

Novel algorithms for efficiently
accumulating, analysing and
visualising full-waveform LiDAR in
a volumetric representation with
applications to forestry

submitted by

Milto Miltiadou

for the degree of Doctor of Engineering

of the

University of Bath

Centre for Digital Entertainment

and of the

Plymouth Marine Laboratory

NERC Airborne Research Facility

April 2017

COPYRIGHT

Attention is drawn to the fact that copyright of this thesis rests with its author. This copy of the thesis has been supplied on the condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the prior written consent of the author.

This thesis may be made available for consultation within the University Library and may be photocopied or lent to other libraries for the purposes of consultation.

Signature of Author.....

Milto Miltiadou

Abstract

no more than 300 words

NOTES:

Blue colour: additions according to Neill's feedback,

Purple colour: addition/corrections according to Mike's comments

Red colour: notes

Gray colour: text that is going to be modified

To be added on top

Abstract

This study focuses on enhancing the visualisations and classifications of forested areas using coincident full-waveform (fw) LiDAR data and hyperspectral images. The ultimate aim is use both datasets to derive information about forests and show the results on a 3D virtual, interactive environment. Influenced by Persson et al (2005), voxelisation is an integral part of this research. The intensity profile of each full-waveform pulse is accumulated into a voxel array, building up a 3D density volume. The correlation between multiple pulses into a voxel representation produces a more accurate representation, which confers greater noise resistance and it further opens up possibilities of vertical interpretation of the data. The 3D density volume is then aligned with the hyperspectral images using a 2D grid similar to Warren et al (2014) and both datasets are used in visualisations and classifications.

Previous work in visualising fw LiDAR has used transparent objects and point clouds, while the output of this system is a coloured 3D-polygon representation, showing well-separated structures such as individual trees and greenhouses. The 3D density volume, generated from the fw LiDAR data, is polygonised using functional representation of object (FReps) and the marching cubes algorithm (Pasko and Savchenko, 1994) (Lorensen and Cline, 1987). Further, an optimisation algorithm is introduced that uses integral volumes (Crow, 1984) to speed up the process of polygonising the volume. This optimisation approach not only works on non-manifold object, but also a speed up of up to 51% was achieved. The polygon representation is also textured by projecting the hyperspectral images into the mesh. In addition, the output is suitable for direct rendering with commodity 3D-accelerated hardware, allowing smooth visualisation.

In future work, the effects of combining both hyperspectral imagery and fw LiDAR in classifications and visualisations are examined. At first, two pixel wise classifiers, a support vector machine and a Bayesian probabilistic model, will be used for testing the effects of the combination in generating tree coverage maps. Higher accuracy classification results are expected when metrics from both datasets are used together. Regarding the visualisations, the differences of applying surface reconstruction versus direct volumetric rendering will be discussed and an ordered tree structure with integral sums of the node values will be used for speeding up the ray-tracing of direct volumetric rendering and improving memory management of aforementioned optimisation algorithm with integral volumes. Further, deferred rendering is suggested for testing the visual human perception of projecting multiple bands of the hyperspectral images on the FW LiDAR

polygon representations. At the end of this project the combination of the datasets will be used along with the watershed algorithm for tree segmentation, which is useful for measuring the stem density of a forest and for tree species classifications.

from EDE:

Firstly, a new and fast way of aligning the FW LiDAR with Remotely Sensed Images has been developed in DASOS and by generating tree coverage maps it was shown that the combination of those datasets confers better remote survey results. This work was presented at the 36th ISRSE International Conference.

Secondly, automated detection of dead trees in native Australian forests has a significant role in protecting animals, which live in those trees and are close to extinction. DASOS allow the generation of 3D signatures characterising dead trees. A comparison between the discrete and FW LiDAR is performed to demonstrate the increased survey accuracy obtained when the FW LiDAR are used.

Finally, the last application is for improving visualisations for foresters. Foresters have a great knowledge about forests and can derive a wealth of information directly from visualisations of the remotely sensed data. This reduces the travelling time and cost of getting into the forests. This research optimises visualisations by using the new FW LiDAR representations and a speed of up to 51% has been achieved.

