK Limited, trading as Taylor & Francis Group

2015), basal area (Bright et al. 2013), and tree species (Guan et al. 2015; Ko et al. 2016). Traditionally, the vegetation information is acquired through expensive, labour-intensive, and time-consuming field inventories. Recent developments in remote sensing technology, particularly full-waveform LiDAR technique, take forest inventory to new heights, as it has capabilities to penetrate dense vegetation canopies, reflecting the multilayers in complex forests. Additionally, the full-waveform light detection and ranging (LiDAR) data include abundant geometric, radiometric, and full-waveform features, which play a fundamental role in accurate object recognition and scene understanding. In this paper, we aim to automatically detect the vegetation cover from full-waveform LiDAR data. Throughout this paper, the term ‘vegetation’ refers to only trees and shrubs. That is vegetation does not include grasslands.

Over the last few years, LiDAR has been widely used for object recognition, 3D building model reconstruction, 3D vegetation mapping, and forest inventory. To date, LiDAR sensors have undergone a huge leap from the traditional discrete-return recording system to full-waveform LiDAR system. Most of the commercially available discrete-return sensors allow for a few, typically four or multiple returns to be recorded for each emitted pulse. However, the full-waveform LiDAR sensor has the ability to record the full profile of backscattered signal at a fixed sampling interval. Although the processing of full-waveform data needs a dramatically higher level of complexity, it can obtain very high density point clouds with high vertical resolution, which are most suitable for describing the vertical structure of vegetation. Moreover, after full-waveform processing, abundance of full-waveform attributes, e.g. the calibrated amplitude echo width, backscatter cross-section, backscatter cross-section per illuminated area, and backscatter coefficient, can be derived. These additional attributes play an irreplaceable role in enhancing the segmentation and recognition of vegetation.

Contributions: In this paper, we thoroughly explore the potentials of full-waveform airborne laser scanning (ALS) data, by presenting a novel supervoxel-based methodology for vegetation detection from small-footprint, full-waveform ALS data. Based on the concepts of OBPA (object-based point cloud analysis), our algorithm can obtain more homogeneous results, on the other hand, the full-waveform data include more abundant information by decomposition and modelling, which can contribute to obtain the more discriminative features, improving the overall accuracy of vegetation classification. Given that our methodology is built on some of previous works, we specifically state the contributions of this work below:

- **Decomposition and modelling of full-waveform profile framework:** We propose an optimisation framework that directly couples the local fitting data and the global priors over multiple kernels to decompose and model ALS full-waveform profiles. The local fitting of the model is determined by how well the inferred model agrees to the complex backscattered profiles, while the global shape of the inferred model is progressively learned and enforced via a constrained non-convex optimisation over multiple kernels. The proposed framework is open and flexible, which means other potential kernels from the extensible kernel library can be easily integrated into current framework without altering the structure of energy formulation.

- **Construction of supervoxels via enhanced probability density clustering (PDC) algorithm:** We present an enhancement to the PDC algorithm (Poullis 2013) for generating supervoxels. The algorithm only requires one input parameter whose value is insensitive to the data sets. The entire process of supervoxel generation is simple but efficient, achieving a good balance between the complexity of supervoxel generation and the accuracy of vegetation classification.
- **Supervoxel feature generalisation via latent Dirichlet allocation (LDA) model:** We introduce the LDA model to generalise the point-based features into the high-level features over supervoxels, thereby improving the discriminative power of the feature descriptors.

2. Related works

In remote sensing, there exist a large body of literature on vegetation extraction from different data sources. Data sources mainly include optical images (Yu et al. 2006; Lu and Weng 2007; Zhang et al. 2011; Rapinel et al. 2014), synthetic aperture radar data (Ersahin, Cumming, and Ward 2010; Yu, Qin, and Clausi 2012; Neumann, Ferro-Famil, and Reigber 2010; Neumann et al. 2012; Ghulam, Porton, and Freeman 2014; Soja, Persson, and Ulander 2015; De Almeida Furtado, Silva, and De Moraes Novo 2016), and LiDAR data (Chen et al. 2013; Zhang et al. 2016). In this section, we focus only on the works relevant to our topic, i.e. vegetation classification techniques from LiDAR data. Discrete-return LiDAR point clouds and full-waveform LiDAR data have given rise to a vast body of literature on 3D vegetation mapping, object segmentation and recognition, and forest inventory. From a technical point of view, the object segmentation (involving vegetation segmentation) techniques can be generally categorised into (1) unsupervised clustering methods, (2) machine learning-based methods, (3) probability-based methods, (4) optimisation-based methods, (5) ensemble learning methods, and (6) object-based methods. We review them in the subsequent sections.

2.1. Unsupervised methods

Haala and Brenner (1999) combine laser altimeter data with multispectral images into an integrated classification. The unsupervised classifier of iterative self-organising data analysis is used for extracting streets, grass-covered areas, trees, buildings, and shadows in urban areas. Reitberger, Krzystek, and Stilla (2008) propose an approach for the classification of deciduous and coniferous trees using features derived from small-footprint full-waveform data. In their method, the simple unsupervised classifiers, e.g. k-means and expectation maximisation, are employed for tree species classification under leaf-on and leaf-off conditions with accuracy of 85% and 96%, respectively. Generally, the unsupervised methods are efficient; however, they generally need to explicitly specify the number of classified objects in advance.

2.2. Machine learning techniques

Machine learning explores the study and construction of complex models that can learn and predict from the sample inputs. Wagner et al. (2008a) explore the understanding of

the scatter characteristics of vegetation and the underlying terrain by distinguishing vegetation from terrain points. The decision tree classifier is used to process small-footprint full-waveform ALS points and the results show that the echo width and total cross-section are useful indicators for separating vegetation and terrain echoes, reaching an overall accuracy of 89.9% and 93.7% for a dense natural forest and baroque garden areas, respectively. Alexander et al. (2010) utilise a decision tree classifier to segment the urban areas into vegetation, roads and building roofs. They demonstrate that the backscatter coefficient attribute is especially useful for separating road and grass. However, when construing a decision tree classifier, the determination of decision thresholds is significantly influenced by study areas and sample size inputs. Horvat, Žalik, and Mongus (2016) combine the non-linear distributions of vegetation point clouds with some other types of context information to derive discriminative attributes as inputs of decision tree for vegetation detection. Salah and Trinder (2010) introduce a Support Vector Machine (SVM)-based method to classify complex, urban landscapes by fusing ALS point clouds with multispectral images. Mallet et al. (2011) study the potentials of full-waveform data by segmenting urban areas into building, ground, and vegetation using SVM classifier. The results show that echo amplitude, cross section, and backscatter coefficient significantly contribute to the high accuracy of classification; however, limited ground truth in their experiment overestimates the classification accuracy. Guan et al. (2015) utilise deep learning to generate high-level discriminative features for tree species classification from mobile laser scanning (MLS) point clouds.

2.3. Probability and optimisation-based classifiers

Although machine learning performs well in high-dimensional feature space and with small training sample size, it does not consider the spatial correlation among multiple labels. Therefore, probability-based classifiers, e.g. conditional random fields (CRF) or Markov random fields (MRF), and optimisation-based classifiers, e.g. graph cuts or normalised cuts, which incorporate the context information of neighbourhood, are frequently used to obtain a more homogeneous classification results. Zhang and Sohn (2010) present a novel MRF-based stochastic framework for individual trees detection from ALS point clouds. In this flexible framework, other features or constraints can be easily integrated into the current model without having to alter the structure of algorithm. Niemeyer et al. (2011) apply CRF probabilistic model to label multiple objects, namely ground, building, and vegetation from full-waveform ALS data. The local context and full-waveform metrics are incorporated into CRF model, obtaining an overall accuracy of 94.3% and 89.3% for Bonnland and Kiel data sets. Yang and Dong (2013) present a shaped-based segmentation method for urban scenes classification using MLS point clouds. In their method, they use normalised cuts to refine over-segmented results, thereby achieving a global optimal segmentation. Vosselman, Coenen, and Rottensteiner (2017) jointly use the multiple segment techniques to segment the point clouds into different size patches, representing the planar as well as arbitrary shapes. After that, a series of context have been utilised and coupled into CRF model to finalise the segment-based contextual classification of ALS point clouds. In the context of fallen tree extraction, Polewski et al. (2015) use the high mean point density ALS points (30 points/m^2) to detect of fallen tree stems. To achieve this, the digital terrain

model points are eliminated, and the probability of each object point is calculated according to the likelihood of each point belonging to the fallen tree. All the stem segment candidates are generated and selected, followed by a merge operation via normalised cut. This work was further enhanced in Polewski et al. (2017) by optimising the stem candidate selection step by applying a constrained CRF, which not only consider each candidate individually but also incorporate the context pairwise interactions into the selection process, thereby improving the stem detection rate.

2.4. Ensemble learning

Considering that the probability- and optimisation-based methods have high time complexity for training their models, some emerging researches tend to use ensemble learning for vegetation classification. Ensemble learning is also known as decision fusion, which obtain the final decision according to statistical analysis of results of multiple classifiers. The representative algorithms such as AdaBoost, JoinBoost, and random forest (RF) belong to this category. Carlberg et al. (2009) introduce a cascade of binary classifier for identifying water, ground, roof, and tree from ALS points, each of which are based on an unsupervised region growing, followed by supervised segment-wise classification using RF classifier. The result shows that the maximum precision and recall for tree classification can reach 93.7% and 95.4%, respectively. Samadzadegan, Bigdeli, and Ramzi (2010) establish a method for fusion of multi-class SVM classifiers that are trained on different subset of features to segment building, terrain, and vegetation from ALS points. This method is based on the concept of decision fusion, which takes full advantage of each classifier and gets the final decision by the majority voting. Compared to each of the single SVM classifier, accuracy and performance are significantly improved. Guo et al. (2011) propose a framework to evaluate the relevance of features extracted from multispectral images and ALS point clouds using the RF classifier. The buildings, vegetation, artificial ground, and natural ground regions are well separated. Meanwhile, the significance of relevant features is evaluated during the procedure of feature selection. Wang et al. (2015) use the AdaBoost to separate TLS point clouds with varying point density, occlusions, and missing data. Similarly, the AdaBoost classifier is also used in Yao and Wei (2013) for vegetation classification, followed by individual tree segmentation using spectral clustering technique through analysing the graph structure of point clouds. This methodology needs both LiDAR and imagery data sources to define the most discriminative features related to trees which are provided as input to the AdaBoost. Guo et al. (2015) propose a supervised classification in a complex landscape including JointBoost classification, followed by graph cuts refinement. The ALS point clouds are segmented into five classes: vegetation, building, ground, power-line, and pylon. Matikainen et al. (2017) fully exploit the potentials of multispectral ALS data in land cover classification and change detection using segment-based RF classifier. They achieve 96% overall accuracy of land cover classification with six classes, directly demonstrating the multispectral ALS points are useful for land cover classification especially for the elevated and ground-level classes.

2.5. Object-based point cloud analysis

To date, OBPA methods are more appropriate for image and point cloud segmentation than traditional pixel-based and point-based classification techniques as it is resistant to discontinuities and the poor quality of the input point clouds. Moreover, the methods based on OBPA are flexible. That is to say, the users can build their own OBPA frameworks to highlight their unique functions and features. The OBPA has proven to be superior to other traditional classification techniques not only on satellite images but also point clouds (Yan, Shaker, and El-Ashmawy 2015). Secord and Zakhor (2007) propose an OBPA method to eliminate trees for reconstructing 3D polyhedral building roof models by fusing multispectral images with ALS data. The region-growing segmentation is first employed, followed by an object-based classification using the SVM classifier. Rutzinger et al. (2008) present an OBPA method, combining segmentation and classification using full-waveform ALS data for detecting vegetation in urban environments. More specifically, the pre-segmentation is performed with the echo width, followed by a statistical classification of trees, reaching completeness and correctness greater than 90%. Chen and Zakhor (2009) propose an OBPA workflow for tree detection in large-scale North American and European ALS data sets. The method consists of region-growing segmentation, followed by RF classification. For each data set, the precision and recall rates reach over 95%. Höfle, Hollaus, and Hagenauer (2012) integrate the 2D vector map, e.g. alley tree and building cadastre map, and full-waveform ALS data, for urban vegetation mapping. In their method, the polygon segments derived from edge-based segmentation are used as input for a decision tree and an artificial neural network classifier, achieving correctness of 98% for vegetation classification. However, the low point density leads to noises due to the limited number of echoes within the segment. To validate the effects of multiple entities in point cloud classification, Xu, Vosselman, and Elberink (2014) introduce the multiple-entity (point, planer segments, and segments derived by mean shift) strategy for urban scene classification. The experiment shows that the proposed method has obvious advantage for vegetation extraction, compared to the traditional pointwise (PW)- and planar-segment-based methods.

