Int J Appl Earth Obs Geoinformation 67 (2018) 135–147

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

journal homepage: www.elsevier.com/locate/jag

Detection of dead standing Eucalyptus camaldulensis without tree delineation for managing biodiversity in native Australian forest

Milto Miltiadou[a,b,c,d,*], Neil D.F. Campbell[a], Susana Gonzalez Aracil[c], Tony Brown[e], Michael G. Grant[b]

[a] Department of Computer Science, University of Bath, Bath, United Kingdom
[b] Airborne Research Facility, Plymouth Marine Laboratory, Plymouth, United Kingdom
[c] Interpine Group Ltd, Rotorua, New Zealand
[d] Department of Civil Engineering and Geomatics, Cyprus University of Technology, Limassol, Cyprus
[e] Forestry Corporation of NSW, Wauchope, Australia

ARTICLE INFO

Keywords:
Full-waveform (FW) LiDAR
Biodiversity
Eucalyptus camaldulensis
Native forest

ABSTRACT

In Australia, many birds and arboreal animals use hollows for shelters, but studies predict shortage of hollows in near future. Aged dead trees are more likely to contain hollows and therefore automated detection of them plays a substantial role in preserving biodiversity and consequently maintaining a resilient ecosystem. For this purpose full-waveform LiDAR data were acquired from a native Eucalypt forest in Southern Australia. The structure of the forest significantly varies in terms of tree density, age and height. Additionally, Eucalyptus camaldulensis have multiple trunk splits making tree delineation very challenging. For that reason, this paper investigates automated detection of dead standing Eucalyptus camaldulensis without tree delineation. It also presents the new feature of the open source software DASOS, which extracts features for 3D object detection in voxelised FW LiDAR. A random forest classifier, a weighted-distance KNN algorithm and a seed growth algorithm are used to create a 2D probabilistic field and to then predict potential positions of dead trees. It is shown that tree health assessment is possible without tree delineation but since it is a new research directions there are many improvements to be made.

1. Introduction

1.1. The importance of dead wood

The value of dead trees from a biodiversity management perspective is large. Once a tree dies, its woody structure remains for centuries and it contributes to forest regeneration while providing resources for numerous surrounding organisms (Franklin et al., 1987). More than 4000 species inhabit dead wood in Finland (Siitonen, 2001), where an estimate of 1000 species are threatened (Hanski, 2000). These species include animals, birds and other organisms, like fungi. Fungi contributes to wood decaying, formation of hollows and biodiversity, which supports the resilience of our ecosystem (Peterson et al., 1998).

In Australia, tree hollows play a significant role in managing biodiversity (Lindenmayer et al., 1997; Bennett et al., 1994). Nearly all arboreal mammals rely on hollows with the exception of the Koala (*Phascolarctos cinereus*) and perhaps Ringtail Possums (*Pseudocheirus peregrinus*) that preferentially make a stick nest. Additionally, numerous Australian bird species use hollows for shelters (Gibbons and Lindenmayer, 2002). Nevertheless, Australia has no real hollow creators unlike the northern hemisphere (e.g. Woodpeckers), and therefore it relies predominantly on natural processes of limb breakage, insect and fungal attack when access points are provided through damage caused by wind, storms and fire. This kind of hollows takes hundreds of years to form (Wormington and Lamb, 1999).

According to Gibbons et al. (2000), hollows are more likely to exist on dead trees or trees in poor physiological condition. In Australia, studies predict shortage of hollows for colonisation in the near future (Lindenmayer and Wood, 2010; Goldingay, 2009). A sample list of species that rely on hollows, provided by Forestry Corporation of NSW, is depicted at Fig. 1. Three of them are threatened (New South Wales Government, 2016). Consequently, automated detection of dead trees plays a substantial role in managing biodiversity.

As explained above, monitoring dead trees is essential for preserving a resilient ecosystem. Nevertheless, their distribution significantly varies making detection of them difficult (Kim et al., 2009). Remote sensing automates the process of monitoring forest and increases the spatial resolution of the monitored area.

* Corresponding author at: Department of Computer Science, University of Bath, Bath, United Kingdom.
E-mail address: milto_miltiadou@hotmail.com (M. Miltiadou).

https://doi.org/10.1016/j.jag.2018.01.008
Received 31 October 2017; Received in revised form 11 January 2018; Accepted 15 January 2018

Fig. 1. Some species that uses tree hollows for shelters. The red ones/bold ones are threatened: Kookaburra, Sulphur Crested Cockatoo, Corella, Crimson Rosella, Eastern Rosella, Galah, Rainbow Lorikeet, Musk Lorikeet, Little Lorikeet, Red-winged Parrot, **Superb Parrot**, Cockatiel, Australian Ringneck (Parrot), Red-rumped Parrot, Powerful Owl, Sooty Ow, Barking Owl, **Masked Owl**, **Barn Owl**, White-throated Treecreeper, Hollow Owl, Brush-tailed Possum. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

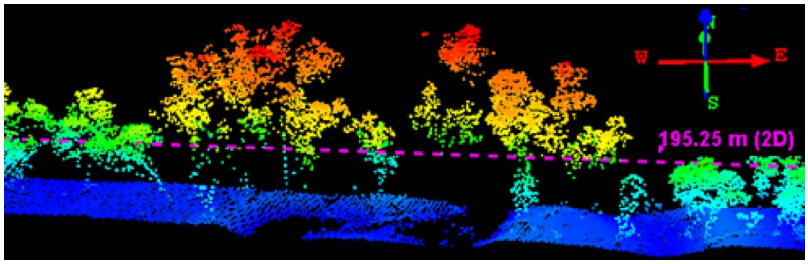


Fig. 2. Discrete LiDAR point cloud indicating the missing information about the trunks.

1.2. Related work

Remote sensing was introduced for automating detection of dead trees since fieldwork is a time consuming task, considering the variance spread of trees and the spatial resolution of the area of interest. From a classification perceptive, the task of identifying dead standing and dead fallen trees is different. Fallen trees are identified by detecting segments or line-like features on the terrain surface using LiDAR (Polewski et al., 2015; Mcke et al., 2013). Regarding standing dead trees, their shape (reduced number of leaves or broken branches) (Yao et al., 2012) and light reflectance (less green light illuminated) (Pasher and King, 2009) are important factors for identifying them.

Previous work on dead standing trees detection performs single tree crown delineation before health assessment (Yao et al., 2012; Shendryk et al., 2016b). Tree crown delineation is usually done by detecting local maxima from the canopy height model (CHM) and then segmenting trees using the watershed algorithm (Popescu et al., 2003). Improvements has been achieved by introducing markers controlled watershed (Jing et al., 2012) and structural elements of tree crowns with different sizes (Hu et al., 2014). Additionally, Popescu and Zhao (2008) analyse the vertical distribution of the LiDAR points in conjunction with the local maximum filtering of CHM.

In the case of Eucalyptus camaldulensis, tree delineation is a challenge due to their irregular structure and multiple trunk splits. Local maxima filtering, used for tree detection, leads to over-segmentation because each tree trunk split forms a local maxima. Shendryk et al. (2016a) published a Eucalyptus delineation algorithm that performs segmentation from bottom to top; the trunks point cloud is separated from the leaves and individual trunks are identified before the segmentation. Nevertheless, the density resolution starts from 12 points/m^2 and goes up to 36 points/m^2 around forested areas. For small research projects capturing this high resolution is acceptable, but for larger areas, the density of the emitted pulses is above the optimal resolution for a cost effective versus quality acquisition (Lovell et al., 2005). The investigated area of this paper is 95,196 ha and the entire area was scanned. The resolution of our acquired data has a minimum specification of four pulses per square meter. This is considered an optimal resolution in respect to the cost. Nevertheless, due to tree heights (up to 43 m) and acquired pulse density, there are not enough peak returns from trunks to enable bottom to up delineation. An example of the missing information about the trunks is depicted in Fig. 2.