FW LiDAR has great potentials in forestry and this research has already started to have an impact in the FW LiDAR community by making those huge datasets easier to handle. DASOS is now used at Interpine Group Ltd, a world leading Forestry Company in New Zealand and it has been tested from a PhD student at Bournemouth University who looks into estimating bird distribution in the New Forest. In the future, it is expected that DASOS will be widely used in remote forest surveys (i.e. estimating the commercial value of a forest and detecting infected trees at early stages for treatment).

Acknowledgements

Above all, I would like to express my great gratitude to my industrial supervisors Dr. Michael Grant who had supported me continuously during my research and gave me the freedom to create a project of my own interest.

Then, I would like to thanks Dr. Matthew Brown, who helped me during my first years of my studies by giving me valuable and informative feedback. He was always there to keep me working on the right track.

Equally important is my current supervisor Dr. Neil D.F. Campbell and he is not to be missed from the acknowledgements.

Furthermore, special thanks are given to Dr. Mark Warren, Dr. Darren Cosker, MSc Susana Gonzalez Aracil and Dr. Ross Hill who occasionally advised me during my studies.

It further worth giving credits to my data providers, the Natural Environment Research Council's Airborne Research Facility (NERC ARF) and Interpine Group Ltd.

Last but not least, I am extremely grateful to my funding organisations, the Centre for Digital Entertainment and Plymouth Marine Laboratory, who supported financially and consequently made this research possible.

Abbreviations and Glossary

AGC	Automatic Gain Controller
ALS	Airborne Laser Scanning
APL	Airborne Processing Library
ARF	Airborne Research Facility
CG	Computer Graphics
CHM	Canopy Height Model
CUDA	parallel computing platform available on nvidia graphic cards
DASOS	(δασος=forest in Greek), the open source software implemented for managing FW LiDAR data
DBH	Diameter at Breast Height
DEM	Digital Elevation Model
DTM	Digital Terrain Model (DTM)
FN	False Negative
FP	False Positive
FW	Full-Waveform
GB	Gigabyte
K-NN	K-Nearest Neighbour
LiDAR	Light Detection And Ranging
MRI	Magnetic Resonance Imaging
NASA	National Aeronautics and Space Administration
NDVI	Normalised Difference Vegetation Index
NERC	Natural Environment Research Council
NIR	Near-Infrared Region of the electromagnetic spectrum
QGIS	Quantum Geographic Information System
SIMD	Single Instruction, Multiple Data
TB	Terabyte
TP	True Positive
TN	True Negative
VIS	Visual Spectrum
VLR	Variable Length Records
WPDF	Waveform Packet Descriptor Format
UK	United Kingdom

Publications

DASOS-User Guide, M. Miltiadou, N.D.F Campbell, M. Brown, S.C. Aracil, M.A. Warren, D. Clewley, D.Cosker, and M. Grant, Full-waveform LiDAR workshop at Interpine Group Ltd, Rotorua NZ, 2016

Improving and Optimising Visualisations of full-waveform LiDAR data, M. Miltiadou, M. Brown, N.D.F Campbell, D. Cosker, M. Grant, *EuroGraphics UK, Computer Graphics & Visual Computing*, 2016

University of Bath Alignment of Hyperspectral Imagery and Full-Waveform LiDAR data for visualisation and classification purposes, M. Miltiadou, M. A. Warren, M. Grant, and M. Brown, *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, no. 7, p. 1257, 2015.

Reconstruction of a 3D Polygon Representation from Full-Wavefrom LiDAR data, M. Miltiadou, M. Grant, M. Brown, M. Warren, and E. Carolan,*RSPSoc Annual Conference, New Sensors for a Changing World*, 2014.

Awards

EDE and Ravenscroft Prize - Finalist: Selected as one of the five finalists for this is a prestigious prize that recognises the work of best postgraduate researchers.

Student Poster Competition at Silvilaser.