In contrast to the aforementioned work for multiple object classes extraction, in our scenario, we design an algorithm to solely classify the vegetation based on the following principles and intentions:

- For the optimum use of supervoxel-based features in vegetation classification, on the one hand, the clustering algorithm is capable of generating largest possible segments with different supervoxel size rather than the traditional homogeneous supervoxels with uniform size. To do this, the feature descriptors could be more rich, stable, and discriminative, thereby enhancing the vegetation separability. On the other hand, supervoxels clustering algorithm as a preprocessing should be simple, effective, and as fast, thereby highlighting the advantages of computational efficiency of OBPA-based classification algorithm.
- Considering the advantage of full-waveform laser technique capable of describing vegetation vertical structures, this paper aims to fully exploit the potentials of full-waveform features and integrate them into our algorithm as much as possible to

improve vegetation separability. For the selection of other non-waveform features, we should choose the target-oriented features that are most relevant to vegetation. Particularly, we try our best to primarily select the statistical features (e.g. Spin image and SPFH features) due to their stability, robustness, and discriminative power.

- Given that our algorithm aims to classify the vegetation only, we try to achieve a good trade-off between the complexity of algorithm and the accuracy of vegetation classification, thus making the algorithm scalable for processing large-scale ALS point clouds.

3. Methodology

A detailed block diagram of the system model and the main functions used to distinguish the vegetation and non-vegetation from airborne full-waveform LiDAR data is shown in [Figure 1](#). It consists of four steps: (1) the full-waveform data processing, (2) the supervoxel generation, (3) extraction attributes over supervoxels, and (4) supervoxel-based vegetation classification. All the specific operations for implementing the proposed methodology are presented in [Sections 3.1–3.4](#). The pipeline starts by smoothing the full-waveform backscattered profiles via Savitzky–Golay (SG) algorithm and then decomposing and modelling the backscattered profiles by solving the kernel selection and optimisation problems. Through this step, the dense point clouds and their associated full-waveform attributes are accurately generated, and provided as input to the supervoxel construction step, which incrementally classify the scene into homogeneous supervoxels via spatial and intensity proximity criteria. Using the LDA model, the features of each supervoxel are acquired by generalising the point-based features within the corresponding supervoxel. The vegetation supervoxels are finally recognised from non-vegetation supervoxels by SVM and RF classifiers. Each stage of the vegetation extraction pipeline is elaborated in the subsequent sections.

3.1. Fitting of full-waveform backscattered signals

We propose a two-step algorithm to obtain sparse 3D point clouds and associated features from the full-waveform ALS data. More specifically, the backscattered full-waveform profiles are first smoothed by SG algorithm (Savitzky and Golay [1964](#)). Then, we formulate a convex energy function involving local fitting errors and the global priors over multiple kernels to calculate the position of multiple echoes (full-waveform decomposition) and retrieve their relevant features (full-waveform modelling). Our signal fitting model provides a uniform framework by simultaneously solving full-waveform decomposition and modelling.

(1) **SG smoothing.** In our setting, the backscattered signals generally contain a certain degree of noise, which affect the accurate extraction of the potential components (echoes) in the signals. To suppress noise, the entire backscattered signal of each emitted pulse should be smoothed in advance. SG smoothing filter is well-suited for handling the signals sampled at equal time intervals and it can also suppress the noise while preserving the signal details, e.g. maximum value, minimum value, and width of components. Our full-waveform backscattered profiles are digitised at equal time

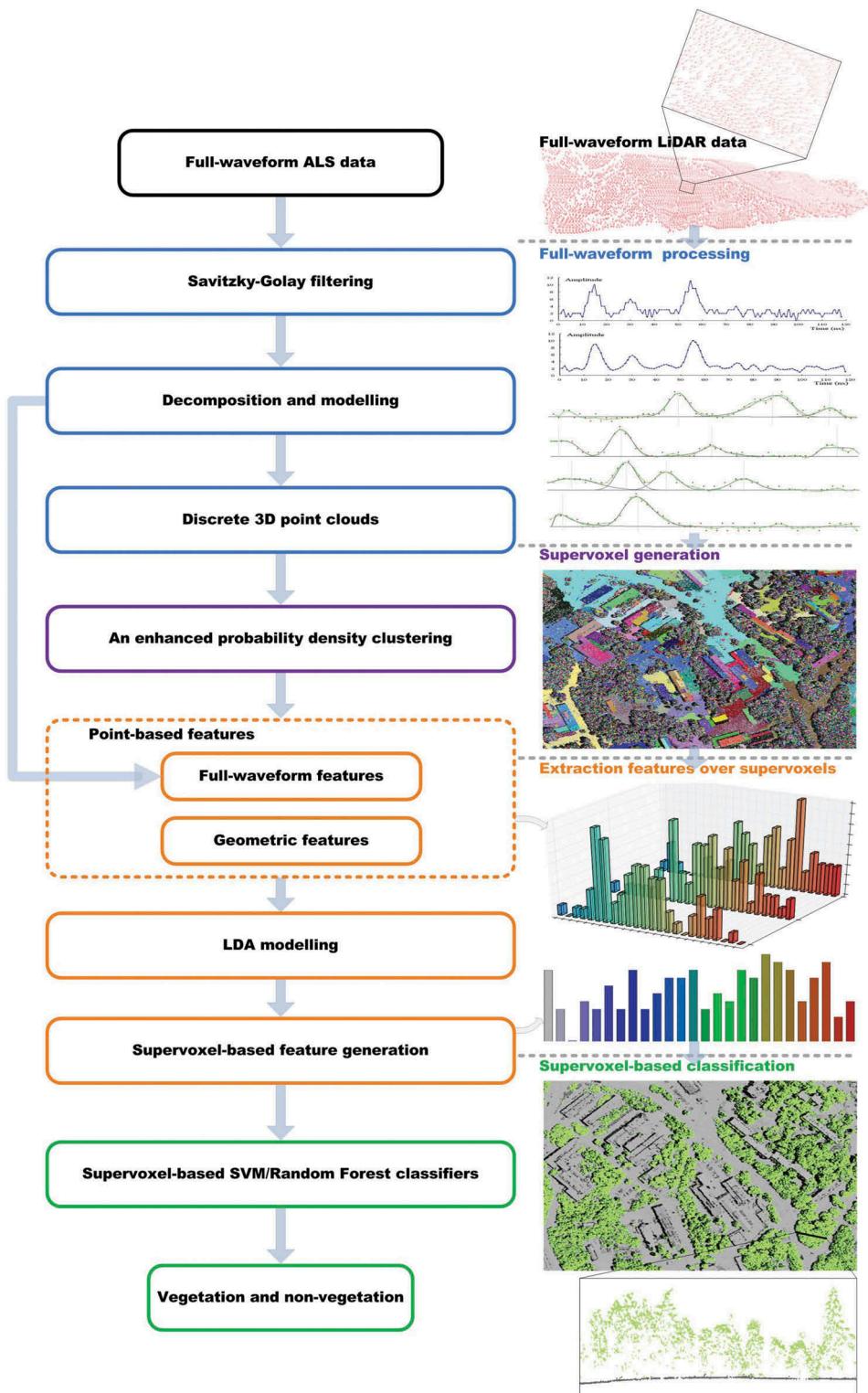


Figure 1. Block diagram showing the proposed methodology. It is to be noted that the different colour rectangles represent the different modules in the pipeline.

intervals and we also want to preserve the details of profile signals; therefore, SG algorithm as an ideal tool is adopted to smooth our full-waveform signals. SG filtering converts the noisy backscattered signals into stable and piecewise-smooth signals, which is then fed as the input to the signal fitting step.

(2) **Full-waveform decomposition and modelling.** As previously mentioned, the original waveform backscattered signal is composed of points $p(x_i, y_i), i = 1, \dots, N$ uniformly spaced at equal time intervals. After SG smoothing, the amplitudes, i.e. intensity of each sampling point are slightly smoothed; however, the topological relationships among sampling points are not changed. Therefore, we formulate the following energy function for decomposition and modelling full-waveform backscattered signals:

$$\theta^* = \operatorname{argmin}_{\theta} \frac{1}{N-t} \sum_{i=1}^N \left(y_i - \sum_{j=1}^n P_j(x_i, \theta) - \epsilon \right), \quad (1)$$

where N denotes the number of samples. The parameter n represents the number of targets (components or echoes) within the travel path of the laser pulse. The parameter t is the number of parameters of energy function. $P_j(\cdot)$ and θ represent the given kernel functions and associated parameters. In practice, we only use four types of kernel functions for decomposing and modelling the backscattered laser signals. More specifically, the Gaussian and general Gaussian kernels are used to fit symmetrically reflected echoes within backscattered profiles. However, the Lognormal and Nakagami kernels are collectively employed to model the reflected echoes with left-skewed or right-skewed behaviours. To minimise Equation (1), the following two critical issues should be considered: (1) how to select the optimal combination of kernel functions to minimise the fitting error and (2) how to determine the initial value θ of the multi-kernels to guarantee Equation (1) converges to a global rather than a local minimum. Since the kernel composition constantly varies, the traditional Gaussian–Newton and Levenberg–Marquardt algorithms cannot solve non-convex energy Equation (1). Inspired by the work in Höfle, Hollaus, and Hagenauer (2012), we introduce the reversible-jump Markov Chain Monte Carlo (MCMC) sampler (Hastings 1970) coupled with simulated annealing to find the global minimum of Equation (1). A MCMC sampler associated with relevant proposition kernels allows us to avoid exhaustive explorations of high-dimensional configuration spaces, no matter whether they are continuous or discrete. Particularly, a MCMC sampler is efficient when the number of kernel functions is unknown.

After solving Equation (1), the parameters of relevant kernels can be obtained, which are regarded as the full-waveform attributes (see Section 3.3). In addition, the accurate time position (T_j for j th echo) of multiple reflected echoes within backscattered profile can be also acquired. The 3D coordinate X_j of j^{th} potential target can be further calculated below:

$$X_j = X_{\text{start}} + cd_{\text{start}}(T_j - T_{\text{start}}), \quad (j = 1, \dots, n), \quad (2)$$

where

X_{start} represents the first sample of start pulse waveform,

d_{start} represents the emitted laser pulse direction,

T_{start} denotes the offset of first sample of surface return waveform in bins, measured from the first sample of the start pulse waveform,

parameter n is the number of detected targets per emitted laser pulse and c is the speed of light.

The detailed diagram can be found in [Figure 2](#).

3.2. Construction of supervoxels

Inspired by the [PDC algorithm](#) in Poullis (2013) for rooftop segmentation from ALS point clouds, in this section we employ an enhanced PDC algorithm to generate supervoxels from the derived sparse point clouds. Compared to the PDC algorithm that relies on high dimensional features, we jointly use a simple height difference (spatial proximity) and intensity difference (intensity proximity) between the current point and the mean value of the cluster as clustering criteria. The reason is that the multiple quantities in high-dimensional space do not have the same scale which

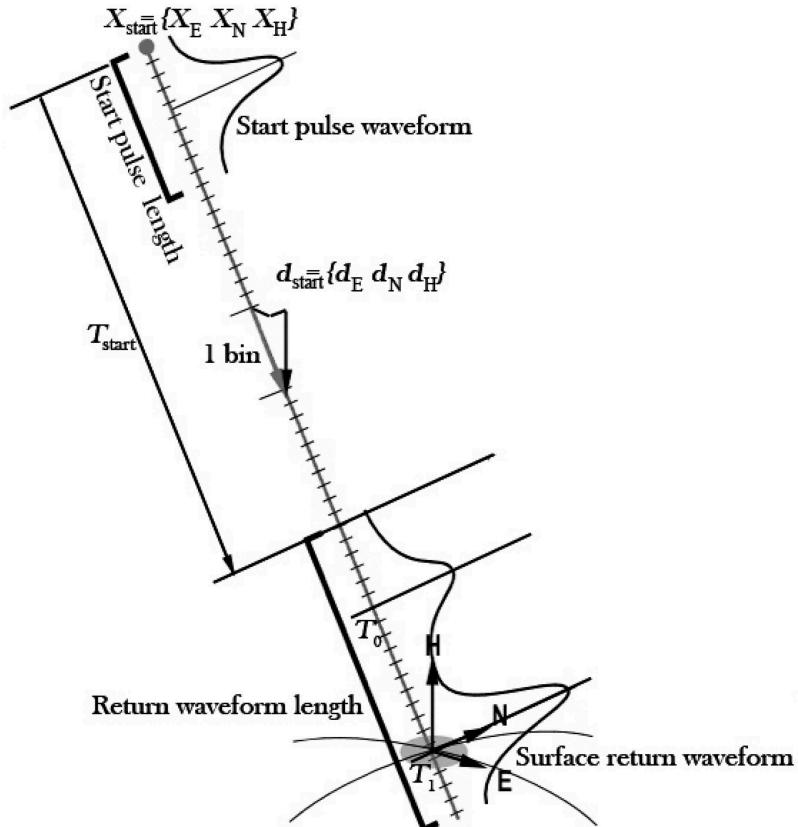


Figure 2. The diagram of discrete 3D point cloud generation. Triplet{E,N,H} represents geocoded positions of surface reflected echoes. The subscript triplet $\{d_E, d_N, d_H\}$ denotes easting, northing, and elevation components of measurement ray in units of waveform samples. Triplet $\{X_E, X_N, X_H\}$ represents the time position of the first sample of start pulse waveform. T_i ($i=0, 1$) represents the time position of the reflected echoes.