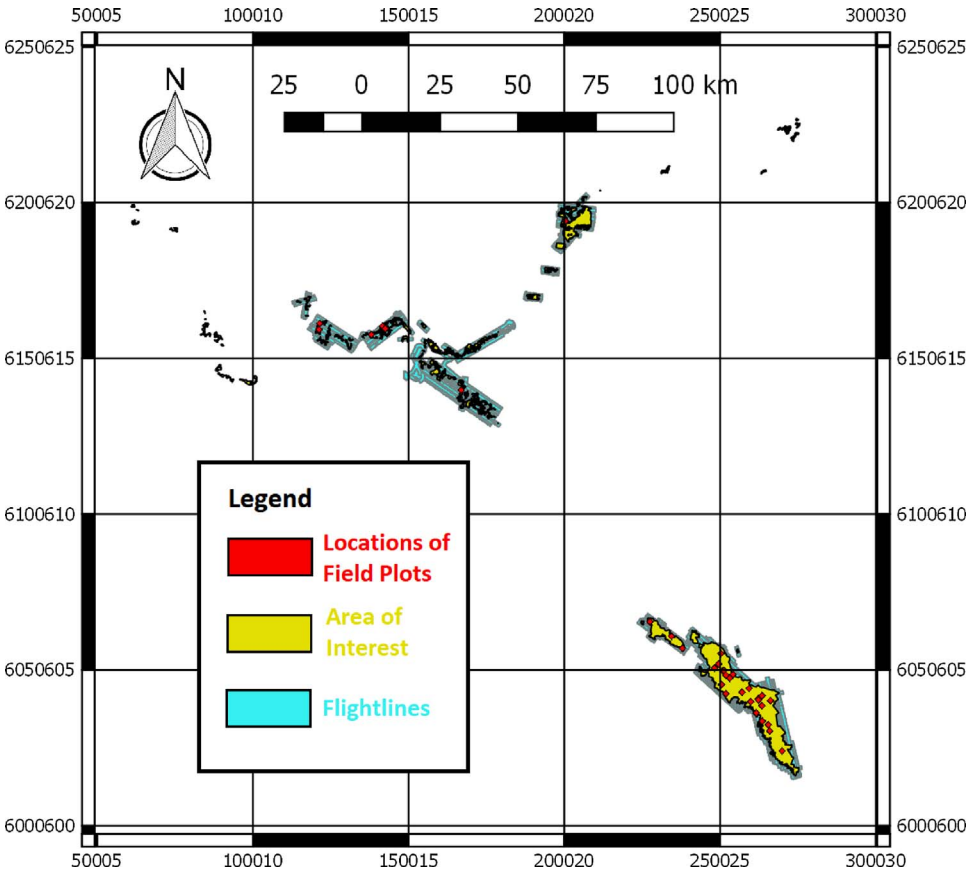


Fig. 3. Study area.

Fig. 4. Example of dead trees indicating their variance in shape.

Table 1
Number of trees according to their DBH.

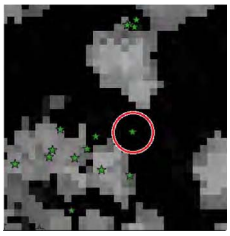
DBH (mm)	Dead trees	Alive trees
> 2000	0	1
1000–2000	7	21
600–1000	8	146
400–600	26	290
300–400	32	286
200–300	50	462
100–200	125	904
< 100	11	16
Total	260	2126

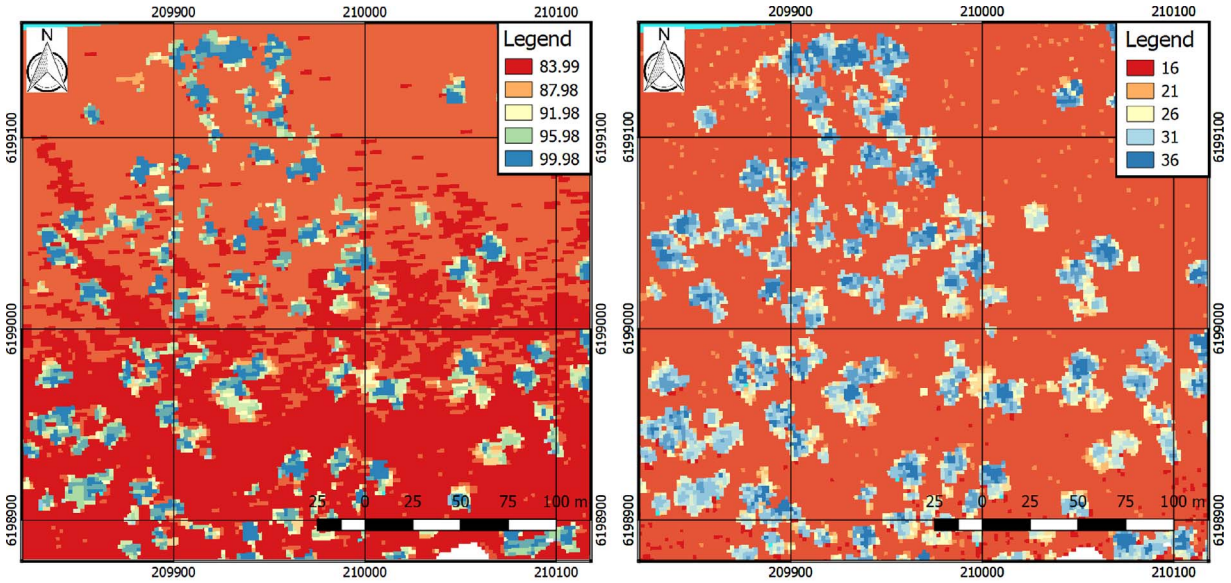
2. Materials

2.1. Study area

The study area (Fig. 3) is a native River Red Gum (Eucalypt camaldulensis) forest of size 95,196 ha^2 in south-eastern Australia. The regeneration of the Eucalyptus camaldulensis is extremely dependant on floods and therefore, their distribution in respect to density, health and age is highly variance (Kerle, 2005). Additionally, the height of Eucalyptus camaldulensis reaches up to 40 m and their structural complexity is high with multiple trunk splits (Wilson and of N.S.W, 1995). Fig. 4 shows the structure of the forest and the shape variations of dead trees.

Table 2
The classification challenges of automated detection of dead Eucalyptus camaldulensis.

Study area	Acquired data	Field data
• The study area is a native Eucalypt forest. Native forests contain trees of different ages and heights. The height of a dead tree could be within the range of [1.5,40] meters. • There is a high variance in the density of the forest. The testing/training samples of small trees may contain information from nearby alive trees or ground. • A tree may have dead branches but still be alive. • Eucalyptus camaldulensis have irregular shapes and multiple trunk splits making tree delineation to require very dense acquired data.	• The pulse density of the acquired data does not allow bottom to top tree delineation. Additionally, crown detection from DEM leads to over-segmentation due to the multiple trunk splits. • Dead trees are identifiable by their light reflectance, but the acquired data do not contain this kind of information. Therefore, the classifier depends predominantly on tree shape, which is not an independent factor of identifying dead trees; a tree may not have leaves but still be alive.	• If a tree has a trunk split below the 1.3 m height, then it is recorded as multiple trees; inconsistent meaning of the "one tree" term. • Field data contain small trees, that are non detectable within the FW LiDAR. • Unknown accuracy of the geo-locations of the trees; Some trees appear to be on the ground, once visualised on top of DEM: 



(a) The DEM before subtracting the DTM (b) The DEM after subtracting the DTM

Fig. 5. The difference of the DEM before and after subtracting the DTM.

2.2. Acquired full-waveform (FW) LiDAR

FW LiDAR are acquired by RPS Australia East Pty Ltd from 900 m above ground level, using a Trimble AX60 Airborne LiDAR sensor. The wavelength of the emitted laser was 1062 nm, the maximum scan angle was 60°, and the pulse rate was 400 kHz. The acquisition was held from March 6th till March 31st, 2015. The collected LiDAR were delivered into 206 flightlines, of which 13 are cross runs for geometric correction. There is also a 30% of swath overlap. The footprint spacing of the pulses is 4.3 footprints per square meter.

Traditional ways of interpreting FW LiDAR, suggests extraction of a denser points cloud using Gaussian decomposition (Neuenschwander et al., 2009; Reitberger et al., 2008). This project implements on the open source software DASOS (Miltiadou et al., 2015), developed by the authors of this paper. Influenced by Persson et al. (2005), who used voxelisation to visualise the waveforms, DASOS uses voxelisation for both visualisations and classification. It further normalises the intensities so that equal pulse length exists inside each voxel, making intensities more meaningful. The literature is also moving towards voxelisation with promising results obtained by recent publications (Cao et al., 2016; Hancock et al., 2017). Generally, DASOS aims to ease manipulation of FW LiDAR in a volumetric representation. Its first two features (polygal mesh extraction and aligned metrics with hyperspectral imagery) was presented at the ISRSE conference in 2015. This paper presents a new feature implemented and how it is useful for assessing forest health.

2.3. Field data

The field data were collected in July 2015 during the winter season of Australia by Interpine Group Ltd and Forestry Corporation of NSW. There are 33 plots with radius 35.68 m and area 0.4 ha allocated randomly inside the study area. A total of 2386 trees were individually measured. Tree measurements include the geo-location, the trunk diameter at the breast height (1.3 m), species and health conditions (i.e. dead or alive). The geo-location of each tree was calculated at post-processing using the recorded magnetic bearing from the centroid of the plot and the distance from the centroid. Here, it is worth mentioning that a single tree may be recorded as multiple trees if there is a trunk split bellow the breast height of 1.3 m. Furthermore, 91.59% of the

Table 3
Explanation of a sample, the most relevant processed features, exported by DASOS.