Conference Presentations

Remote Sensing Cyprus (RSCy) Conference, 2017 , Paphos, Cyprus - Oral Presentation

ForestSAT Conference,2016 , Santiago, Chile - Oral Presentation

Computer Graphics & Visual Computing (CGVC),2016, Bournemouth, United Kingdom - Poster Presentation

Silvilaser, 2015, La Grant Motte, France - Oral Presentation

International Symposium of Remote Sensing of the Environment (ISRSE), 2015, Berlin, German - Oral Presentation

Remote Sensing and Photogrammetry Society (RSPSoc) Conference, New Sensors for a Changing world , 2014, Aberystwyth, United Kingdom - Oral Presentation

Workshops

Full day workshop about FW LiDAR and DASOS at *Interpine Ltd Group*, 2016,
Rotorua, New Zealand

Demonstration of DASOS_v2 at the practical LiDAR session at *the NERC ARF annual workshop*, 2017, Plymouth, United Kingdom

Contents

Abstract	i
Acknowledgements	iii
Abbreviations and Glossary	iv
Publications	v
Awards	v
Conference Presentations	v
Workshops	vi
List of Figures	ix
1 Introduction	1
1.1 Forest Monitoring: Importance and Applications	1
1.2 Background Information about Remote Sensing and Airborne Laser Scanning Systems	1
2 Acquire Data	2
3 Overview of Thesis	3
3.1 Problem	3
3.2 Aims and Objectives	4
3.3 Overview	5
3.4 Thesis Structure	6
4 The open source software DASOS and the Voxelisation Approach	7
5 Surface Reconstruction from Voxelised FW LiDAR Data	8
6 Optimisation Attempts for the Surface Reconstruction	9
7 Alignment with Hyperspectral Imagery	10

8 Detection of Dead Standing Eucalyptus For Managing Biodiversity in Native Australian Forest	11
8.1 Introduction	11
8.2 Materials	15
8.3 Classification Challenges *** NEILL : everything from here is new *** .	19
8.4 Methods and Algorithms	20
8.5 Evaluation *** The following are not ready yet	27
8.6 Discussion	27
8.7 Future Work	28
9 Overall Results	30
10 Conclusions	31
10.1 Contributions	32
Bibliography	32
Appendices	i
A DASOS's user guide, released on the 20th of January 2017	ii
A.1 Introduction	ii
A.2 License	iii
A.3 Installation Guide	iv
A.4 Instructions	v
A.5 Exercises	xx
A.6 Limitations	xxv
A.7 Related Forums and Social Media	xxvi
B Case Study: Field Work in New Forest	xxvii
B.1 Introduction	xxvii
B.2 Validation Data Collected	xxviii
B.3 Landscape types	xxxii
B.4 Conclusions and Discussion	xxxv

List of Figures

3-1	The pipeline of the thesis	6
8-1	Animals Closes to Extinction	13
8-2	LiDAR point cloud showing that there are very limited points reflected from tree trunks.	14
8-3	The study area is depicted by green ($542km^2$), the yellow strips are the LiDAR flightlines and the red dots are the position of the field plots. <i>**Note: this image many need to be removed due to confidentiality of the company. I will talk with them and hopefully it will be ok.</i>	16
8-4	Structure of Red Gum Forest in south-eastern Australia.	17
8-5	Example of dead trees indicating their variance in shape.	17
8-6	Example of a dead tree in relation to the discrete LiDAR point cloud. .	18
8-7	Parameters used in Quick Terrain Modeller to obtain the DTM used here. .	21
8-8	Before and after subtracting the DTM.	21
8-9	This figure shows what priors were created for testing and how they are divided for cross validation.	24
8-10	Application of K-NN using the most significant features identified by Random Forest and filtering.	25
8-11	Removing the ground pixels	27
8-12	Thresholding, filtering, segmentation and calculating the dead trees' position.	28
A-1	Selecting Region of Interest	ix
A-2	Effect of modifying the user defined parameters; voxel length, isolevel and noise level.	ix
A-3	Example of fieldplot input	xv
A-4	Example of .csv files with a list of feature vectors exported.	xvii
B-1	The first area of interest and the related maps.	xxx
B-2	The second area of interest and the related maps.	xxxi

B-3	Trees that have been cut down	xxxii
B-4	Grass with a few scattered trees	xxxii
B-5	Dense forest	xxxii
B-6	Trees that have been cut down	xxxiii
B-7	Lakes and rivers	xxxiii
B-8	Trees that have been cut down	xxxiv
B-9	Animals in New Forest	xxxiv
B-10	Trees, which are mixed together	xxxv

Chapter 1

Introduction

- 1.1 Forest Monitoring: Importance and Applications
- 1.2 Background Information about Remote Sensing and Airborne Laser Scanning Systems