makes it biased towards one or the other if simply put together in a high-dimensional distribution. The simplified criteria improve the efficiency and enhance the flexibility of the proposed algorithm. The four steps of the enhanced PDC algorithm are described below:

- (1) **Choose the initial seed point with highest saliency.** As the defect-laden point clouds with missing data, noise, and outliers, random selection of clustering seed point can make the generated supervoxels inhomogeneous. To obtain the reliable and homogeneous supervoxels as much as possible, we first calculate the saliency of each point and select the point with highest saliency as a current seed point for clustering. To this end, we analyse the eigenvalues ($\lambda_i^1 \leq \lambda_i^2 \leq \lambda_i^3$) of the covariance matrix of points within a small sphere of radius \mathcal{R} centred at each point, and calculate each point saliency according to equation below:

$$P_{\text{saliency}} = P_{\text{flatness}} \times P_{\text{uniformity}}, \quad (3)$$

$$\text{s.t. } P_{\text{saliency}} = 1 - 3\lambda_i^1 / \sum_{j=1}^3 \lambda_i^j, \quad P_{\text{uniformity}} = \lambda_i^2 / \lambda_i^3,$$

where P_{flatness} term measures the quality of fit of a local tangent plane estimate at p_i , whose value of 1 indicates a perfect fit, whereas value of 0 denotes the worst possible distribution. The $P_{\text{uniformity}}$ term measures local sampling uniformity at p_i , whose value of 1 means a uniform disc distribution; however, the value of 0 denotes the points lie on a line. Intuitively, the point with the highest saliency generally arises from the interior of the plane primitives.

- (2) **Search neighbourhood of current supervoxel.** Based on k -d tree indexing and breadth-first search technique, the current supervoxel's adjacent neighbourhood points within a sphere of radius \mathcal{R} are searched. Then, we execute step (3) to determine whether or not these adjacent points should be merged into the current supervoxel.
- (3) **Determine the clustering criteria.** Two criteria namely spatial proximity and intensity proximity are used to determine whether or not the candidate point should be merged into the current supervoxel. Two probability density distributions of the candidate point are calculated by

$$P_r(D_p) \propto \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(D_p - \mu)^2}{2\sigma^2}\right\} \quad \&& \quad P_r(I_p) \propto \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(I_p - \mu)^2}{2\sigma^2}\right\}, \quad (4)$$

where D_p is the Euclidean distance from the current candidate point to the fitted primitive, and I_p is the intensity of the candidate point. It should be noted that the point intensity is the corrected intensity as discussed in [Section 3.3](#). The parameters μ and σ are the current supervoxel's mean and variance with respect to D_p and I_p , respectively. According to Equation (4), if these two probability density distributions of the candidate point are both greater than the current supervoxel's probability density defined by $P_r(\mu \pm k\sigma)$ (k is the correlation coefficient, which will be discussed in [Section 4.3](#)), the candidate point will be merged into the current supervoxel.

Meanwhile, the current supervoxel's mean value μ and variance σ with regard to D_p and I_p are updated simultaneously.

(4) **Generate the next supervoxel.** The remaining unvisited point with highest saliency is selected and steps (1) through (4) are repeated until all the unvisited remaining point clouds have been processed.

3.3. Supervoxel feature generation

In this section, a series of full-waveform and geometric features of each point are first extracted. Thereafter, each point is associated to a set of features comprising full-waveform features as well as geometric features derived from discrete 3D point clouds. These point-based features are further used to train an LDA model for generating more discriminative features of supervoxels at high level.

(1) **Full-waveform features.** The full-waveform feature consists of two parts: the simple full-waveform features acquired by minimising Equation (1) and other derivatives of high-level full-waveform features, e.g. echo width, backscatter cross-section, backscatter cross-section per illuminated area, and backscatter coefficient. These two part features collectively constitute the complete full-waveform features of each point. These selected full-waveform features are described as below:

- **The simple full-waveform features:** As previously mentioned in Section 3.1, through solving Equation (1), the optimal full-waveform features, e.g. time position T_j , amplitude I_j , skewness η_j , flatness ζ_j , and the standard deviation s_j of each kernel that best fits the j^{th} relevant reflected echo are directly derived. These features can be regarded as the direct full-waveform features.
- **The corrected amplitude:** Since the attribute I_j is the function of the laser power, the distance from the source to the target, the incidence angle, the target reflectivity, and the absorption by the atmosphere, the I_j shows the characteristics of inhomogeneity among different strips. Thus the attribute I_j needs to be corrected to eliminate the fluctuation effects of the emitted power E , the incidence angle a , and the range R . We use the following corrected procedure proposed by Bretar et al. (2009): $I'_j = (I_j / \cos a)(I_{\text{mean}} / I_{\text{current}})(R^2 / R_s^2)$, where I_{mean} is the average amplitude value of all emitted pulses and I_{current} is the amplitude of the emitted pulse of the current peak. R , R_s and a are the range between the current target and the sensor, a standard range for the whole survey area, and the incident angle between the laser beam and the estimated local plane, respectively (Höfle and Pfeifer 2007).
- **Echo width:** The width of echoes fitted with the general Gaussian kernels can be defined by the full width at half maximum (FWHM) echo w_j , as defined by $w_j = 2\sqrt{2 \ln 2} s_j$. FWHM is a useful indicator for separating vegetation and the terrain echoes (Wagner et al. 2008a).
- **Backscatter cross-section:** Based on the radar equation, the cross section σ_j of j^{th} detected echo is defined as $\sigma_j = C_{\text{cal}} R^4 P_j S_j$, where P_j is the amplitude of the echo from scattering echo j , C_{cal} is the calibration constant, which is calculated

by the process of Lehner and Briese (2010) where a portable reflectometer and spectralon targets are used to estimate calibration constant of *in situ* radiometric reference targets (e.g. asphalt areas close to nadir view for each ALS strips).

- **Backscatter cross-section per illuminated area:** Since the cross section of σ_j can be varied according to different flight height and beam divergence, to make comparisons of the measurements of radar system with different spatial resolution more easily, the cross section normalised by the illuminated area A_j is described as $\sigma'_j = \sigma_j/A_j$ [m^2m^2] (Wagner et al. 2008b; Wagner 2010).
- **Backscatter coefficient:** The corresponding illuminated area A_j can be changing with incidence angle a ; therefore, the σ'_j is also related to the incoming beam. It might be more convenient to normalise the backscatter cross-section to the illuminated areas at zero angle of incidence. The backscatter coefficient γ is σ'_j corrected from the incidence angle a : $\gamma = 4\sigma_j/\pi R^2 \beta^2$. When the reflected target is an area-extend target, i.e. the size is larger than the footprint size of the sensor, the γ can also be substituted with $\gamma = \sigma_j/A_j \cos a$ [m^2m^2].

(2) **Geometric features.** We use eigenvalue-based features, FPFH (fast point feature histograms, FPFH) features and spin image features to distinguish the relationships among point clouds in 3D scene. Note that the FPFH feature is originally developed from the PFH feature which describes the local geometry for 3D point clouds. As robust multiple-dimensional features, FPFH retains most of the discriminative power of the PFH while reducing the computational times significantly for real-time applications. For more details, we refer the reader to Rusu, Blodow, and Beetz (2009). We give the detailed geometric feature descriptions below:

- **Eigenvalue-based features:** Eigenvalues represent the extent of a 3D ellipsoid along its principal axes, which provide additional features and can help discriminate planes, edges, corners, lines and volumes. The features based on eigenvalues describe the local spatial distribution of 3D points. To obtain the eigenvalue-based features, typically, the covariance matrix of the point \mathbf{p} and its neighbourhood point set P with a sphere of radius of \mathcal{R} is defined as

$$\mathbf{C} = \frac{1}{|P|} \sum_{\mathbf{p}_i \in P} (\mathbf{p}_i - \mathbf{q})(\mathbf{p}_i - \mathbf{q})^\top, \quad (5)$$

where \mathbf{q} is the mean value of the set P ; λ_1 , λ_2 , and λ_3 ($\lambda_1 \leq \lambda_2 \leq \lambda_3$) are the eigenvalues of covariance matrix \mathbf{C} . We take the following combined 6D vector:

$$[\sqrt[3]{\prod_{i=1}^3 \lambda_i}, (\lambda_3 - \lambda_1)/\lambda_3, (\lambda_2 - \lambda_1)/\lambda_3, \lambda_1/\lambda_3, -\sum_{i=1}^3 \lambda_i \ln(\lambda_i), (\lambda_3 - \lambda_2)/\lambda_3]$$

obtained by eigenvalues as our eigenvalue-based features. According to Gross and Thoennesen (2006), the 6D vector represents the structure tensor omnivariance, structure tensor anisotropy, structure tensor planarity, structure tensor sphericity, structure tensor eigenentropy, and structure tensor linearity, respectively.

- **FPFH features:** The FPFH attempts to find surface variations by exploring the relationships between a point's k -neighbours and their estimated surface normals. Through representations of the geometrical properties around a specific point in high dimensional space, FPFH provides an informative signature for the feature representation. According to the principles proposed by Rusu, Blodow, and Beetz (2009), FPFH feature of point p is calculated as follows:

- (1) Based on k -d tree indexing, search the neighbourhood point set P of current point p with a sphere of radius \mathcal{R} .
- (2) Construct the local coordinate system (e.g. $U-V-W$ frame) according to current point p and its neighbourhood point set P .
- (3) Calculate a set of triplets $\langle a, \phi, \theta \rangle$ for each pair $\langle p, p_k \rangle$, ($p_k \in P$) according to the relationships depicted in Figure 3.
- (4) Quantise a set of triplets into a histogram by counting the number of occurrences of each components in a specific subinterval. The final histogram is also called simple point feature histograms (SPFH) of point p .
- (5) Redetermine the neighbourhood of current point p , and the neighbour's SPFHs are used to weigh the final histogram by the equation below:

$$F_{\text{FPFH}}(p) = F_{\text{SPFH}}(p) + \frac{1}{k} \sum_{k=1}^{i-1} \frac{1}{\omega_k} F_{\text{SPFH}}(p_k), \quad (6)$$

where the weight ω_k represents a distances between a current point p and a neighbour point p_k within a neighbourhood point set P .

- **Spin image features:** The spin image is a 2D histogram image, which is generated by projecting the 3D point clouds onto 2D image defined in object-oriented coordinate system with a specific support point and the associated normal. As the spin image encodes the 3D coordinates of object with respect to the local reference system, it is a local description of the global shape of the object and is invariant to the rigid transformation (Johnson and Hebert 1999). Combined with the schematic diagram of spin image illustrated in Figure 4, the complete spin-image algorithm is described as below:

- (1) Set the parameters of width W and bin size b of the generated spin image.

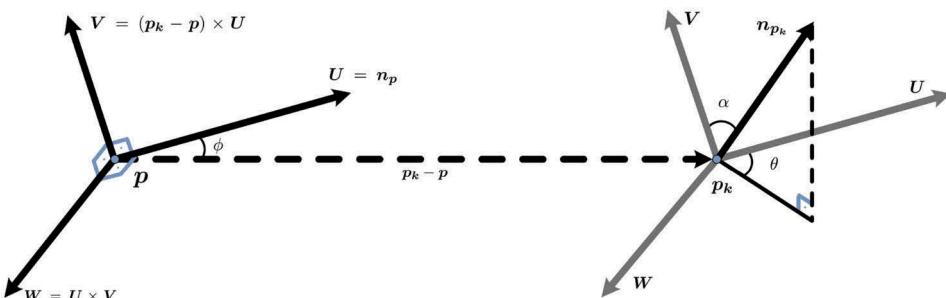


Figure 3. A local $U-V-W$ frame and the relationships for a pair of points p and p_k with their associated normals n_p and n_{p_k} .