No	Label	Description
1	Height_Middle_Column	The height of the middle column of the shape
	Height_Mean	The Mean height of all the columns
	Height_Median	The Median height of all the columns
1	Height_Std	The Standard Deviation of the heights
2	Top_Patch_Len_Std	The Standard Deviation of all the top patches
3	Dis_Std	The Standard Deviation of the distances between the central voxel and every non-empty voxel
4	Per_Int_Above_Iso	Percentage of voxels that contain an intensity above the isolevel
5	Top_Patch_Len_Mean	The Mean length of all the top patches
	Top_Patch_Len_Median	The Median length of all the top patches
7	Dis_Mean	Mean distance from the central voxel to every non-empty voxel
8	Dis_Median	Median distance from the central voxel to every non-empty voxel
9	Sum_Int_Diff_Z	The Mirror Summed Difference of the intensities using the middle column in the *z*-axis as the axis of symmetry
10	Sum_Int_Diff_X	The Mirror Summed Difference of the intensities using the middle column in the *x*-axis as the axis of symmetry

trees are River Red Gum.

The field data contain information for 260 dead trees. Nevertheless, not all of them are considered useful for biodiversity. Dead trees with big Diameter at Breast Height (DBH) are more likely to contain hollows. Additionally, trees with DBH smaller than the footprint spacing of the LiDAR are not identifiable from the FW LiDAR. Table 1 shows the number of dead and alive trees in respect to their DBH. Please note that for training the classifier equal numbers of dead and alive trees have been used to reduce statistical bias.

Within the field data, some plots exist on two flightlines due to flights overlap. Overlaps exist at the edges of the flightlines and their scan angle significantly varies. For that reason, each unique set of field plot and flightline is considered as a test/training plot. This results into 50 plots used for training and cross-validation.

3. Classification challenges

Table 2 underlines the challenges faced while working on the detection of dead standing Eucalyptus camaldulensis. The challenges are categorised into three groups (the nature of the study area, the acquired data, the field data) and they all influence the quality of the classifier.

4. Methods and algorithms

This sections explains the methodology applied. In a few words, Random Forest is applied for identifying the most significant features. Then, a weighted KNN algorithm is used for generating a probabilistic field. Once ground pixels are removed, a seed growth algorithm segments potential dead trees and positions are finally assigned. More details, including intermediate steps for noise reduction, are explained below.

But before proceeding to the details, it worth highlighting the reasons of choosing those algorithms. As shown in Fig. 4, the shapes of the dead trees significantly vary from one to another. Therefore, during statistical analysis they may form multiple clusters with variant shapes and tangle between the clusters of the alive trees. For that reason, it is preferable to use a classification algorithm that supports non-linear boundaries and makes no assumptions of them. All parametric classifiers (e.g. Support Vector Machine and Bayesian) create models that summarise the properties of the class of interests, while the *k*-nearest neighbour (KNN) algorithm preserves all the information in some way. Nevertheless, KNN is prone to noisy parameters and for that reason the Random Forest Algorithm is used first for identifying the most significant features and reducing dimensionality.

4.1. Subtract DTM from FW LiDAR

For this paper, a new feature on DASOS was added for subtracting pre-calculated DTMs generated using the Quick Terrain Modeller. The

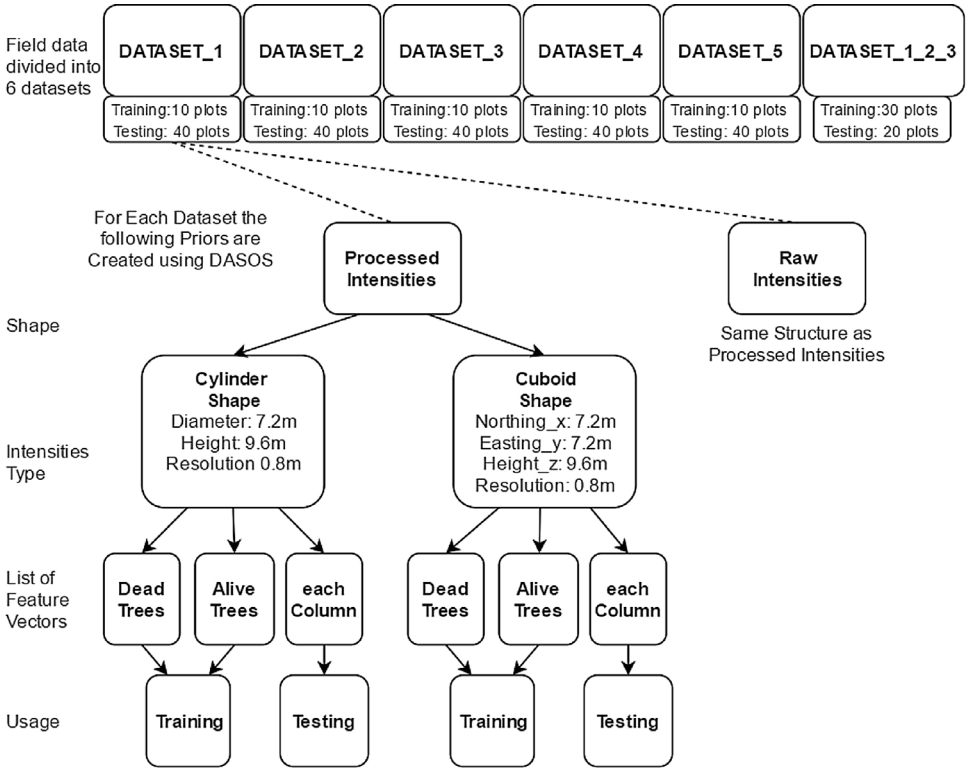


Fig. 6. Information about the feature vectors created.

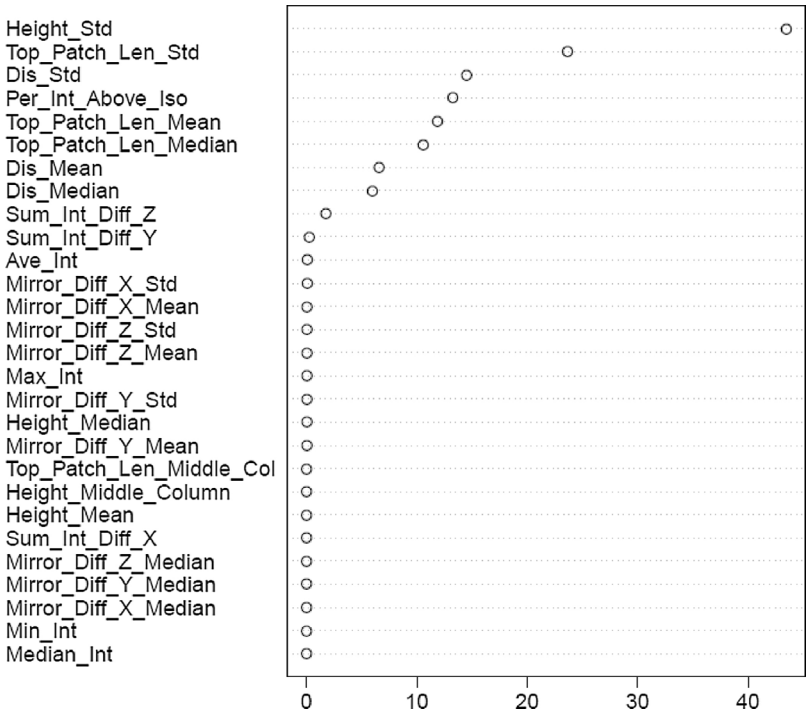


Fig. 7. Importance of variables, as identified by Random Forest.

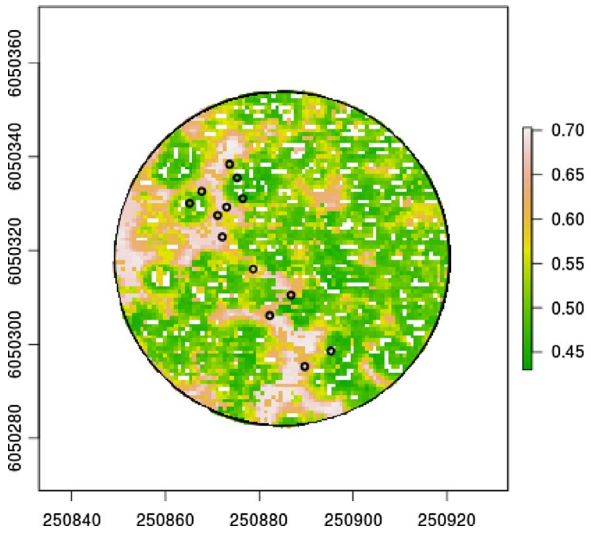


Fig. 8. The results of the K-NN algorithm.

subtraction of the DTM is done during the voxelisation; each terrain value is subtracted from the position of each waveform sample and not from the origin of the pulse since the terrain at the origin and the terrain at the position may vary. Fig. 5 shows an example of a DEM generated before and after the subtraction using DASOS.