Chapter 2

Acquire Data

Chapter 3

Overview of Thesis

3.1 Problem

FW LiDAR systems have been available for a number of years but there still very few uses of FW LiDAR data. NERC-ARF has been acquiring airborne data for the UK and overseas since 2010 and it has more than 100 clients of new and archived data. Many clients request FW LiDAR data to be acquired, but despite the significant number of requests, the majority of research still only uses discrete LIDAR. Some of the factors regarding this slow intakes are:

- Typically FW datasets are 5 – 10 times larger than discrete data, with data sizes in the range of 50GB – 2.5TB GB for a single area of interest. NERC-ARF's datasets are up to 100GB each because most clients are research institutes but for commercial purposes each FW dataset is a couple of TB.
- Existing workflows are only able to work with the discrete data since the increased amount of information recorded within the FW LiDAR makes handling the quantity of data very challenging.

3.2 Aims and Objectives

This thesis explores visualisation and data-understanding for FW LiDAR systems and the overarching aim is to increase the accessibility FW LiDAR in remote forest surveying. The objectives are listed in Table 3.1 and they are associated with the Sections that tackles them.

No.	Objective	Related Chapters
1	Enable forestry experts with no computer science expertise to visualise and work with the FW LiDAR data.	5
2	Enable forest understanding through 3D visualisations of FW LiDAR data.	5
3	Improve and optimise visualisations of FW LiDAR data and hyperspectral images.	6 & 7
4	Enable browsing of very large scale datasets and spectral bands in an efficient manner.	6 & 7
5	Investigate data structures for faster iso-surface extraction of large volumetric datasets and efficient management of voxels.	6
6	Estimate tree coverage and investigate the potential of integrating multiple remote sensing datasets in forestry.	7
7	Dead tree detection in comparison to human detection and remote surveying with discrete LiDAR that will benefit biodiversity management.	8
8	Research whether terrain classification can be improved by the inference of high quality 3D information, for example, using priors over the space of 3D elements.	8

Table 3.1: Values of divisible sides

3.3 Overview

*** the following text has been taken from the IAA2 funding application

To address the limitations of existing workflows for using FW data we developed the open source software DASOS (from $\delta\alpha\sigma\omega\varsigma$ meaning forest in Greek) and novel algorithms that allow users, without computer science expertise, to work with and visualise large volumes of FW LiDAR data. Our open source software DASOS aims to remove the barriers preventing the use of FW LiDAR. Its contributions, and those of the new representations of the FW LiDAR, are demonstrated in three applications:

- Firstly, foresters can exploit their domain expertise to derive a wealth of information by observing the FW LiDAR data. We therefore improve visualisations for deriving information directly from the data, thus reducing travelling time and the associated expenses of getting into the forests. This cost includes appropriate cars and sometimes helicopters depending on the accessibility of the forests. While previous work on FW LiDAR visualisation talks about point cloud visualisation [1] and transparent voxels [2], DASOS is able to reconstruct the surfaces from the scanned area in 3D. This research further optimises visualisations by using the new FW LiDAR representations to accelerate this process by ****%. ***
I will complete the percentage once related test are completed
- Secondly, a fast way of aligning the FW LiDAR with Remotely Sensed Images has been developed in DASOS. Subsequently, by generating tree coverage maps, it has been shown that the combination of these datasets confers better remote survey results [3].
- Finally, DASOS allows the generation of 3D priors. An example usage of this information is characterising dead standing Eucalyptuses, which as explained at Section 1.1 are extremely beneficial for managing biodiversity in native Australian forests. This is work in progress and a comparison between the discrete and FW LiDAR will be performed to demonstrate the increased survey accuracy obtained when the FW LiDAR is used.

In summary, FW LiDAR has great potential to improving automated surveying accuracy and consequently reduce the expensive fieldwork conducted in forestry and this research has already started to have an impact in the FW LiDAR community. DASOS is now used at Interpine Group Ltd, a world leading Forestry Company in New Zealand, and a PhD student at Bournemouth University is evaluating it for use in the estimation of bird distributions in the New Forest in the UK.