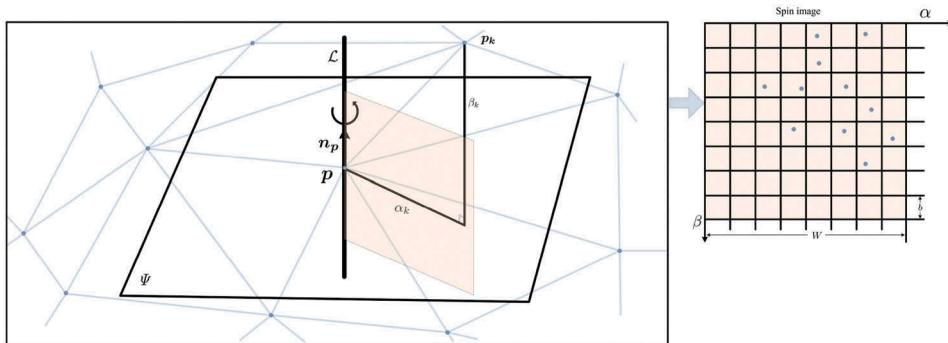


Figure 4. Creation of the 2-D array representation of a spin image.

- (2) Construct the local reference system according to point \mathbf{p} and its normal n_p . For an arbitrary neighbour point \mathbf{p}_k , two cylindrical coordinates can be generated: the radial coordinate a_k , defined as the perpendicular distance to the line \mathcal{L} through the surface normal, and the elevation coordinate β_k , defined as the signed perpendicular distance to the tangent plane Ψ determined by normal and position of point \mathbf{p} . The 2D coordinate (a_k, β_k) is calculated by $(a_k, \beta_k) = (\sqrt{\|\mathbf{p}_k - \mathbf{p}\|^2 - (n_p \cdot (\mathbf{p}_k - \mathbf{p}))^2}, n_p \cdot (\mathbf{p}_k - \mathbf{p}))$.
- (3) Determine the exact projected location (bin) of a pair coordinate (a_k, β_k) in spin image based on the parameters b and W , and increment the accumulator of this bin.
- (4) The above procedure is repeated for all 3D points within the extent of the spin image. The resulting point occurrence accumulation can be regard as the final spin image.

After the abundance of point-based features combined geometric and full-waveform information have been derived, these point-based features are then provided as input to the LDA modelling step which generalises these point-based features into the high-level features over supervoxels, thereby improving the discriminate power of the feature descriptors. More precisely, all the above full-waveform and geometric features are concatenated to form a high-dimensional point descriptor F_{point} . To calculate the supervoxel's features, we cannot simply average F_{point} of the points in supervoxel but instead, should generalise them. Although the number of points varies in different supervoxels, the feature length of each supervoxel should be kept identical, so that these features can be conveniently employed by the subsequent classifiers. To this end, we use LDA model to generalise the features of supervoxels. The LDA is originally proposed to classify the different size of documents by generating a certain number of latent topics. Each document is represented by a vector that denotes the probability of each latent topic exists in the corresponding document. In this paper, each point-based feature F_{point} can be regarded as a word and each supervoxel is taken as a document. The feature of each supervoxel is expected to be generated by the following steps:

- (1) We first linearly scale each attribute of F_{point} to the range $[-1,1]$. This can void the attributes in greater numeric ranges, dominating other attributes in smaller numeric ranges. In addition, it can also eliminate numerical difficulties during the calculation.
- (2) The fuzzy k -means algorithm is used to aggregate all point feature F_{point} into k clusters. After clustering, each point feature F_{point} is replaced with the nearest cluster centre. Actually, the cluster centres are taken as the latent topics of the supervoxels.
- (3) The frequency vector of supervoxel is generated by counting the number of occurrences of the same latent topic within corresponding supervoxel. In this way, each supervoxel is represented by a frequency vector of latent topic in k -dimensional space.
- (4) The LDA model is employed and it works on the frequency vectors to obtain the features of supervoxels consisting of the probability of each latent topic.

The generalized features over supervoxels with same length contain the relations of points within each supervoxel. Meanwhile, the features of supervoxels inherit the orientation invariance of F_{point} and reduce its sensitivity to the varying point density.

3.4. Supervoxel-based classification

Non-parametric machine learning methods, i.e. SVM classifier and ensemble tree RF, are employed to classify the supervoxels using the generalized supervoxels' features. SVM performs a robust non-linear classification of samples by introducing the mechanism of kernel functions, and can effectively avoid over-fitting and under-fitting problems. In this paper, the radial basis function (RBF) kernel is used to achieve a balance between margin maximisation and training error minimisation. RF algorithm is also implemented, from which 50 decisions trees are built. For each tree, the samples are randomly selected from the training set with replacement, so that every decision tree can grow from different set of samples and at each node features are randomly selected to split the node. Each tree is grown to the largest extent possible without any pruning. Through solving the issues of binary classification, the vegetation and non-vegetation supervoxels are obtained after supervoxel-based SVM and RF classification.

4. Experimental evaluation

4.1. Data sets

ALS data sets of Helsinki and Dayekou are used to evaluate the robustness of the proposed approach. The data set from the Helsinki occupies an area of about 3 km^2 and is located in Espoo of southern Finland, a few kilometres west of Helsinki, Finland. The prominent urban scene, which covers 1.08 km^2 with university and office buildings, flat terrain and a small forest, is chosen to evaluate the classification algorithm. The corresponding ALS data contain 4,708,555 point clouds with a point density of 4.3 points/ m^2 . Dayekou data set is acquired from a mountainous area dominated by 15–20 m high spruce and multilevel understoreys. This forest site is located in Qilian

Mountain Dayekou, Zhangye City, Gansu Province, China, with an area of about 14.65 km², containing 26,380,782 point clouds with a mean point density of 1.75 points/m² after full-waveform composition and modelling.

The full-waveform system (Optech instrument system) is used to conduct the Laser campaigns of Helsinki under leaf-on conditions. Data acquisition is performed with a ALTM 3100 system, which uses near-infrared (1064 nm) laser pulses with a pulse width of 4 ns. Further settings include a pulse repetition frequency of 100 kHz, scan angle range of 22.5°, and a beam divergence of 0.8 mrad. After full-waveform data decomposition, the discrete point clouds with a density around 5 points/m² are obtained. Dayekou study area is captured by Y-12 aeroplane, which carries the LiteMapper 5600 full-waveform ALS system at 1 GHz sampling frequency. The average flight height of the aeroplane is approximately 760 m, with a speed of 227 km h⁻¹. The LiteMapper system adopted in this paper uses a Riegl LMS-Q560 laser scanner with near-infrared (1550 nm) laser pulses and a transmitted pulse width of 4 ns. It has a laser beam divergence angle of 0.5 mrad, which produces a footprint diameter of approximately 0.38 m on the target at nadir from a flying height of 760 m. The laser scanning pulse rate is 50 kHz and the maximum scan angle is 22.5° from nadir. After full-waveform data decomposition, the discrete point clouds with a moderate density 1.75 points/m² are obtained. The detailed sensor and data set specifications are summarised in Table 1.

4.2. Full-waveform decomposition and modelling evaluation

For the test data sets, the first step is to decompose and fit the full-waveform back-scattered profiles to obtain sparse 3D point clouds and their corresponding full-waveform features. SG filtering algorithm is employed to filter out the noise while preserving the details of backscattered signals. The smoothed backscattered signals are taken as inputs to the waveform decomposition and modelling. Figure 5 shows six typical smoothed backscattered signals reflected from the bare terrain and forest. Please be noted that despite being smoothed these signals still preserve their prominent characteristics. Different shapes of backscattered profiles reflect the terrain roughness of the surface which the emitted laser pulse interacts with. For example, the signals in Figure 5 (a-d) are most probably reflected from the bare grounds or rooftops as they include only one echo. However, the reflected signals in Figure 5(e, and f) include two or more echoes and hence they are most likely reflected from the objects with complex structures, especially in forest with multilayered canopies. The detected echoes within the backscattered profile represent the individual detected targets in the direction of

Table 1. Sensor configurations and data set specifications.

Data set characteristics	Helsinki data set (Finland)	Dayekou data set (China)
LiDAR sensor	Optech ALTM 3100	IGI LiteMapper 5600
Wavelength	1064 nm	1550 nm
Beam divergence	0.8 mrad	0.8 mrad
Laser repetition rate	100 kHz	50 kHz
Maximum scan angle	22.5°	22.5°
Operating altitude	1100 m	760 m
Testing areas	1.08 km ²	14.65 km ²
Mean point density	4.30 points/m ²	1.75 points/m ²

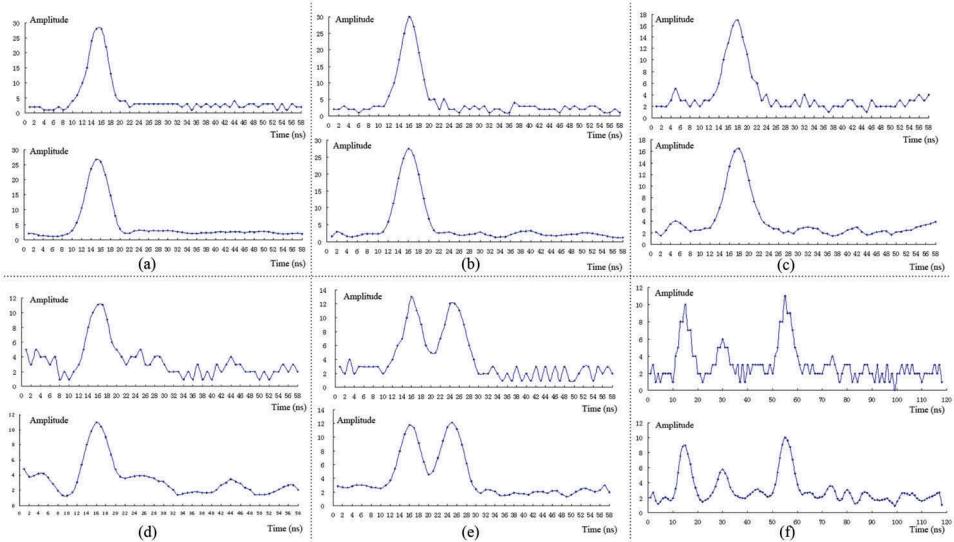


Figure 5. Before-and-after smoothing of some typical backscattered full-waveform pulses. Subfigures (a–d) are the backscattered signals most probably from terrains or rooftops with different roughness. However, the subfigures (e) and (f) indicate the signals that are most likely reflected from the forest with multilayered canopies.

emitted laser pulse. After smoothing, the geometric shape pattern of the profiles becomes more prominent, which provides an ideal reflected profile as input for back-scattered signal decomposition and modelling.

After the full-waveform smoothing, we generate the corresponding dense 3D point clouds and their associated attributes by solving Equation (1). As can be seen from Figure 6, some typical reflected pulses are selected to demonstrate the effectiveness of our kernel selection scheme. The raw digitised signals indicated by red points are accurately smoothed and they are shown in green lines. Thereafter, the smoothed signals are well modelled by the proposed four types of kernels and the results are shown in black grey lines. It is obvious that most of the reflected echoes are symmetrical and well fitted by the Gaussian and general Gaussian kernels. In fact, for airborne full-waveform LiDAR data with small footprints, more than 90% echoes can be successfully modelled by the Gaussian kernel (Wagner et al. 2006). However, we still observe that there exist some asymmetric and weak echoes reflected from the targets with complex 3D structures and/or specific materials. By introducing Lognormal, Nakagami kernels into our model, we can precisely fit these complex asymmetric and weak echoes by free selection of the most appropriate kernels to minimise the fitting error. It is to be noted that in Figure 6, the backscattered profile components labelled with the arrows with different colours represent the corresponding components modelled by different kernels. The readers can refer to the legend to trace the specific kernel used in signal modelling.