4.2. The new feature of DASOS; feature vectors

This paper presents the new feature of DASOS, which is useful for characterising object inside the 3D space (e.g. trees). For each column of interest, inside the voxelised FW LiDAR, information from around its local area are exported as a feature vector. Multiple feature vectors are listed within.csv files for easy interpretation within statistical software packages. There are two types of exported information: processed and raw. The processed option lists information like distribution of non-empty voxels and standard deviation of heights (Table 3). The raw option lists the voxel intensity values of the local area. Additionally, there are two shapes of the local area from where the features are obtained (cuboid and cylinder). The size of each shape is user defined. Here, this new feature of DASOS is used for generating feature vectors used as a likelihood in the classifier.

Fig. 6 depicts the divisions of the datasets and related feature vectors generated. As aforementioned, there are 50 plots (including overlaps). For cross-validation, these plots were randomly divided into 5 equal training datasets. Another one was created by merging three of them to check whether the increased training samples improves classification accuracy. The feature vectors generated for each field plot are divided into two categories (processed and raw intensities) and two sub-categories (cylinder and cuboid shape), resulting into four types of feature vectors per plot. For each type, three.csv files are generated: dead trees, alive trees and columns from unknown population. The first two are used for training and the last one for testing. Each feature vector represents the area within either a cuboid or cylinder. The dimensions of those shapes are slightly smaller than the average size of the dead trees to reduce noise. Additionally, the values of the vectors are normalised to belong in the range [0, 100].

4.3. Random Forest

Random Forest generates multiple regression trees by randomly sampling the data at its nodes and chooses the best predicting variables for each sampled data. The variable importance is defined by the influence it has to the classification once this variable is modified and the rest remain unchanged (Liaw and Wiener, 2002). This paper uses the R package for finding the most relevant features exported by DASOS (Section 4.2 in identifying dead trees).

It worth highlighting that Random Forest failed to find any relation between the feature vectors with the “Raw Intensities” due to the irregular shapes of Eucalyptus camaldulensis and the variant scan angle of each field plot. Nevertheless, “Raw Intensities” may be useful in classifying trees with smaller shape variance (e.g. pine trees).

Regarding the “Processed Intensities”, Fig. 7 shows a list with the feature importance according to Random Forest and Table 3 gives the explanation of each important feature identified. The most important one is the standard deviation of heights. This is reasonable since the

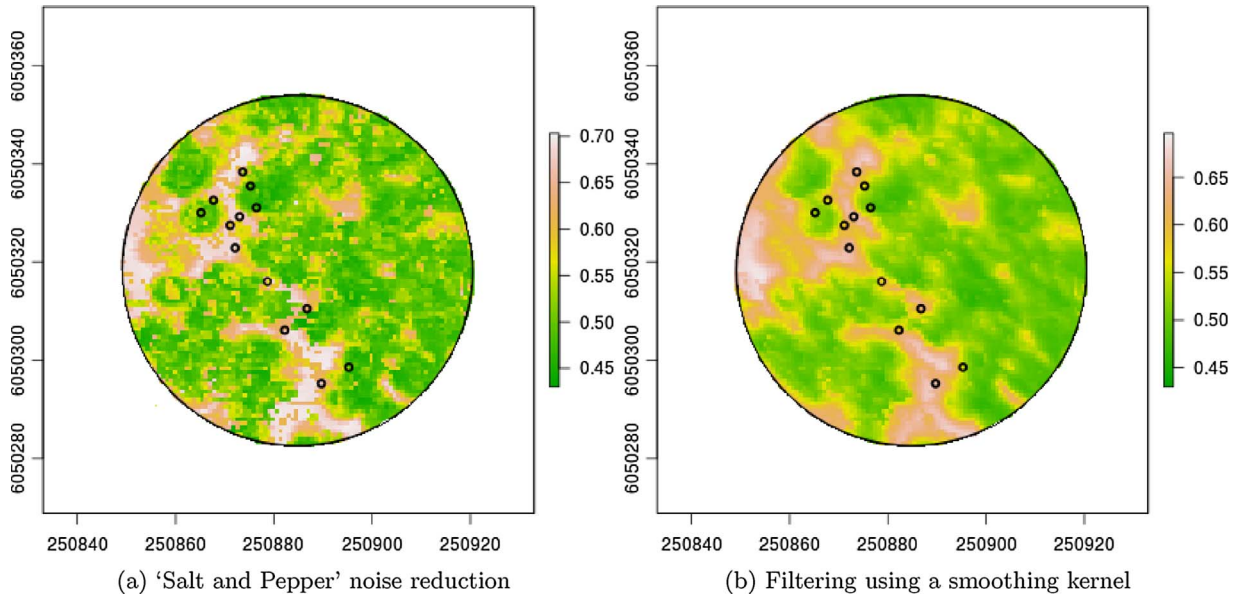


Fig. 9. Filtering the results of the K-NN algorithm.

canopy of dead trees has bigger height variance in comparison to alive trees whose canopy is leafy. Please note that the order of the significant features slightly vary according to the sub-dataset used.

4.4. Probabilistic field derived from weighted K-Nearest Neighbours Algorithm

Once the ten most significant features are identified, a weighted KNN is applied to generate a probabilistic field. As mentioned in Section 4.2, positive training feature vectors from dead trees and negative from alive trees has been created. To reduce bias, the same number of dead and alive trees is used at each training dataset.

Let's assume that T is a training dataset with n feature vectors:

$$T: (x_n, f(x_n)), \quad n = 1...N. \tag{1}$$

The function $f(x_n) \in \{0, 1\}$ return 0 for alive trees and 1 for dead trees. Therefore, the dataset T has this form:

$$T: (\boldsymbol{t}_1, 1), (\boldsymbol{t}_2, 0), (\boldsymbol{t}_3, 0), (\boldsymbol{t}_4, 1). .....(\boldsymbol{t}_n, 1) \tag{2}$$

Every feature vector $\boldsymbol{t_q} \in T$ contains the 10 most important features, as identified by Random Forest ($\boldsymbol{t} = \{t_1, t_2, \ldots, t_{10}\}$). Every feature is associated with a weight value according to its importance ($\boldsymbol{w} = \{w_1, w_2, ..., w_{10}\}$). Additionally:

$$\boldsymbol{t_q}\begin{Bmatrix} t_1 \\ t_2 \\ \cdots \\ t_{10} \end{Bmatrix} \in R^d \qquad \boldsymbol{w}\begin{Bmatrix} w_1 \\ w_2 \\ \cdots \\ w_{10} \end{Bmatrix} \in R^d \tag{3}$$

Let's define a data vector $\boldsymbol{x} = (x_1, \ldots, x_{10})$ of an unknown population. How do we calculate the probability of vector $\boldsymbol{x}$ to belong to the dead trees population? At first, the weighted Euclidean distance from $\boldsymbol{x}$ to every $\boldsymbol{t_q} \in T$ is calculated as follow:

$$d(\boldsymbol{t_q}, \boldsymbol{x}) = \sqrt{\sum_{i=1}^{10} (w_i \times (t_{\mathrm{qi}} - x_i)^2)} \tag{4}$$

Then the $k-$nearest training samples are selected. Here, $k = 7$ was considered sensible, but testing different values of k could lead to the optimal k value. The nearest 7 indices of the training samples are selected as follow:

$$q = \underset{\boldsymbol{t} \in T}{\operatorname{argmin}} d(\boldsymbol{t}, \boldsymbol{x}) \tag{5}$$

The dataset $V = \{\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_7\}$ is a subset of the training samples T and contains the k-nearest indices to $\boldsymbol{x}$, which could be positive, negative or both.

For each $\boldsymbol{v_i} \in V$ a distance-weight $\boldsymbol{u_i}$ is calculated:

$$u_i = \frac{1}{d(\boldsymbol{t_i}, \boldsymbol{i})} \tag{6}$$

Finally, the probability of being a dead tree is calculated as follow:

$$P(\mathrm{dead}) = \frac{\sum_{i=1}^{k} (u_i \times \delta(1, f(\boldsymbol{v_i})))}{\sum_{i=1}^{k} (u_i \times \delta(1, f(\boldsymbol{v_i}))) + \sum_{i=1}^{k} (u_i \times \delta(0, f(\boldsymbol{v_i})))} \tag{7}$$

where the function $\delta(a, b)$ returns 1 if a is equal to b and 0 otherwise.

For each column of the voxelised FW LiDAR, a testing data vector $\boldsymbol{x}$ is created and its probability of being dead is calculated. Fig. 8 shows the probability field of each column to belong to the dead trees population. The big circle indicates the limits of the field plot and the small circles the actual dead trees locations.