*** end of copied text

3.4 Thesis Structure

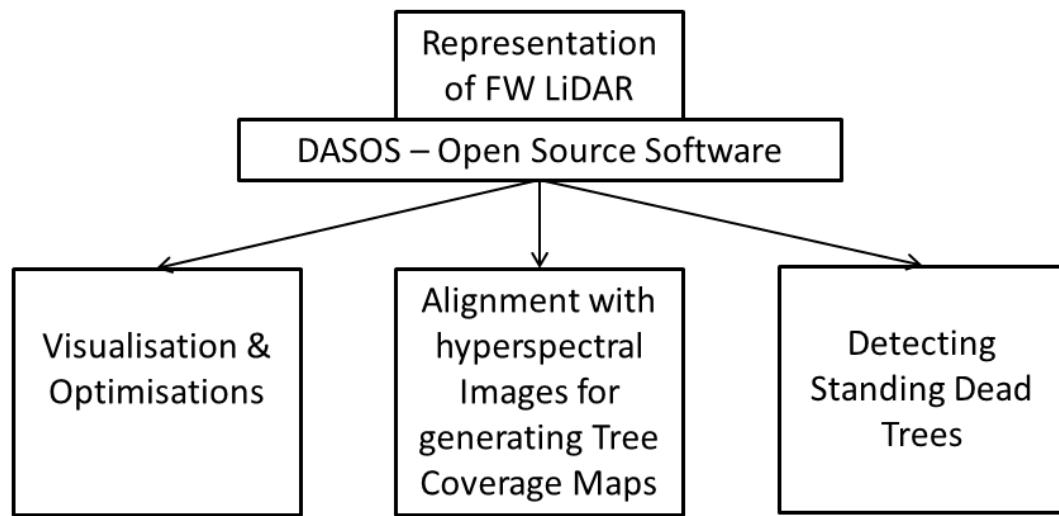


Figure 3-1: The pipeline of the thesis

Chapter 4

The open source software DASOS and the Voxelisation Approach

Chapter 5

Surface Reconstruction from Voxelised FW LiDAR Data

Chapter 6

Optimisation Attempts for the Surface Reconstruction

Chapter 7

Alignment with Hyperspectral Imagery

Chapter 8

Detection of Dead Standing Eucalyptus For Managing Biodiversity in Native Australian Forest

8.1 Introduction

8.1.1 The Importance of Dead Wood

The value of dead trees from a biodiversity management perspective is large. Once a tree dies, its contribution to our ecosystem continues. The woody structure remains for centuries and it contributes to forest regeneration while providing resources for numerous surrounding organisms [4]. As an indication, more than 4000 species inhabit dead wood in Finland [5], where an estimate of 1000 species has been extinct [6]. These species do not only include animals and birds but also organisms, like fungi. Fungi contributes to wood decaying, formation of hollows and biodiversity, which is an important factor for a resilient ecosystem [7]. Observing the changes of fungal diversity on decaying wood has an increased interest in science [8] [9] [10] in order to ensure the continuous existence of decaying wood in forests.

** NEill comma where: Specifically, in Australia, tree

Specifically in Australia, tree hollows play a significant role in managing biodiversity. Nearly all arboreal mammals rely on hollows with the exception of the Koala and perhaps Ringtail Possums that preferentially make a stick nest, but they use hollows as well. Additionally, a large number of Australian bird species rely on hollows for shelters

[11]. 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 take hundreds of years to form and because of that it is more likely to exist on dead trees. In Australia, studies predict shortage of hollows for colonisation in the near future [12] [13]. Therefore automated detection of them plays a significant role in protecting those animals. As an indicator of the importance of hollows in managing biodiversity, a list of a few of the species that rely on hollows was provided by the Forestry Corporation of NSW. Those species are shown at Figure 8-1. According to the Department of the Environment of Australian Government and the Government of Western Australia, six of them are protected, threatened or close to extinct [14] [15]. Figure 8-1 shows the species from the provided list and the six protected species have a red border and their names are bold in the description.

For the aforementioned reasons, monitoring dead trees is essential for having a resilient ecosystem. Nevertheless, the distribution of dead trees significantly varies making detection of them difficult [16]. Remote sensing approaches has been introduce to automate the process of monitoring forest and further increase the spatial resolution of the monitored area. The following section gives an overview of the related work undertaken in Remote Sensing.

8.1.2 Related Work

Remote Sensing was introduced for automatically detecting dead trees, because field-work is time consuming considering their variance spread and the size of the relevant forests. 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 data [17] [18]. Regarding standing dead trees, their shape (reduced number of leaves or broken branches) [19] and light reflectance (less green light illuminated) [20] are important factors for identifying them.