To further prove the effects of the proposed multiple kernels, we make a quantitative evaluation of the fitting error by progressively adding the individual kernel into our

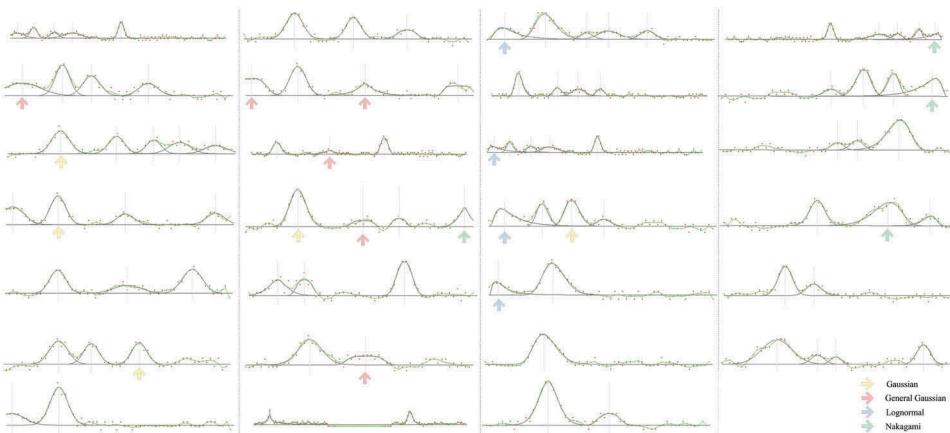


Figure 6. Typical full-waveform ALS pulses decomposition and modelling. The red points denote the raw digitalised backscatter signal points and the green dashed lines represent the smoothed signals after SG smoothing. The components of each backscattered signal are well fitted by black grey lines through selection of the most appropriate kernel composition.

fitting model. The result is shown in Figure 7, from which we can see that the back-scattered profiles cannot be well fit by independently using skewed kernels such as Nakagami and Lognormal. By further adding the symmetrical Gaussian kernel, the residual error decreases significantly, which means most of the airborne full-waveform laser backscattered signals can be better modelled by the symmetrical kernels. The results are further refined by taking into account general Gaussian kernel. In this way, we can easily distinguish which kernels are relatively more important than others, thereby providing a practical solution for the selection of the most appropriate kernels to expand the kernel library. It should be noted that our full-waveform process model is

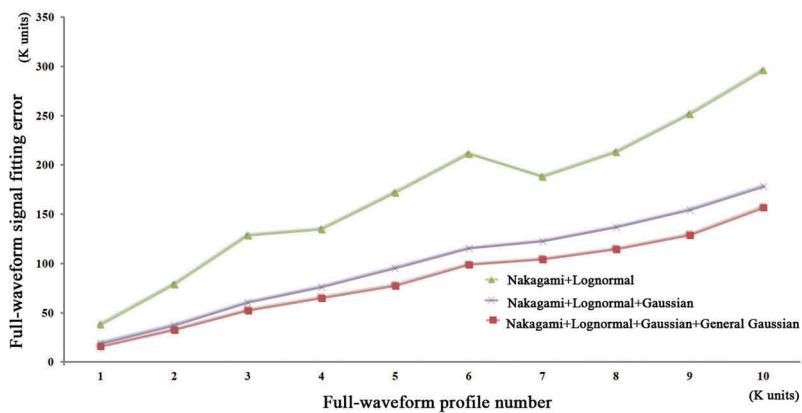


Figure 7. Full-waveform fitting error evaluation by progressively adding kernels. Note that we format the axis values as thousands (K). The quantity of the profile samples is uniform increased from 1K to 10K by random selection of 10 times. For any two arbitrary selections, the samples are not overlapped.

an open-form framework which allows the user to expand their own appropriate kernel library according to the different applications without altering the structure of the energy function. Thus obtained parameters of kernels and the optimal time locations of the backscattered components are taken as the full waveform attributes for aiding the vegetation classification.

4.3. Supervoxel generation evaluation

Next, we evaluate the performance of enhanced PDC algorithm to generate supervoxels from the sparse point clouds. In the proposed PDC method, the point clouds are merged into the current supervoxel only if their probability densities with regard to D_p and I_p simultaneously satisfy the criteria of spatial and intensity proximities. This requires the candidate's probability density to be fallen within $\mu \pm k\sigma$ of the Gaussian distribution of current supervoxel. We use two separate individual clustering criteria rather than we concatenate them together into high dimensional features by defining 2D (D_p, I_p) or 4D (X_p, Y_p, Z_p, I_p) feature vectors as we find that it will yield inconsistencies in clustering behaviours for different supervoxel sizes. More precisely, for large supervoxel sizes, the spatial distance outweighs intensity proximity. As a consequence, the generated supervoxels do not adhere well to intensity boundaries. For smaller supervoxel sizes, the converse is true. Therefore, we can have a proper balance by splitting the clustering criteria into spatial proximity and intensity proximity. Moreover, this organisation overcomes the deficiency of multiple quantities in high dimensional feature spaces with different scales, biasing towards one or the other.

For more in-depth evaluation of our enhanced PDC algorithm, as shown in Figure 8, we further compare our clustering algorithm with other mainstream clustering algorithms, e.g. region growing, mean shift (Cheng 1995) and efficient RANSAC (Schnabel, Wahl, and Klein 2007). Region growing is a classic algorithm which is extensively used to solve the computer vision problems such as image segmentation and classification. The algorithm is simple; however, it needs the seed points to start the program. Different seed point selection strategies can yield different clustering results. In addition, the algorithm is sensitive to clustering thresholds such as Euclidean distance and normal. Inappropriate parameter setting will undoubtedly cause over-segmentation and/or under-segmentation, as shown in Figure 8(a). Although mean shift algorithm does not require the number of clusters, we have to specify the termination condition and a bandwidth value of the adopted kernel function. Mean shift is extremely sensitive to the selection of bandwidth. A small bandwidth can slow down the convergence. In contrast, a large bandwidth can speed up convergence but might merge two models. Moreover, in high dimensional feature spaces, convergence of clustering still remains to be proved. In our scenario, we use point clouds' geometry coordinates as a clustering criterion to generate the supervoxels of nearly uniform size, as shown in Figure 8(b). Efficient RANSAC is another commonly used algorithm for object detection and segmentation. It is computationally intensive, especially when the number of parameters of the assumption model is high. Although this algorithm is capable of extracting the point clouds that are consistent with the predefined model, it almost fails to recognise vegetation and other isolated points, causing the generated supervoxels in these areas

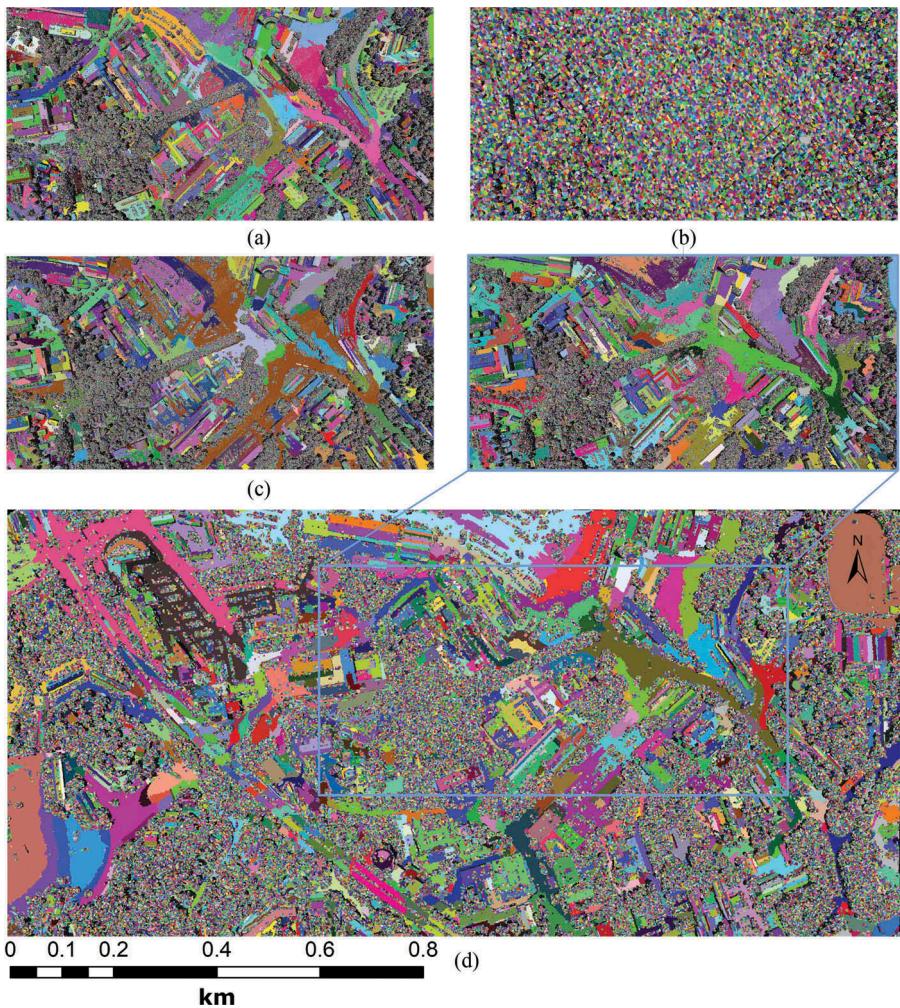


Figure 8. Supervoxel generation algorithm comparison. Subfigures from (a) to (d) are the supervoxels generated by region growing, mean shift, efficient RANSAC, and the enhanced PDC algorithms. Note that (a) and (d) have similar results but supervoxels in (d) are more meaningful, and the entire clustering procedures only need one insensitive parameter. The results in (b) can guarantee the uniform size of the supervoxels but are insufficient to describe the shapes of the supervoxels. Since efficient RANSAC only extracts objects that comply with the predefined models, thus the points from vegetation and isolated objects in (c) seem extremely trivial.

extremely trivial, as evident in Figure 8(c). In contrast to the above algorithms, our enhanced PDC algorithm performs clustering, starting from the seed point with highest saliency, which makes the clustering result identical at each program execution, guaranteeing the repeatability and reproducibility of the algorithm. In addition, our enhanced PDC algorithm does not need to adjust clustering threshold parameters by introducing only one correlation coefficient parameter k . Parameter k controls the strictness of the clustering criterion. Selecting a extremely small or large value will cause over-segmentation or under-segmentation, respectively. Fortunately, k is less

sensitive over a wide range of values. Intuitively, parameter k is encouraged to be set with a relatively small value so as to produce homogeneous supervoxels. In contrast, with extremely large value of k , the generated supervoxels are more heterogeneous. In this case, the high classification error rate is doomed to happen regardless of whatever these supervoxels are labelled as vegetation or non-vegetation. We find that a range of values from 0.5 to 2.5 can both achieve a good balance between the under-segmentation and over-segmentation. The detailed comparisons of those clustering algorithms can be found in [Table 2](#). Parameter d represents Euclidean distance from the point to the cluster. The parameter n represents the normal difference between the point and cluster. b and ϵ denote the bandwidth and the termination condition, respectively. ρ is the mean point density. For time complexity, parameter N represents the number of point clouds, M is the number of hypothesis models evaluated and T determines the number of iterations.

Essentially, the enhanced PDC algorithm is a type of region growing algorithm. Through introducing the probability density kernel and the point saliency, we decrease the number of input parameters, keeping the parameter threshold insensitive and making the clustering results more reproducible. Meanwhile, the generated supervoxels are more meaningful through a natural clustering style in feature spaces. As the size of a homogeneous supervoxel expands, more points are contained in the supervoxel and the calculation of μ and σ becomes more stable and robust. Our clustering result in Figure 8 (d) shows the building rooftop and the ground have strong homogeneity with large-sized supervoxels; however, the point clouds in vegetation areas have low spatial autocorrelation with many trivial fragments.

4.4. Training and classification

In this section, we first introduce the principle to select the training samples, and then give the details of classifier model training. After that, we quantitatively evaluate the accuracy of the supervoxel-based vegetation classification using SVM and RF classifiers. We finally compare our results with the results generated by state-of-the-art methods.

4.4.1. Training sample collection

After the generalisation of supervoxel-based features using LDA model, each supervoxel gets a discriminate feature with the uniform size, which is used for the final supervoxel-based SVM and RF classification. To achieve this goal, a certain number of training samples are first labelled. The registered aerial images and the commercial Terrasolid¹ software

Table 2. A comparison of clustering algorithms.

package are collectively utilised to assist the sample labelling in two- and three-dimensional data space to improve the reliability of training data selection. In practice, we first use the stratified sampling scheme for selection of the most representative point clouds. The stratified sampling guarantees that the selected point clouds are uniformly distributed and derived from different objects. Once the typical points are selected, visual inspection with the registered aerial images and interactive operations in two- and three-dimensional data space using Terrasolid software package are used to complete the reliable labelling of point clouds. When the point cloud samples are accurately labelled, the associated supervoxels can also be labelled naturally based on the principle of a majority vote of point labels within supervoxel. In our scenario, 362,215 and 293,576 point clouds are labelled from Helsinki and Dayekou data sets, respectively. The percentage of Helsinki supervoxels identified as labelled is 10.73%, i.e. 169,952 labelled ones out of 1,582,681 supervoxels. The Dayekou data set includes 10,007,142 supervoxels, from which 156,269 supervoxels are identified as labelled supervoxels, only occupying 1.56% of all of Dayekou supervoxels. These labelled supervoxels are divided into two parts. In the first part, approximately 2.0% of the labelled supervoxels of each data set are randomly selected as training data. The remaining ones are regarded as the reference ground truth, which are used to qualitatively evaluate the accuracy of classifiers.

4.4.2. Classifier training

We jointly use the supervoxel-based SVM and RF classifiers to comprehensively evaluate the accuracy of supervoxels on two data sets. The classifier training as a critical step determines how well the unknown supervoxels are labelled. In our case, for SVM classifier, we use RBF kernel to map the features of supervoxels into the higher dimensional feature spaces to improve supervoxel separability. The RBF kernel have two parameters C and γ , and we use the tenfold cross-validation to obtain the best values. More specifically, we divide 2.0% training set into 10 subsets of equal size. Sequentially, one subset is verified using the classifier trained on the remaining nine subsets. Through this procedure, each training supervoxel is predicted once. The cross-validation rate is the percentage of supervoxel that are correctly classified. We adopt the grid-search strategy (Chang and Lin 2011) on C and γ with tenfold cross-validation and try various pairs of (C, γ) , the value with the best cross-validation rate is determined as an optimal pair to constitute SVM model. In our experiment, an optimal pair is $(2^3, 2^{6.25})$ for Helsinki data set and $(2^{1.25}, 2^5)$ for Dayekou data set.