4.5. Filtering

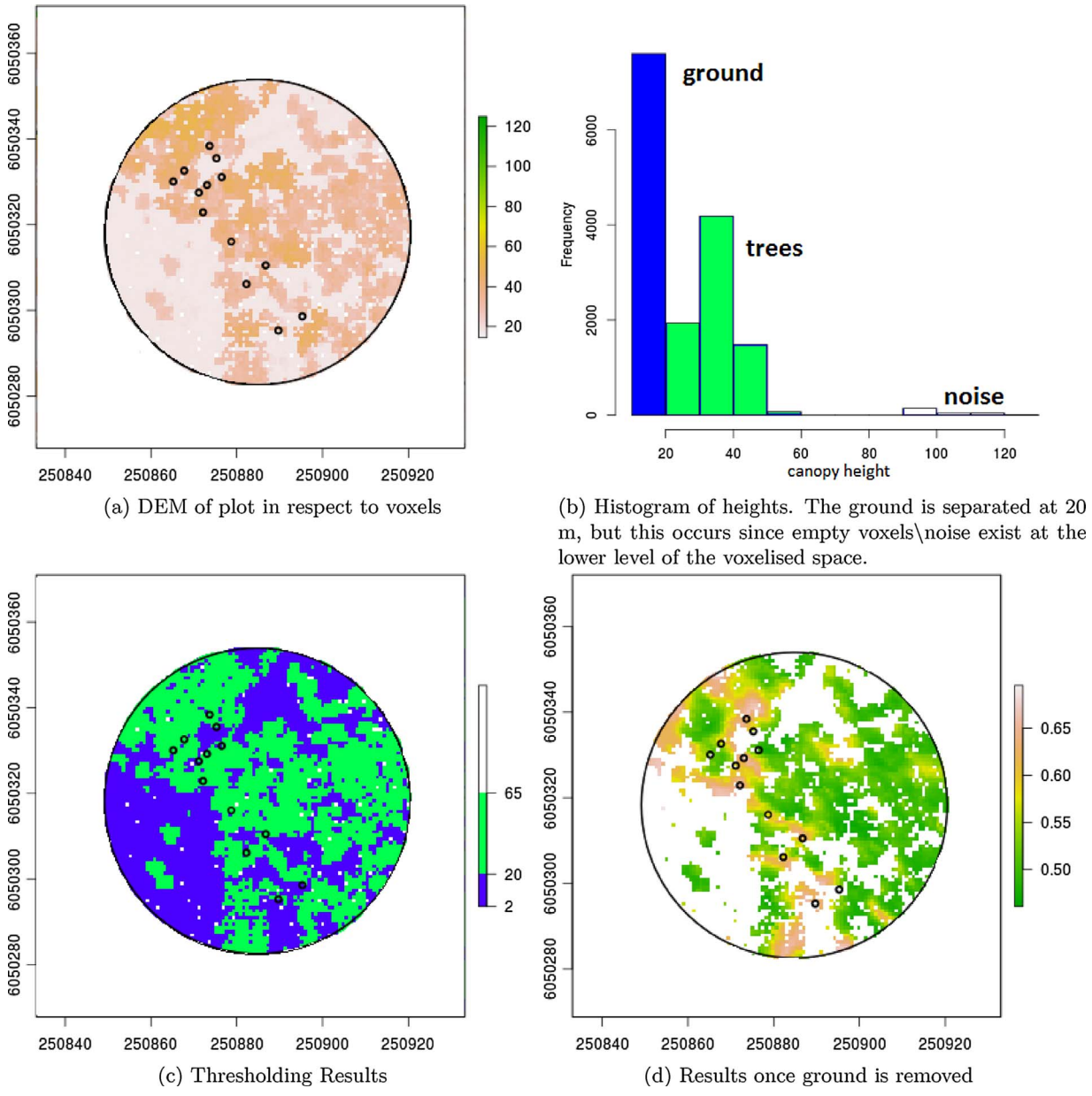
As shown in Fig. 8, there is 'Salt and Pepper' noise. This occurs when no pulses pass through the corresponding column of the voxelised space or when the height of the shape used to calculate the corresponding feature vector is bigger than the canopy height (common at ground). It is removed using a median 3×3 filter which assigns to every empty pixel the median value of its non-empty neighbouring pixels (Fig. 9a). A smoothing filter is afterwards applied for further noise reduction (Fig. 9b).

4.6. Removing ground pixels

Removing the ground pixels is a trivial task because the DTM has already subtracted. A histogram of the heights, defined by the top non-empty voxels times voxel size, was generated. There are three well-defined classes: ground, trees and noise (Fig. 10b). Ground and noise are removed and this processed is illustrated in Fig. 10.

4.7. Dead and alive thresholding, filtering, segmentation and position assignment

Fig. 11a show how the the dead trees are thresholded from the alive



(a) DEM of plot in respect to voxels

(b) Histogram of heights. The ground is separated at 20 m, but this occurs since empty voxels\noise exist at the lower level of the voxelised space.

(c) Thresholding Results

(d) Results once ground is removed

Fig. 10. The process of removing the ground pixels.

trees. To reduce over-detection of dead trees, the outliers are filtered out using a 3 × 3 median kernel (Fig. 11b). Afterwards, the pixels are grouped using a seed growth algorithm (Algorithm 1, Fig. 12a). Each group corresponds to a predicted dead tree. Its predicted location is the average geo-spatial location of the pixels that belong to that tree (Fig. 12b).

Algorithm 1. Seed growth algorithm for grouping nearby pixels of dead trees

1: $P \longleftarrow$ all pixels classified as dead
2: $s \longleftarrow 0$
3: **while** not reached the end of set P **do**
4: get next pixel $\boldsymbol{p} \in P$ that is not assigned to a segment
5: assign pixel $\boldsymbol{p}$ to segment s
6: find $K(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n)$ such that $K \subseteq P$ and every $\boldsymbol{p}_i \in K$ is a neighbour of $\boldsymbol{p}$
7: $\forall\, \boldsymbol{p}_i \in K$, $\boldsymbol{p} \longleftarrow \boldsymbol{p}_i$ and repeat from line 5
8: all pixels of segment s has been labelled
9: $s \longleftarrow s + 1$

5. Evaluation

The results are cross-validated according to the predicted locations of the dead trees and their distance from the actual dead trees. Three different test were undertaken during evaluation: does the increased training samples improve dead tree detection? What shape (cylinder or cuboid) performs better? Are the predictions better than a random prediction? The random prediction was generated by uniformly distributing locations with density equal to the average density of the

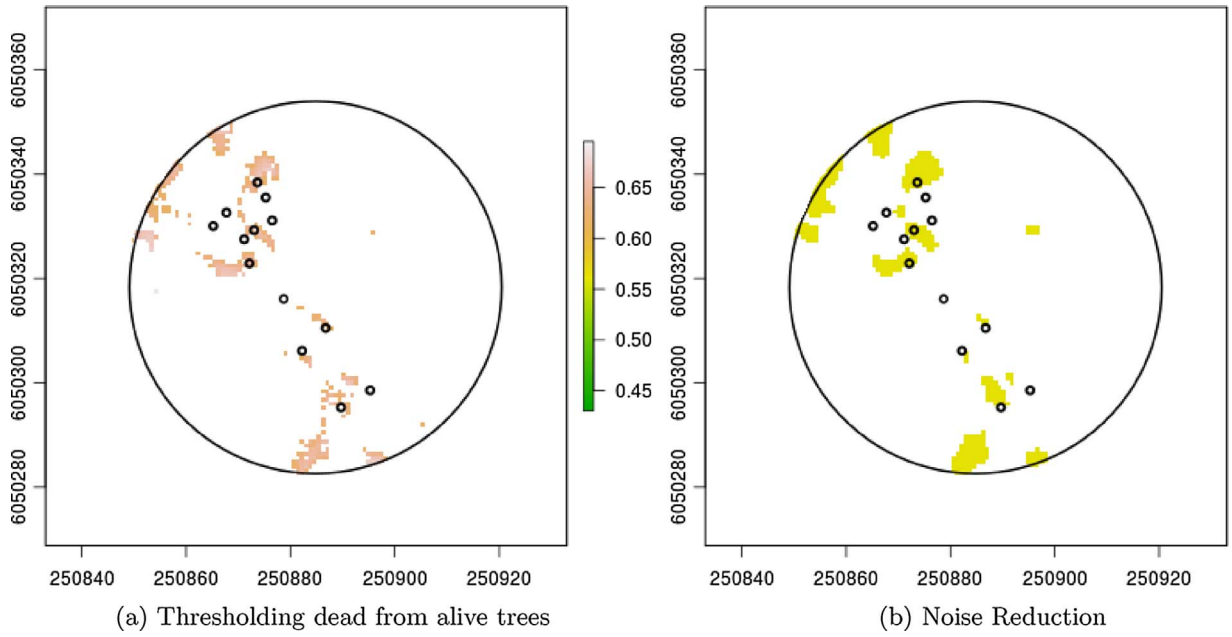


Fig. 11. Thresholding and filtering.