For a supervoxel-based RF classifier, we use 50 decisions trees to build the model. For each tree, the supervoxel fetures are randomly selected with replacement so that every decision tree can grow from different set of samples. At each node, the features are randomly selected to split the node. Each tree is grown to the largest extent possible without any pruning. In our case, we still use grid-search strategy with tenfold cross-validation to find the optimal parametrization of the RF. More specifically, the number of trees is set to 50 although more trees yield slightly better accuracy but also linearly increase the computational cost. The depth of RF is estimated with cross-validation over a parameter from 20 to 60 with a stride length of 3. Through trial-and-error experiment, we find that the depth of approximately 30 with Gini-index as splitting criterion can yield the best accuracy.

4.4.3. Evaluation metrics

In this section, we quantitatively compare the proposed supervoxel-based RF and SVM algorithms with the traditional PW algorithm so as to reflect the superiority of the supervoxels by using the same classifiers. The quantitative classification confusion matrix results for RF, SVM, and PW on Helsinki and Dayekou data sets are shown in [Table 3](#).

RF, SVM, and PW respectively represent the three different classifiers namely RF, SVM, and PW classifiers.

From [Table 3](#), more rigorous performance evaluation results are further derived as reported in [Table 4](#). Although our supervoxel-based SVM classifier has a relatively higher error rate for the Helsinki data (1.08%) than the RF classifier (1.02%), it performs well for the Dayekou data set with an error rate of 2.06%. This result indicates that SVM classifier is more efficient than RF for the trivial fragments classification. However, for relatively large-sized supervoxels in Helsinki data set, RF is slightly superior to SVM with only 1.02% error rate, which means the RF is more suitable to describe the supervoxels with rich descriptors. Compared to the PW classification, the supervoxel-based algorithms show better results either in Helsinki or Dayekou data set although the advantage is not very obvious probably because the PW algorithms also use our point-based features combining the geometric and full-waveform information. This directly proves that the proposed supervoxel-based features generalized by LDA algorithm have better discriminative ability among different classifiers. In addition, the classification results of Helsinki data in all the three classifiers are better than those of the Dayekou data set due to the existence of relatively large-sized supervoxels in Helsinki data set (Helsinki data set has a relatively lower vegetation coverage ratio). In contrast, to some extent the supervoxel-based classifier for processing Dayekou data set is reduced to a PW classifier due to the presence of too many smaller-sized supervoxels. Additionally, the Dayekou data set contains many understorey shrubs which make them more difficult to distinguish between

Table 3. The confusion matrix for our two data sets using different algorithms.

		Helsinki (Reference)		Dayekou (Reference)	
Classified supervoxels		Vegetation	Non-vegetation	Vegetation	Non-vegetation
RF	Vegetation	209,968 (57.97%)	1072 (0.30%)	246,994 (84.13%)	2606 (0.89%)
	Non-vegetation	2622 (0.72%)	148,553 (41.01%)	6439 (21.93%)	37,537 (12.79%)
SVM	Vegetation	209,828 (57.93%)	1212 (0.33%)	247,162 (84.19%)	2438 (0.83%)
	Non-vegetation	2728 (0.75%)	148,447 (40.98%)	3445 (1.17%)	40,531 (13.81%)
PW	Vegetation	208,170 (57.47%)	2870 (0.78%)	246,249 (83.88%)	3351 (1.14%)
	Non-vegetation	5358 (1.48%)	145,817 (40.26%)	7624 (2.60%)	36,352 (12.38%)

Table 4. Performance evaluation results of the vegetation classification framework.

Performance	Classifier					
	RF		SVM		PW	
Helsinki	Dayekou	Helsinki	Dayekou	Helsinki	Dayekou	
Branching factor	0.012	0.026	0.013	0.014	0.026	0.031
Missing factor	0.005	0.011	0.006	0.009	0.014	0.014
Completeness (%)	99.49	98.95	99.43	99.02	98.60	98.65
Correctness (%)	98.77	97.45	98.72	98.62	97.49	96.99
Quality (%)	98.27	96.47	98.16	97.67	96.19	95.73
Overall error (%)	1.02	3.08	1.08	2.06	2.27	3.73

understorey shrubs and rough-ground terrain, thereby decreasing the vegetation detection rate, no matter which classifier is used.

RF, SVM, and PW respectively represent the three different classifiers namely RF, SVM and PW classifier.

4.4.4. Comparative study

To further demonstrate the superiority of our algorithm, we compare our methodology with two other classification methods (Zhou and Neumann 2008; Horvat, Žalik, and Mongus 2016) in Helsinki data set, while in Dayekou data set we compare our algorithm with three other methods (Zhang et al. 2003; Zhou and Neumann 2008; Horvat, Žalik, and Mongus 2016). The comparison results for Helsinki and Dayekou data sets are shown in Figures 9 and 10, respectively. As can be seen from Figure 9(a), in Zhou's method, the low shrubs and trees distributed in the middle of road are not correctly detected and they mistakenly classified into non-vegetation (see blue rectangle region). The reason of this omission error is that the elevation of these vegetation is very low and the elevation changes of vegetation are similar to those of rough topographic terrains. Meanwhile, a few point clouds distributed on rooftops are incorrectly labelled as vegetation as evident in red rectangle region. This error occurs because of the complicated composition of materials or rooftop superstructures (chimneys, antennas, etc). Figure 9(b) presents vegetation classification results by Horvat's algorithm. Although the low vegetation and small scattered vegetation are successfully extracted, the point clouds from building edges and complicated rooftop superstructures are misclassified into vegetation (see orange and purple rectangles), thereby severely decreasing the vegetation correctness rate. In contrast to the above two methods, our methodology on one hand guarantees most of the low and isolated trees and shrubs are successfully recognised (see the view of cross section in Figure 9(c)), thereby improving the completeness of vegetation classification rate. On the other hand, thanks to the radiometric and full-waveform features, our method has capability to distinguish the point clouds reflected from different materials. That is to say we can successfully differentiate irregular and mass points clouds on rooftop from vegetation, thus increasing the correctness of vegetation classification rate.

As Dayekou data set is captured from a mountainous region, the problem of vegetation classification is transformed into a filtering problem, i.e. binary objects classification between terrain and vegetation. We make comparisons with three other methods (Zhang et al. 2003; Zhou and Neumann 2008; Horvat, Žalik, and Mongus 2016). These comparison results are shown in Figure 10(c-f). It can be observed that severe commission error (classifies vegetation points as terrain points) occurs by Zhou's and Horvat's methods. That is to say, too much vegetation points are incorrectly classified into terrain, which can be clearly seen in Figure 10(d,e) where some tiny bumps distribute in topographic shaded relief map. This misclassification occurs most probably due to multilevel small understoreys under spruce trees. In contrast, the progressive morphological filter proposed by Zhang et al. (2003) can recognise these low-level understoreys in a steep mountainous regions, as the bumps disappear from the shaded relief map, as evident in Figure 10(f). However, omission error (mistakenly classifies terrain points into vegetation points) occurs, which causes the terrain details impaired. It can be observed that in Figure 10(f) the details of terrain texture become blurred and steep characteristics are chopped off to some extent. Our method results in Figure 10(c), on the one hand, keep the terrain at a very fine level of detail, on the other hand there seldom exist bumps in the shaded relief terrain map. The result demonstrates that we can strike a balance

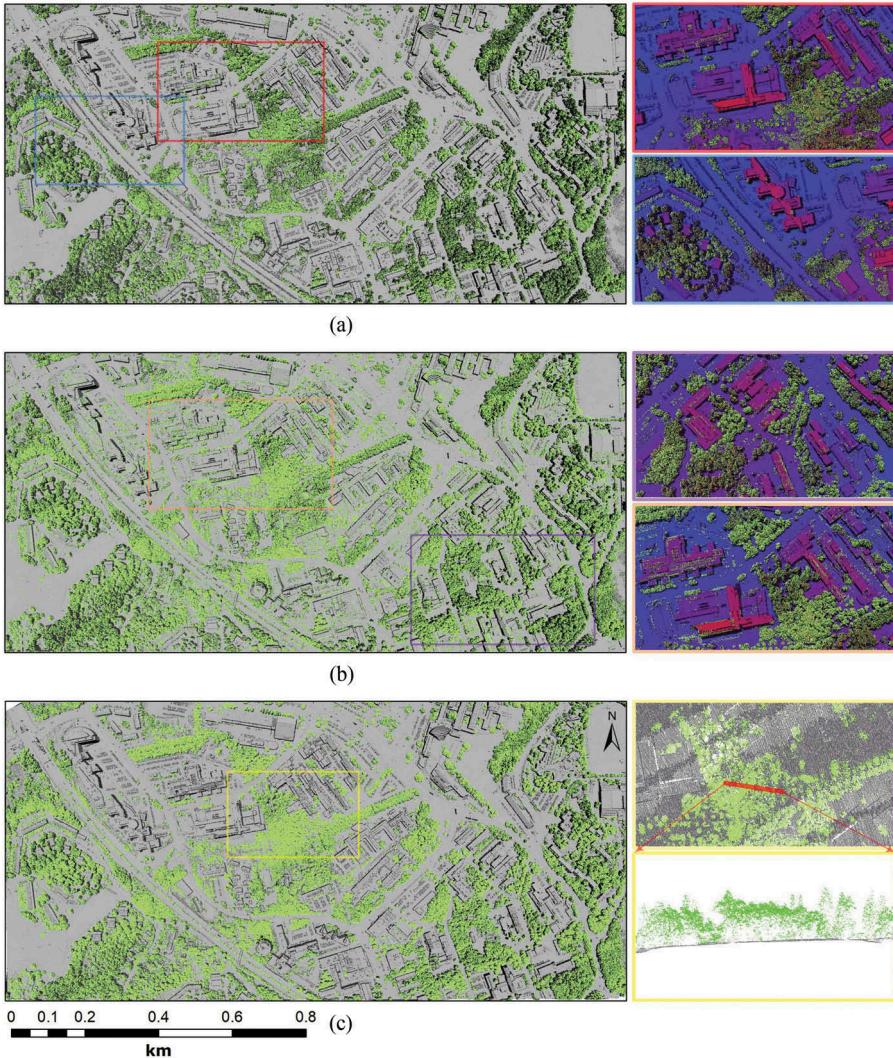


Figure 9. The result comparison of the Helsinki data set. Subfigures (a) and (b) are vegetation classification results by Zhou and Neumann (2008) and Horvat, Žalik, and Mongus (2016), respectively, while (c) is our result. The right column represents the enlarged views of the rectangles from their respective results. The view of cross section in (c) comes from the solid red line. For better visualisation, the extracted vegetation points are overlaid on their corresponding shaded relief maps. The light green points are vegetation, while the grey colour represents non-vegetation.

between commission and omission error and keep the overall accuracy of vegetation classification rate at high level.