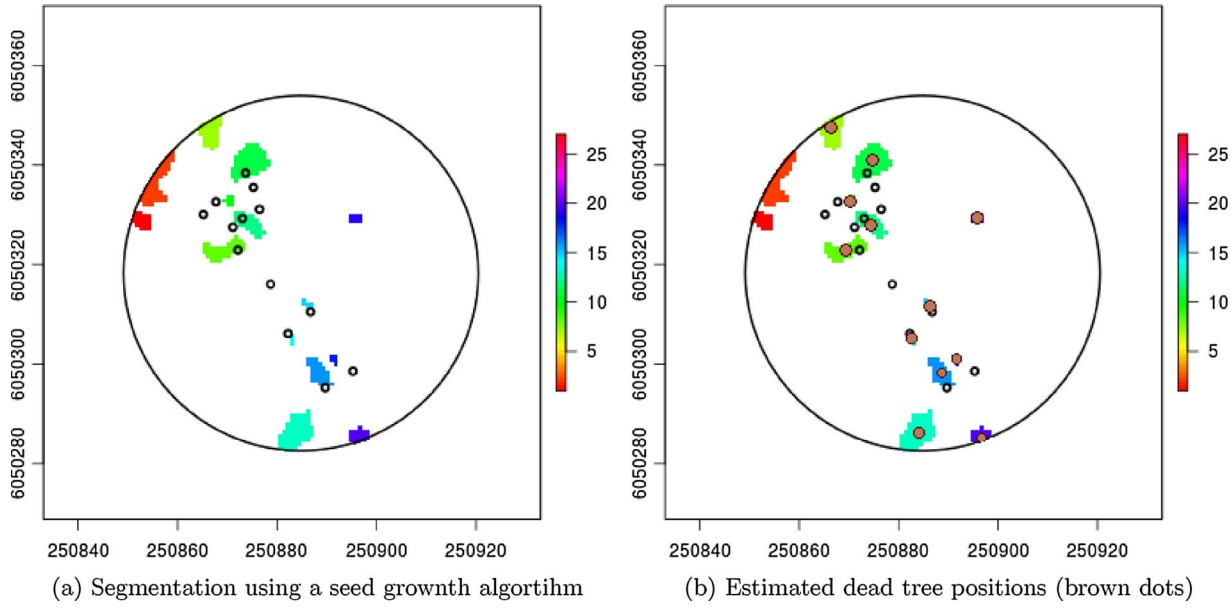


Fig. 12. Segmentation and dead trees' position assignment.

Table 4
Precision achieved using the cylindrical shape to extract features.

Distance	1 m	2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m
D1	7.29	12.15	16.1	24.31	32.21	38.9	47.11	49.84	56.23	58.35
D2	2	3.67	8.36	18.39	25.08	33.11	35.78	40.13	46.48	50.5
D3	1.48	5.46	14.2	23.32	29.08	36.6	40.23	46.15	51.38	56.28
D4	0.96	7.24	20.04	28.26	33.09	40.09	44.68	52.17	56.28	62.07
D5	0.75	5.26	8.27	12.03	14.28	21.8	28.57	39.84	47.36	55.63
D_1_2_3	0	8.69	13.04	19.13	24.34	29.56	34.78	36.52	41.73	55.65

Table 5
Recall achieved using the cylindrical shape to extract features.

Distance	1 m	2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m
D1	2.55	9.26	18.84	32.26	45.04	53.35	56.23	58.78	63.25	68.05
D2	8.22	13.48	23.02	31.25	38.48	51.31	62.17	65.78	66.11	69.07
D3	6.69	14.49	26.55	34.77	40.97	48.6	56.61	59.18	60.71	63.41
D4	5.16	15.5	30.09	38.29	43.46	45.89	51.06	52.58	55.31	57.75
D5	0.89	4.45	12.75	24.03	29.37	35.9	41.83	49.85	59.34	61.12
D_1_2_3	0	7.22	14.45	20.07	45.19	50.6	51.8	62.65	63.85	73.49

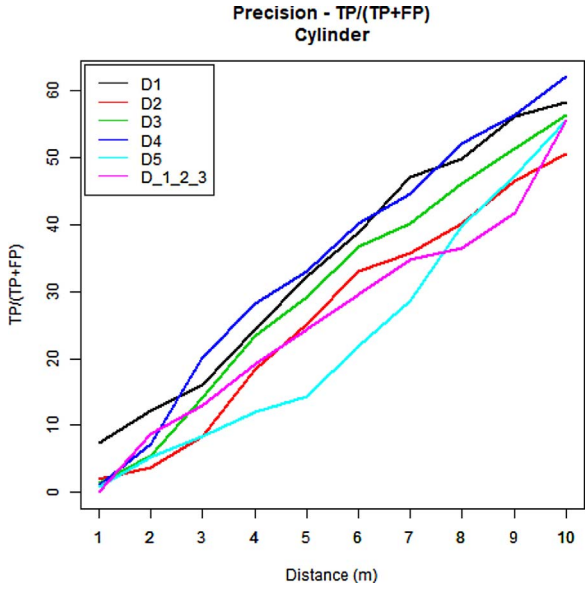


Fig. 13. Precision obtained using a cylindrical shape to extract features. Dataset D_1_2_3 contains more training samples than the rest.

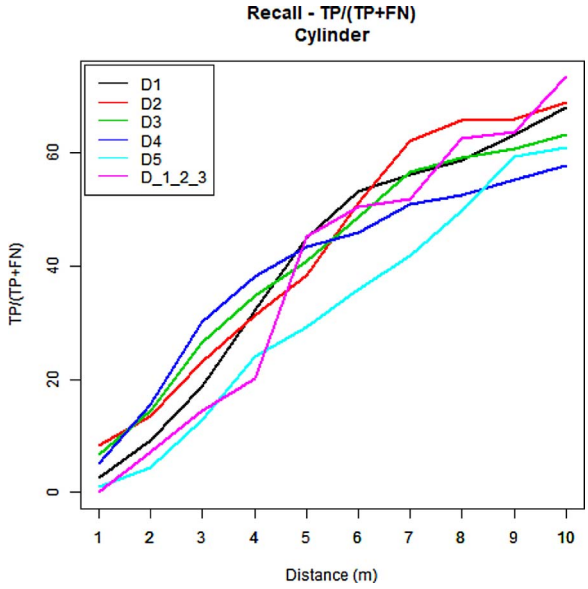


Fig. 14. Recall obtained using a cylindrical shape to extract features. Dataset D_1_2_3 contains more training samples than the rest.

actual dead trees.

Tables 4 and 5 shows the precision and recall percentage achieved using a cylindrical shape to extract features for the likelihood. The D1, D2, …, D5 corresponds to the five divided datasets in the cross-validation. The D_1_2_3 uses all the training samples from D1, D2 and D3 to train the classifier. As shown in the corresponding charts of precision (Fig. 13 and recall Figure 14), the increased amount of training samples do not improve the prediction. The bigger the training set is, the shorter the distances to the k nearest neighbours should be. But the field data contain noise (Section 3), which compensates the value of the increased samples. A less noisy training dataset with trees of similar heights should improve the results.

The second comparison between using a cuboid and cylindrical shape to extract features. Tables 4 and 5 and Figs. 13 and 14 show the results of the prediction using a cylindrical shape to extract features, while Tables 6 and 7 and Figs. 15 and 16 show the results obtained using a cuboid shape. The average results are similar but the cuboid shape has a wider range of predicted results. Trees do not have corners and therefore the information retrieved with the cuboid shape are less meaningful. Nevertheless, the cuboid shape is slightly bigger and therefore collects better information from big trees but more noise from small trees. This justifies the wider range of good/bad results.

Finally, there is a comparison between the average results obtained and the evaluation of the random prediction are shown in Tables 8 and 9 and Figs. 17 and 18. During the evaluation a predicted tree is classified as TP if there is a dead tree within the given distance. As the minimum distance to a dead tree increases, the number of FP reduces. The precision of the random classifier significantly increases after the 6m distance from an actual dead tree, because its distribution is uniform. Additionally, the dimensions of the shapes used are 7.2 m diameter and width for cylinder and cuboid respectively. Therefore, evaluation above this distance is meaningless.

From the aforementioned figures, it is shown that the methodology proposed performs better than random. This indicates that forest health assessment is possible without tree delineation. Additionally, field surveying could be improved by better planning using the results of this systems. Of course, this is a new research direction and there are many improvements to be made (e.g. extracting features resistant to tree height variations).

6. Conclusions and future work

The importance of dead wood in our ecosystem is large and it is monitored for managing biodiversity. The study area of these projects is a native Australian forest with Eucalyptus camaldulensis, where shortage of hollows available for colonisation is predicted. Dead trees are more likely to be aged and contain hollows and therefore detecting them is essential for protecting their inhabitants.

This paper proposed a new direction for detecting dead standing Eucalyptus camaldulensis and presents the new feature of DASOS (extraction of feature vectors). Previous work on forest health assessment uses tree delineation but this leads to over-segmentation when applied to Eucalyptus camaldulensis due to their irregular shapes and multiple trunk splits. Additionally the density of the acquired LiDAR makes bottom to top delineation impossible since information about the trunks are missing. Therefore, this paper investigates the possibility of detecting dead trees from voxelised FW LiDAR without tree delineation.

Field data were provided by Forestry Corporation of NSW, Australia and Interpine Group Ltd, New Zealand. The GPS positions of the dead and alive trees within the plots are given but these data contain noise. Additionally, some small trees listed within the field data are not

Table 6
Precision achieved using the Cuboid shape to extract features.

Distance	1 m	2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m
D1	7.78	10.47	14.07	24.85	32.63	40.11	47.9	53.59	59.28	61.97
D2	3.66	12.5	16.37	22.84	26.72	34.26	42.02	46.76	51.72	56.68
D3	1.24	3.42	20.56	25.23	29.28	36.76	41.74	43.92	50.15	53.27
D4	8.36	19.86	33.79	36.58	39.02	44.25	50.52	55.05	63.76	66.89
D5	1.96	1.96	5.88	15.68	19.6	25.49	35.29	41.17	45.09	62.74

Table 7
Recall achieved using the Cuboid shape to extract features.

Distance	1 m	2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m
D1	9.58	19.16	35.14	44.4	50.47	56.86	62.93	65.49	69.96	74.76
D2	10.52	20.39	33.22	37.82	47.36	62.17	67.1	70.39	74.01	75.65
D3	5.48	20	31.93	38.7	46.12	54.83	60.64	67.09	72.9	77.09
D4	7.9	17.93	24.92	26.74	33.13	38.29	45.28	47.41	50.75	52.27
D5	4.74	4.74	5.34	11.86	16.02	16.32	21.95	23.73	27.29	31.75

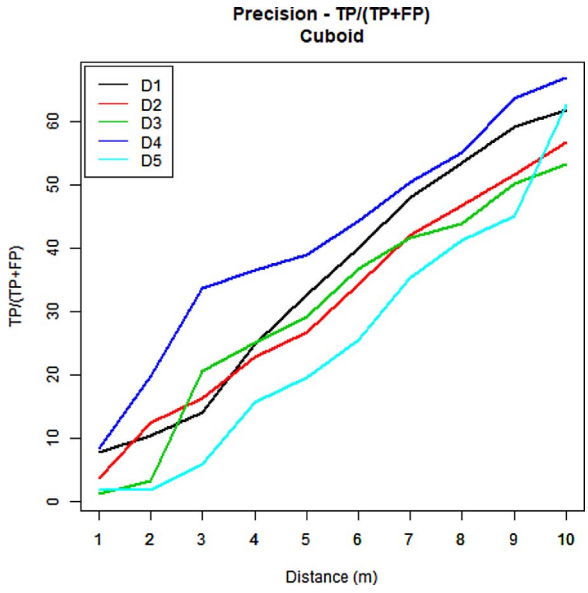


Fig. 15. Precision results obtained using a cuboid shape to extract features.

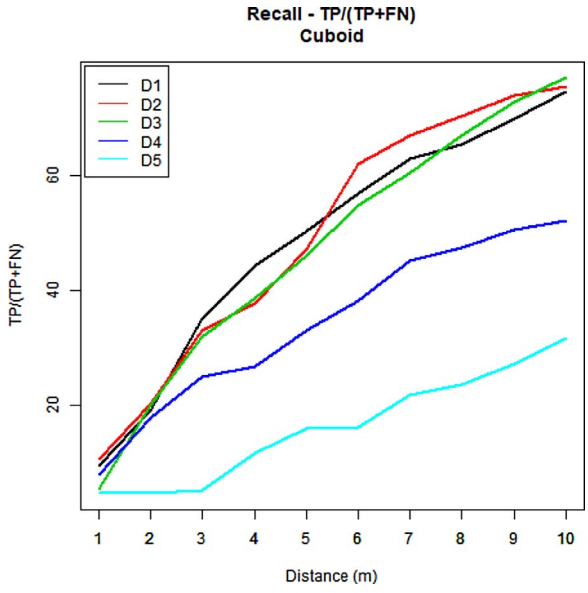


Fig. 16. Recall obtained using a cuboid shape within to extract features.

detectable from the FW LiDAR.

Regarding the methodology, feature vectors characterising dead and alive trees are generated. Then those feature vectors are run over the volume to generate a 2D image for each testing plot with the probability of a pixel to be from a dead tree or not. Salt and pepper noise removal and smoothing filters are applied. The ground is afterwards removed and a threshold is defined to separate pixels containing dead and alive trees. Then extra filtering is applied and a seed growth algorithm is used to label each segment. Each segment corresponds to a dead tree prediction and its estimated location is the average position of the pixels that belong to the segment.

The results have been cross-validated. Overall, there are three outcomes. The increase amount of training samples does not improve the results of the classification because of the noise within the field data. The feature vectors derived from cylindrical shape are more reliable because the range of the recall and precision is smaller. By the end, the most important outcome is that the results was clearly better than the random prediction, justifying that it is possible to identify dead trees without tree delineation.

Nevertheless, this is the first attempt to assess health forest without tree delineation and therefore many improvements could be made. Fore example:

- Check and improve accuracy of the field data using visualisations of the FW LiDAR.
- Adjust features extracted so that they are not height dependant.
- Categorise the trees according to their size and derive a sets of training feature vectors according to their height.
- Once the seed growth algorithm is applied, check the size and shape of the segments and the possibility of dividing big segments into multiple dead trees.
- Add negative samples from the ground and the edges of alive trees, because the system is confused at the edges of the alive trees.

Acknowledgements

Special thanks are given to the Centre for Digital Entertainment and Plymouth Marine Laboratory, who supported this project financially and consequently made it possible. Additional credits are given to Forestry Corporation of NSW for the project idea and Interpine Group Ltd, who hosted the main author for three months and provided useful knowledge about forestry.

Table 8
Average precision of each shape (Cylinder and Cuboid) and the Random prediction.

Distance	1 m	2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m
Cylinder	2.50	6.76	13.40	21.27	26.75	34.10	39.28	45.63	51.55	56.57
Cuboid	4.60	9.645	18.14	25.04	29.45	36.18	43.50	48.10	54.00	60.31
Random	0	0	2.56	8.06	11.36	23.81	52.75	57.14	72.16	79.49

Table 9
Average recall for each shape (Cylinder and Cuboid) and the Random prediction.

Distance	1 m	2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m
Cylinder	4.71	11.44	22.26	32.13	39.47	47.02	53.58	57.24	60.95	63.88
Cuboid	7.65	16.45	26.11	31.91	38.63	45.70	51.59	54.83	59.00	62.31
Random	0	0	6.18	7.34	8.88	10.04	12.74	20.85	21.62	28.59

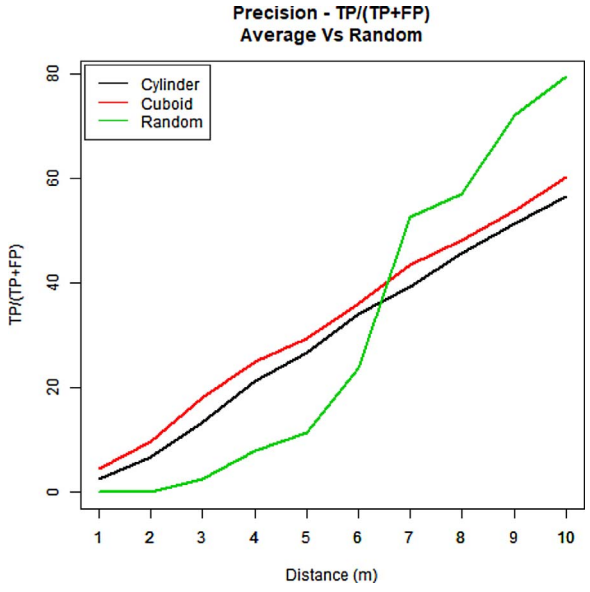


Fig. 17. Average recall for each shape (Cylinder and Cuboid) and the random prediction.

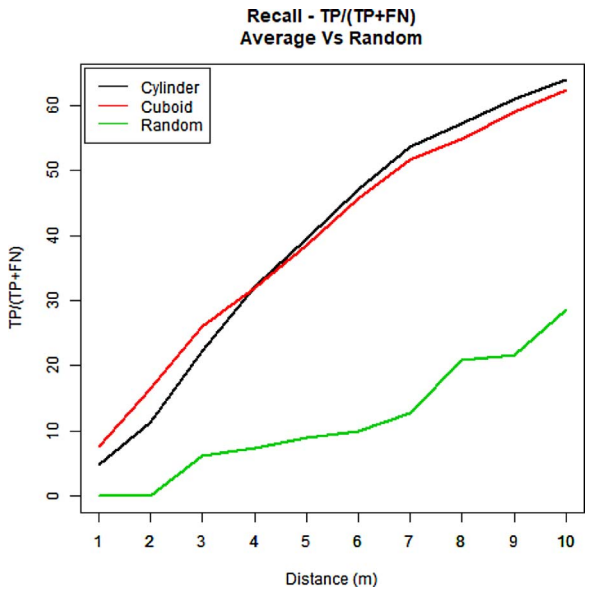


Fig. 18. Average precision for each shape (Cylinder and Cuboid) and the Random prediction.