5. Conclusions and future works

In this paper, we proposed a framework to separate vegetation from non-vegetation points in full-waveform ALS data using a supervoxel-based methodology. The entire

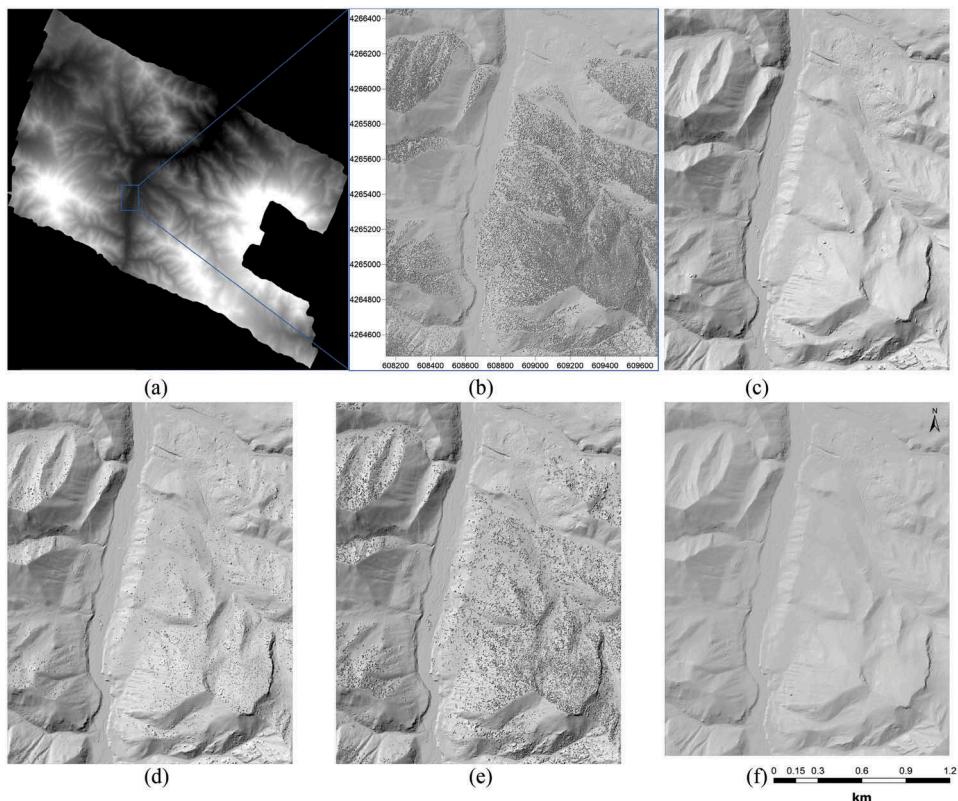


Figure 10. The result comparison of the smaller region in Dayekou data set. (a) The digital surface model of the entire dayekou data set. (b) The shaded relief digital surface map. (c) Our terrain results. Subfigures (d–f) are the shade relief map generated from terrain points by methods of Zhou and Neumann (2008), Horvat, Žalik, and Mongus (2016), and Zhang et al. (2003), respectively.

processing is automatic except the selection of training samples. Specifically, we have formulated a novel optimal energy function, which allows free combination of multiple kernels according to the complexity of backscattered full-waveform profiles. The optimal function can produce accurate discrete point clouds associated with discriminative full-waveform features. Moreover, we have also enhanced existing PDC unsupervised technique for clustering the points into supervoxels, which relies on the probability density function analysis of the point clouds and the results indicate its stability and the robustness for the two data sets tested with only one input parameter. Finally, the point-based features within each supervoxel are generalised by LDA as each supervoxel's feature. These supervoxel features with uniform size are then provided as input to supervoxel-based SVM and RF classifiers. Our experiments on Helsinki and Dayekou data sets outperform state-of-the-art methods and achieves at least 96.47% overall vegetation detection accuracy through a trade-off between missing alarm rate and false alarm rate.

Although extracting the discriminative features of object from point clouds that are corrupted with noise, outliers, and missing data is a key step to producing excellent classification results, the classifier itself is also another significant factor affecting classification precision

and accuracy. Therefore, in future work, we focus on the following two issues: what type of classifiers can help in achieving more promising results and which strategies (e.g. segmenting a specific object at each time or multiple objects simultaneously) can help in improving the overall accuracy of point clouds. In addition, for full-waveform modelling algorithm, we will further verify and incorporate other kernels possibly Gumbel and Weibull kernels into our kernel library to further decrease the fitting error, though this has to be traded against the computational complexity of solving the model equation

Note

1. <https://www.terrasolid.com/home.php>.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported in part by the National Natural Science Foundation of China under Grant 41301521 and Grant 41701533, and in part by the Open Funds of the State Key Laboratory of Remote Sensing Science under Grant OFSLRSS201516, and Beijing Key Laboratory of Urban Spatial Information Engineering under Grant 2017205.

ORCID

Dong Chen  <http://orcid.org/0000-0001-8118-3889>

References

- Alexander, C., P. K. Bøcher, L. Arge, and J.-C. Svenning. 2014. "Regional-Scale Mapping of Tree Cover, Height and Main Phenological Tree Types Using Airborne Laser Scanning Data." *Remote Sensing of Environment* 147: 156–172. doi:[10.1016/j.rse.2014.02.013](https://doi.org/10.1016/j.rse.2014.02.013).
- Alexander, C., K. Tansey, J. Kaduk, D. Holland, and N. J. Tate. 2010. "Backscatter Coefficient as an Attribute for the Classification of Full-Waveform Airborne Laser Scanning Data in Urban Areas." *ISPRS Journal of Photogrammetry and Remote Sensing* 65 (5): 423–432. doi:[10.1016/j.isprsjprs.2010.05.002](https://doi.org/10.1016/j.isprsjprs.2010.05.002).
- Bretar, F., A. Chauve, J.-S. Bailly, C. Mallet, and A. Jacome. 2009. "Terrain Surfaces and 3-D Landcover Classification from Small Footprint Full-Waveform Lidar Data: Application to Badlands." *Hydrology and Earth System Sciences* 13 (8): 1531–1544. doi:[10.5194/hess-13-1531-2009](https://doi.org/10.5194/hess-13-1531-2009).
- Bright, B. C., A. T. Hudak, R. McGaughey, H.-E. Andersen, and J. Negrón. 2013. "Predicting Live and Dead Tree Basal Area of Bark Beetle Affected Forests from Discrete-Return Lidar." *Canadian Journal of Remote Sensing* 39 (sup1): S99–S111. doi:[10.5589/m13-027](https://doi.org/10.5589/m13-027).
- Carlberg, M., P. Gao, G. Chen, and A. Zakhor. 2009. "Classifying Urban Landscape in Aerial LiDAR Using 3D Shape Analysis." In *Image Processing (ICIP), 2009 16th IEEE International Conference On*, November 7-12, Cairo, Egypt. 1701–1704. Piscataway, NJ: IEEE.
- Chang, C.-C., and C.-J. Lin. 2011. "LIBSVM: A Library for Support Vector Machines." *ACM Transactions on Intelligent Systems and Technology Software* available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm> 2: 27:1–27:27. doi:[10.1145/1961189.1961199](https://doi.org/10.1145/1961189.1961199).



- Chen, D., L. Zhang, Z. Wang, and H. Deng. 2013. "A Mathematical Morphology-Based Multi-Level Filter of LiDAR Data for Generating DTMs." *Science China Information Sciences* 56 (10): 1–14. doi:[10.1007/s11432-012-4707-3](https://doi.org/10.1007/s11432-012-4707-3)
- Chen, G., and A. Zakhor. 2009. "2d Tree Detection in Large Urban Landscapes Using Aerial Lidar Data." In *Image Processing (ICIP), 2009 16th IEEE International Conference On*, November 7–12, Cairo, Egypt. 1693–1696. Piscataway, NJ: IEEE.
- Cheng, Y. 1995. "Mean Shift, Mode Seeking, and Clustering." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (8): 790–799. doi:[10.1109/34.400568](https://doi.org/10.1109/34.400568).
- De Almeida Furtado, L. F., T. S. F. Silva, and E. M. L. De Moraes Novo. 2016. "Dual-Season and Full-Polarimetric C Band SAR Assessment for Vegetation Mapping in the Amazon Várzea Wetlands." *Remote Sensing of Environment* 174: 212–222. doi:[10.1016/j.rse.2015.12.013](https://doi.org/10.1016/j.rse.2015.12.013).
- Ersahin, K., I. G. Cumming, and R. K. Ward. 2010. "Segmentation and Classification of Polarimetric SAR Data Using Spectral Graph Partitioning." *IEEE Transactions on Geoscience and Remote Sensing* 48 (1): 164–174. doi:[10.1109/TGRS.2009.2024303](https://doi.org/10.1109/TGRS.2009.2024303).
- Ferraz, A., S. Saatchi, C. Mallet, S. Jacquemoud, G. Gonçalves, C. Silva, P. Soares, M. Tomé, and L. Pereira. 2016. "Airborne Lidar Estimation of Aboveground Forest Biomass in the Absence of Field Inventory." *Remote Sensing* 8 (8): 653. doi:[10.3390/rs8080653](https://doi.org/10.3390/rs8080653).
- Ghulam, A., I. Porton, and K. Freeman. 2014. "Detecting Subcanopy Invasive Plant Species in Tropical Rainforest by Integrating Optical and Microwave (Insar/Polinsar) Remote Sensing Data, and a Decision Tree Algorithm." *ISPRS Journal of Photogrammetry and Remote Sensing* 88: 174–192. doi:[10.1016/j.isprsjprs.2013.12.007](https://doi.org/10.1016/j.isprsjprs.2013.12.007).
- Gross, H., and U. Thoennesen. 2006. "Extraction of Lines from Laser Point Clouds." *Symposium of ISPRS Commission III: Photogrammetric Computer Vision PCV06. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36: 86–91.
- Guan, H., Y. Yu, Z. Ji, J. Li, and Q. Zhang. 2015. "Deep Learning-Based Tree Classification Using Mobile LiDAR Data." *Remote Sensing Letters* 6 (11): 864–873. doi:[10.1080/2150704X.2015.1088668](https://doi.org/10.1080/2150704X.2015.1088668).
- Guo, B., X. Huang, F. Zhang, and G. Sohn. 2015. "Classification of Airborne Laser Scanning Data Using JointBoost." *ISPRS Journal of Photogrammetry and Remote Sensing* 100: 71–83. doi:[10.1016/j.isprsjprs.2014.04.015](https://doi.org/10.1016/j.isprsjprs.2014.04.015).
- Guo, L., N. Chehata, C. Mallet, and S. Boukir. 2011. "Relevance of Airborne Lidar and Multispectral Image Data for Urban Scene Classification Using Random Forests." *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (1): 56–66. doi:[10.1016/j.isprsjprs.2010.08.007](https://doi.org/10.1016/j.isprsjprs.2010.08.007).
- Haala, N., and C. Brenner. 1999. "Extraction of Buildings and Trees in Urban Environments." *ISPRS Journal of Photogrammetry and Remote Sensing* 54 (2–3): 130–137. doi:[10.1016/S0924-2716\(99\)00010-6](https://doi.org/10.1016/S0924-2716(99)00010-6).
- Hastings, W. K. 1970. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications." *Biometrika* 57 (1): 97–109. doi:[10.1093/biomet/57.1.97](https://doi.org/10.1093/biomet/57.1.97).
- Höfle, B., M. Hollaus, and J. Hagenauer. 2012. "Urban Vegetation Detection Using Radiometrically Calibrated Small-Footprint Full-Waveform Airborne LiDAR Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 67: 134–147. doi:[10.1016/j.isprsjprs.2011.12.003](https://doi.org/10.1016/j.isprsjprs.2011.12.003).
- Höfle, B., and N. Pfeifer. 2007. "Correction of Laser Scanning Intensity Data: Data and Model-Driven Approaches." *ISPRS Journal of Photogrammetry and Remote Sensing* 62 (6): 415–433. doi:[10.1016/j.isprsjprs.2007.05.008](https://doi.org/10.1016/j.isprsjprs.2007.05.008).
- Horvat, D., B. Žalik, and D. Mongus. 2016. "Context-Dependent Detection of Non-Linearly Distributed Points for Vegetation Classification in Airborne LiDAR." *ISPRS Journal of Photogrammetry and Remote Sensing* 116: 1–14. doi:[10.1016/j.isprsjprs.2016.02.011](https://doi.org/10.1016/j.isprsjprs.2016.02.011).
- Hu, B., J. Li, L. Jing, and A. Judah. 2014. "Improving the Efficiency and Accuracy of Individual Tree Crown Delineation from High-Density LiDAR Data." *International Journal of Applied Earth Observation and Geoinformation* 26: 145–155. doi:[10.1016/j.jag.2013.06.003](https://doi.org/10.1016/j.jag.2013.06.003).
- Johnson, A. E., and M. Hebert. 1999. "Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21 (5): 433–449. doi:[10.1109/34.765655](https://doi.org/10.1109/34.765655).
- Ko, C., G. Sohn, T. Remmel, and J. Miller. 2016. "Maximizing the Diversity of Ensemble Random Forests for Tree Genera Classification Using High Density LiDAR Data." *Remote Sensing* 8 (8): 646. doi:[10.3390/rs8080646](https://doi.org/10.3390/rs8080646).