References

Bennett, A., Lumsden, L., Nicholls, A., 1994. Tree hollows as a resource for wildlife in remnant woodlands: spatial and temporal patterns across the northern plains of Victoria, Australia. Pac. Conserv. Biol. 1 (3), 222–235.

Cao, L., Coops, N., Innes, L., Dai, J., Ruan, H., 2016. Tree species classification in subtropical forests using small-footprint full-waveform LiDAR data. Int. J. Appl. Earth Observ. Geoinf. 49, 39–51. http://dx.doi.org/10.1016/j.jag.2016.01.007.

Franklin, J.F., Shugart, H.H., Harmon, M.E., 1987. Tree death as an ecological process. BioScience 17 (8), 550–556. http://dx.doi.org/10.2307/1310665.

Gibbons, P., Lindenmayer, D., 2002. Tree Hollows and Wildlife Conservation in Australia. CSIRO Publishinghttp://dx.doi.org/10.5860/choice.40-1547.

Gibbons, P., Lindenmayer, D., Barry, S.C., Tanton, M., 2000. Hollow formation in eucalypts from temperate forests in Southeastern Australia. Pac. Conserv. Biol. 6 (3), 218.

Goldingay, R.L., 2009. Characteristics of tree hollows used by Australian birds and bats. Wildl. Res. 36 (5), 394–409. http://dx.doi.org/10.1071/WR08172.

Hancock, S., Anderson, K., Disney, M., Gaston, K.J., 2017. Measurement of fine-spatial-resolution 3D vegetation structure with airborne waveform LiDAR: calibration and validation with voxelised terrestrial LiDAR. Remote Sens. Environ. 188, 37–50. http://dx.doi.org/10.1016/j.rse.2016.10.041.

Hanski, I., 2000. Extinction debt and species credit in boreal forests: modelling the consequences of different approaches to biodiversity conservation. Ann. Zool. Fennici 271–280.

Hu, B., Li, J., Jing, L., Judah, A., 2014. Improving the efficiency and accuracy of individual tree crown delineation from high-density LiDAR data. Int. J. Appl. Earth Observ. Geoinf. 26, 145–215. http://dx.doi.org/10.1016/j.jag.2013.06.003.

Jing, L., Hu, B., Li, J., Noland, T., 2012. Automated delineation of individual tree crowns from LiDAR data by multi-scale analysis and segmentation. Photogramm. Eng. Remote Sens. 78 (12), 1275–1284. http://dx.doi.org/10.14358/PERS.78.11.1275.

Kerle, J., 2005. Collation and Review of Stem Density Data and Thinning Prescriptions for the Vegetation Communities of New South Wales.

Kim, Y., Yang, Z., Cohen, W.B., Pflugmacher, D., Lauver, C.L., Vankat, J.L., 2009. Distinguishing between live and dead standing tree biomass on the north rim of grand canyon national park, USA using small-footprint LiDAR data. Remote Sens. Environ. 113 (11), 2499–2510. http://dx.doi.org/10.1016/j.rse.2009.07.010.

Liaw, A., Wiener, M., 2002. Classification and regression by randomforest. R News 2 (3), 18–22.

Lindenmayer, D.B., Cunningham, R.B., Donnelly, C.F., 1997. Decay and collapse of trees with hollows in Eastern Australian forests: impacts on arboreal marsupials. Ecol. Appl. 7 (2), 625–641. http://dx.doi.org/10.2307/2269526.

Lindenmayer, D.B., Wood, J.T., 2010. Long-term patterns in the decay, collapse, and abundance of trees with hollows in the mountain ash (eucalyptus regnans) forests of Victoria, Southeastern Australia. Can. J. For. Res. 40 (1), 48–54. http://dx.doi.org/10.1139/X09-185.

Lovell, J.L., Jupp, D.L.B., Newnham, G.J., Coops, N.C., Culvenor, D.S., 2005. Simulation study for finding optimal LiDAR acquisition parameters for forest height retrieval. For. Ecol. Manag. 214 (1), 398–412. http://dx.doi.org/10.1016/j.foreco.2004.07.077.

Mcke, W., Dek, B., Schroiff, A., H.M, Pfeifer, N., 2013. Detection of fallen trees in forested areas using small footprint airborne laser scanning data. Can. J. Remote Sens. 139 (s1), S32–S40. http://dx.doi.org/10.5589/m13-013.

Miltiadou, M., Warren, M.A., Grant, M., Brown, M., 2015. Alignment of hyperspectral imagery and full-waveform LiDAR data for visualisation and classification purposes. Int. Arch. Photogramm. Rem. Sens. Spatial Inf. Sci. 40 (7), 1257. http://dx.doi.org/10.5194/isprsarchives-XL-7-W3-1257-2015.

Neuenschwander, A., Magruder, L., Tyler, M., 2009. Landcover classification of small-footprint full-waveform LiDAR data. J. Appl. Remote Sens. 3 (1), 033544. http://dx.doi.org/10.1117/1.3229944.

New South Wales Government, 2016. Biodiversity Conservation Act 2016 No 63, Technical report.

Pasher, J., King, D.J., 2009. Mapping dead wood distribution in a temperate hardwood forest using high resolution airborne imagery. For. Ecol. Manag. 258 (7), 1536–1548.

http://dx.doi.org/10.1016/j.foreco.2009.07.009.
Persson, A., Soderman, U., Topel, J., Ahlberg, S., 2005. Visualisation and analysis of full-waveform airborne laser scanner data. V/3 Workshop, Laser scanning 2005. http://dx.doi.org/10.1117/12.604655.
Peterson, G., Allen, C.R., Holling, C.S., 1998. Ecological resilience, biodiversity, and scale. Ecosystems 1 (1), 6–18. http://dx.doi.org/10.1007/s100219900002.
Polewski, P., Yao, W., Heurich, M., Krzystek, P., Stilla, U., 2015. Detection of fallen trees in als point clouds using a normalized cut approach trained by simulation. ISPRS J. Photogramm. Remote Sens. 105, 252–271. http://dx.doi.org/10.1016/j.isprsjprs.2015.01.010.
Popescu, S.C., Wynne, R.H., Nelson, R.F., 2003. Measuring individual tree crown diameter with LiDAR and assessing its influence on estimating forest volume and biomass. Can. J. Remote Sens. 29 (5), 564–577. http://dx.doi.org/10.5589/m03-027.
Popescu, S.C., Zhao, K., 2008. A voxel-based LiDAR method for estimating crown base height for deciduous and pine trees. Remote Sens. Environ. 112 (3), 767–781. http://dx.doi.org/10.1016/j.rse.2007.06.011.
Reitberger, J., Krzystek, P., Stilla, U., 2008. Analysis of full waveform LiDAR data for tree species classification. Int. J. Remote Sens. 29 (5), 1407–1431.
Shendryk, I., Broich, M., Tulbure, M.G., Alexandrov, S.V., 2016a. Bottom-up delineation of individual trees from full-waveform airborne laser scans in a structurally complex eucalypt forest. Remote Sens. Environ. 173, 69–83. http://dx.doi.org/10.1016/j.rse.2015.11.008s.
Shendryk, I., Broich, M., Tulbure, M.G., McGrath, A., Keith, D., Alexandrov, S.V., 2016b. Mapping individual tree health using full-waveform airborne laser scans and imaging spectroscopy: a case study for a floodplain eucalypt forest. Remote Sens. Environ. 187, 202–217. http://dx.doi.org/10.1016/j.rse.2016.10.014.
Siitonen, J., 2001. Forest management, coarse woody debris and saproxylic organisms: Fennoscandian boreal forests as an example. Ecol. Bull. 11–41.
Wilson, N., of N.S.W, N. C. C, 1995. The Flooded Gum Trees: Land Use and Management of River Red Gums in New South Wales. The Council, Sydney. http://dx.doi.org/10.1046/j.1442-9993.2000.01036.x.
Wormington, K., Lamb, D., 1999. Tree hollow development in wet and dry sclerophyll eucalypt forest in South-East Queensland, Australia. Aust. Forestry 62 (4), 336–345. http://dx.doi.org/10.1080/00049158.1999.10674801.
Yao, W., Krzystek, P., Heurich, M., 2012. Identifying standing dead trees in forest areas based on 3D single tree detection from full-waveform LiDAR data. ISPRS Ann. Photogramm. Rem. Sens. Spatial Inf. Sci. I-7, 359–364. http://dx.doi.org/10.5194/isprsannals-i-7-359-2012.