- Lehner, H., and C. Briese. 2010. "Radiometric Calibration of Fullwaveform Airborne Laser Scanning Data Based on Natural Surfaces." *International Archives Photogramm.Remote Sens Spat Information Sciences* 38 (Part 7B): 360–365.
- Lu, D., and Q. Weng. 2007. "A Survey of Image Classification Methods and Techniques for Improving Classification Performance." *International Journal of Remote Sensing* 28 (5): 823–870. doi:[10.1080/01431160600746456](https://doi.org/10.1080/01431160600746456).
- Mallet, C., F. Bretar, M. Roux, U. Soergel, and C. Heipke. 2011. "Relevance Assessment of Full-Waveform Lidar Data for Urban Area Classification." *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (6): S71–S84. doi:[10.1016/j.isprsjprs.2011.09.008](https://doi.org/10.1016/j.isprsjprs.2011.09.008).
- Matikainen, L., K. Karila, J. Hyppä, P. Litkey, E. Puttonen, and E. Ahokas. 2017. "Object-Based Analysis of Multispectral Airborne Laser Scanner Data for Land Cover Classification and Map Updating." *ISPRS Journal of Photogrammetry and Remote Sensing* 128: 298–313. doi:[10.1016/j.isprsjprs.2017.04.005](https://doi.org/10.1016/j.isprsjprs.2017.04.005).
- Neumann, M., L. Ferro-Famil, and A. Reigber. 2010. "Estimation of Forest Structure, Ground, and Canopy Layer Characteristics from Multibaseline Polarimetric Interferometric SAR Data." *IEEE Transactions on Geoscience and Remote Sensing* 48 (3): 1086–1104. doi:[10.1109/TGRS.2009.2031101](https://doi.org/10.1109/TGRS.2009.2031101).
- Neumann, M., S. S. Saatchi, L. M. H. Ulander, and J. E. S. Fransson. 2012. "Assessing Performance of L- and P-Band Polarimetric Interferometric SAR Data in Estimating Boreal Forest Above-Ground Biomass." *IEEE Transactions on Geoscience and Remote Sensing* 50 (3): 714–726. doi:[10.1109/TGRS.2011.2176133](https://doi.org/10.1109/TGRS.2011.2176133).
- Niemeyer, J., J. D. Wegner, C. Mallet, F. Rottensteiner, and U. Soergel. 2011. "Conditional Random Fields for Urban Scene Classification with Full Waveform LiDAR Data." In *ISPRS Conference on Photogrammetric Image Analysis 2011*, Munich, Germany, October 5-7, edited by Uwe Stilla, Franz Rottensteiner, Helmut Mayer, Boris Jutzi, and Matthias Butenuth. 233–244. Berlin, Germany: Springer Berlin Heidelberg.
- Polewski, P., W. Yao, M. Heurich, P. Krzystek, and U. Stilla. 2015. "Detection of Fallen Trees in ALS Point Clouds Using a Normalized Cut Approach Trained by Simulation." *ISPRS Journal of Photogrammetry and Remote Sensing* 105: 252–271. doi:[10.1016/j.isprsjprs.2015.01.010](https://doi.org/10.1016/j.isprsjprs.2015.01.010).
- Polewski, P., W. Yao, M. Heurich, P. Krzystek, and U. Stilla. 2017. "Learning a Constrained Conditional Random Field for Enhanced Segmentation of Fallen Trees in ALS Point Clouds." *ISPRS Journal of Photogrammetry and Remote Sensing*. doi:[10.1016/j.isprsjprs.2017.04.001](https://doi.org/10.1016/j.isprsjprs.2017.04.001).
- Poullis, C. 2013. "A Framework for Automatic Modeling from Point Cloud Data." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35 (11): 2563–2575. doi:[10.1109/TPAMI.2013.64](https://doi.org/10.1109/TPAMI.2013.64).
- Puttonen, E., C. Briese, G. Mandlburger, M. Wieser, M. Pfennigbauer, A. Zlinszky, and N. Pfeifer. 2016. "Quantification of Overnight Movement of Birch (*Betula pendula*) Branches and Foliage with Short Interval Terrestrial Laser Scanning." *Frontiers in Plant Science* 7: 222. doi:[10.3389/fpls.2016.00222](https://doi.org/10.3389/fpls.2016.00222)
- Rapinel, S., B. Clément, S. Magnanom, V. Sellin, and L. Hubert-Moy. 2014. "Identification and Mapping of Natural Vegetation on a Coastal Site Using a Worldview-2 Satellite Image." *Journal of Environmental Management* 144: 236–246. doi:[10.1016/j.jenvman.2014.05.027](https://doi.org/10.1016/j.jenvman.2014.05.027).
- Reitberger, J., P. Krzystek, and U. Stilla. 2008. "Analysis of Full Waveform LIDAR Data for the Classification of Deciduous and Coniferous Trees." *International Journal of Remote Sensing* 29 (5): 1407–1431. doi:[10.1080/01431160701736448](https://doi.org/10.1080/01431160701736448).
- Rusu, R. B., N. Blodow, and M. Beetz. 2009. "Fast Point Feature Histograms (FPFH) for 3D Registration." In *IEEE International Conference on Robotics and Automation, 2009. ICRA'09*, Kobe, Japan, May 12–17. 3212–3217. Piscataway, NJ: IEEE.
- Rutzinger, M., B. Hfle, M. Hollaus, and N. Pfeifer. 2008. "Object-Based Point Cloud Analysis of Full-Waveform Airborne Laser Scanning Data for Urban Vegetation Classification." *Sensors* 8 (8): 4505–4528. doi:[10.3390/s8084505](https://doi.org/10.3390/s8084505).
- Salah, M., and J. Trinder. 2010. "Support Vector Machines Based Filtering of Lidar Data: A Grid Based Method." In *Proceeding of FIG Congress*, Sydney, Australia. Copenhagen: International Federation of Surveyors (FIG).



- Samadzadegan, F., B. Bigdeli, and P. Ramzi. 2010. "A Multiple Classifier System for Classification of LIDAR Remote Sensing Data Using Multi-Class SVM." In *9th International Workshop, MCS (Multiple Classifier Systems)*, Cairo, Egypt, April 7-9, 254–263. Berlin, Germany: Springer Berlin Heidelberg.
- Savitzky, A., and M. J. E. Golay. 1964. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures." *Analytical Chemistry* 36 (8): 1627–1639. doi:[10.1021/ac60214a047](https://doi.org/10.1021/ac60214a047).
- Schnabel, R., R. Wahl, and R. Klein. 2007. "Efficient RANSAC for Point-Cloud Shape Detection." *Computer Graphics Forum* 26 (2): 214–226. doi:[10.1111/cgf.2007.26.issue-2](https://doi.org/10.1111/cgf.2007.26.issue-2).
- Secord, J., and A. Zakhori. 2007. "Tree Detection in Urban Regions Using Aerial Lidar and Image Data." *IEEE Geoscience and Remote Sensing Letters* 4 (2): 196–200. doi:[10.1109/LGRS.2006.888107](https://doi.org/10.1109/LGRS.2006.888107).
- Soja, M. J., H. Persson, and L. M. H. Ulander. 2015. "Estimation of Forest Height and Canopy Density from a Single InSAR Correlation Coefficient." *IEEE Geoscience and Remote Sensing Letters* 12 (3): 646–650. doi:[10.1109/LGRS.2014.2354551](https://doi.org/10.1109/LGRS.2014.2354551).
- Vosselman, G., M. Coenen, and F. Rottensteiner. 2017. "Contextual Segment-Based Classification of Airborne Laser Scanner Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 128: 354–371. doi:[10.1016/j.isprsjprs.2017.03.010](https://doi.org/10.1016/j.isprsjprs.2017.03.010).
- Wagner, W., M. Hollaus, C. Briese, and V. Ducic. 2008a. "3D Vegetation Mapping Using Small-Footprint Full-Waveform Airborne Laser Scanners." *International Journal of Remote Sensing* 29 (5): 1433–1452. doi:[10.1080/01431160701736398](https://doi.org/10.1080/01431160701736398).
- Wagner, W., J. Hyppä, A. Ullrich, H. Lehner, C. Briese, and S. Kaasalainen. 2008b. "Radiometric Calibration of Full-Waveform Small-Footprint Airborne Laser Scanners." *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 37 (B1): 163–168.
- Wagner, W. 2010. "Radiometric Calibration of Small-Footprint Full-Waveform Airborne Laser Scanner Measurements: Basic Physical Concepts." *ISPRS Journal of Photogrammetry and Remote Sensing* 65 (6): 505–513. doi:[10.1016/j.isprsjprs.2010.06.007](https://doi.org/10.1016/j.isprsjprs.2010.06.007).
- Wagner, W., A. Ullrich, V. Ducic, T. Melzer, and N. Studnicka. 2006. "Gaussian Decomposition and Calibration of a Novel Small-Footprint Full-Waveform Digitising Airborne Laser Scanner." *ISPRS Journal of Photogrammetry and Remote Sensing* 60 (2): 100–112. doi:[10.1016/j.isprsjprs.2005.12.001](https://doi.org/10.1016/j.isprsjprs.2005.12.001).
- Wang, Z., L. Zhang, P. Tian Fang, T. Mathiopoulos, X. Tong, Q. Huamin, Z. Xiao, L. Fang, and D. Chen. 2015. "A Multiscale and Hierarchical Feature Extraction Method for Terrestrial Laser Scanning Point Cloud Classification." *IEEE Transactions on Geoscience and Remote Sensing* 53 (5): 2409–2425. doi:[10.1109/TGRS.2014.2359951](https://doi.org/10.1109/TGRS.2014.2359951).
- Xia, S., C. Wang, F. Pan, X. Xiaohuan, H. Zeng, and H. Liu. 2015. "Detecting Stems in Dense and Homogeneous Forest Using Single-Scan TLS." *Forests* 6 (11): 3923–3945. doi:[10.3390/f6113923](https://doi.org/10.3390/f6113923).
- Xu, S., G. Vosselman, and S. Oude Elberink. 2014. "Multiple-Entity Based Classification of Airborne Laser Scanning Data in Urban Areas." *ISPRS Journal of Photogrammetry and Remote Sensing* 88: 1–15. doi:[10.1016/j.isprsjprs.2013.11.008](https://doi.org/10.1016/j.isprsjprs.2013.11.008).
- Yan, W. Y., A. Shaker, and N. El-Ashmawy. 2015. "Urban Land Cover Classification Using Airborne LiDAR Data: A Review." *Remote Sensing of Environment* 158: 295–310. doi:[10.1016/j.rse.2014.11.001](https://doi.org/10.1016/j.rse.2014.11.001).
- Yang, B., and Z. Dong. 2013. "A Shape-Based Segmentation Method for Mobile Laser Scanning Point Clouds." *ISPRS Journal of Photogrammetry and Remote Sensing* 81: 19–30. doi:[10.1016/j.isprsjprs.2013.04.002](https://doi.org/10.1016/j.isprsjprs.2013.04.002).
- Yao, W., and Y. Wei. 2013. "Detection of 3-D Individual Trees in Urban Areas by Combining Airborne LiDAR Data and Imagery." *IEEE Geoscience and Remote Sensing Letters* 10 (6): 1355–1359. doi:[10.1109/LGRS.2013.2241390](https://doi.org/10.1109/LGRS.2013.2241390).
- Yu, P. A., K. Qin, and D. A. Clausi. 2012. "Unsupervised Polarimetric SAR Image Segmentation and Classification Using Region Growing with Edge Penalty." *IEEE Transactions on Geoscience and Remote Sensing* 50 (4): 1302–1317. doi:[10.1109/TGRS.2011.2164085](https://doi.org/10.1109/TGRS.2011.2164085).
- Yu, Q., P. Gong, N. Clinton, G. Biging, M. Kelly, and D. Schirokauer. 2006. "Object-Based Detailed Vegetation Classification with Airborne High Spatial Resolution Remote Sensing Imagery." *Photogrammetric Engineering & Remote Sensing* 72 (7): 799–811. doi:[10.14358/PERS.72.7.799](https://doi.org/10.14358/PERS.72.7.799).

- Zhang, J., and G. Sohn. 2010. "A Markov Random Field Model for Individual Tree Detection from Airborne Laser Scanning Data." *Proceedings of Photogrammetric Computer Vision (PCV) 2010*: 01–03.
- Zhang, K., S.-C. Chen, D. Whitman, M.-L. Shyu, J. Yan, and C. Zhang. 2003. "A Progressive Morphological Filter for Removing Nonground Measurements from Airborne LIDAR Data." *IEEE Transactions on Geoscience and Remote Sensing* 41 (4): 872–882. doi:[10.1109/TGRS.2003.810682](https://doi.org/10.1109/TGRS.2003.810682).
- Zhang, Y., L. Dengsheng, B. Yang, C. Sun, and M. Sun. 2011. "Coastal Wetland Vegetation Classification with a Landsat Thematic Mapper Image." *International Journal of Remote Sensing* 32 (2): 545–561. doi:[10.1080/01431160903475241](https://doi.org/10.1080/01431160903475241).
- Zhang, Z., L. Zhang, X. Tong, P. Takis Mathiopoulos, B. Guo, X. Huang, Z. Wang, and Y. Wang. 2016. "A Multilevel Point-Cluster-Based Discriminative Feature for ALS Point Cloud Classification." *IEEE Transactions on Geoscience and Remote Sensing* 54 (6): 3309–3321. doi:[10.1109/TGRS.2016.2514508](https://doi.org/10.1109/TGRS.2016.2514508).
- Zhou, Q.-Y., and U. Neumann. 2008. "Fast and Extensible Building Modeling from Airborne LiDAR Data." In *Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Irvine, CA, edited by W. G. Aref, F. Mokbel, H. Samet, M. Schneider, C. Shahabi, and O. Wolfson, 7–15. New York, NY: ACM.