Previous work on dead standing trees detection performs single tree crown delineation before health assessment [19] [21]. Tree-crown delineation is usually done by detecting local maxima from the canopy height model (CHM) and then segmenting trees with watershed algorithm [22]. Improvements has been achieved by introducing markers controlled watershed [23] and structural elements of tree crowns with different sizes [24]. Additionally, Popescu and Zhao analyse the vertical distribution of the LiDAR points in conjunction with the local maximum filtering of CHM [25].

In the case of Eucalyptus, single tree detection is a challenge on its own, due to their



Figure 8-1: A number of species that rely on tree hollows of which the red ones / bold ones are close to extinction: 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 Owl, Barking Owl, **Masked Owl**, **Barn Owl**, White-throated Treecreeper, Hollow Owl, **Brush-tailed Possum** (mammal)¹

¹The images of the birds were taken from the following links (Retrieved on the 27th of April 2016): Kookaburra: <<http://tenrandomfacts.com/blue-winged-kookaburra/>>, Sulphur Crested Cockatoo: <<http://aussiegal7.deviantart.com/art/Sulphur-Crested-Cockatoo-08-153341893>>, Corella: <<http://www.theparrotplace.co.nz/all-about-parrots/long-billed-corella/>>, Superb Parrot: <<http://www.davidkphotography.com/?showimage=637>>, Crimson Rosella: <http://25.media.tumblr.com/tumblr_m3mo89c40r1r4t9h1o1_1280.jpg>, Eastern Rosella: <http://2.bp.blogspot.com/-pYxw51WjSOY/UB-LEFgd2KI/AAAAAAAAGw/9z60PUWE6TE/s1600/_GJS6601-as-Smart-Object-1.jpg>, Rainbow Lorikeet: <https://www.reddit.com/r/pics/comments/328fvc/a_rainbow_lorikeet_found_in_coastal_regions/>, Musk Lorikeet: <http://www.rymich.com/girraween/photos/animals/birds/medium/glossopsitta_concinna/glossopsitta_concinna_001.jpg>, Little Lorikeet: <<http://www.pbase.com/sjmurray/psittacidae>>, Red-winged Parrot: <<https://www.pinterest.com/pin/395894623469889727>>, Cockatiel: <<http://up.parsipet.ir/uploads/Cockatiels-for-sale.jpg>>, Australian Ringneck (Parrot): <<http://ontheroadmagazine.com.au/wp-content/uploads/2015/09/Twenty-eight-parrot-2-min.jpg>>, Red-rumped Parrot: <<http://parrotfacts.net/wp-content/uploads/Red-Rumped-Parrot-on-a-tree.jpg>>, Powerful Owl: <http://farm1.staticflickr.com/219/495796536_f78dac04c1.jpg>, Sooty Owl: <http://www.mariewinn.com/marieblog/uploaded_images/screech2-738532.jpg>, Barking Owl: <<http://www.pcpimages.com/Nature-and-Wildlife/Birds/i-7JKSTp5/1/L/owl%20%281%20of%201%29-L.jpg>>, Masked Owl: <http://www.survival.org.au/images/birds/masked_owl_2_600.jpg>, Galah: <<https://www.pinterest.com/pin/537546905498955709/>>, White-throated Treecreeper: <<https://geoffpark.files.wordpress.com/2011/09/female-white-throated-treecreeper.jpg>>,

irregular structure and multiple trunk splits. In other words, each tree trunks splits create a local maximum leading into over-segmentation when tree crowns are detected by local maxima filtering. Shendryk published a eucalyptus delineation algorithm that starts segmentation from bottom to top. In this paper, the trunks point cloud is separated from the leaves and individual trunks are identified before proceeding to crown segmentation [26]. Nevertheless, for that project only 17 flightlines of LiDAR data were collected. 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 commercial use and larger areas, the density of data collected is above the optimal resolution for a cost effective versus quality acquisition [27]. The project of this thesis is much larger. The resolution of our acquired LiDAR data has an average of four pulses per square meter, which is considered an optimal resolution in relation to the cost. But because of the tree height (up to 43m according to the fieldwork), a small amount of pulse intensity reached the trunks and the recorded waveform do not include enough information for individual trunk detection. An example of this project's discrete LiDAR data is shown in Figure 8-2 and the missing information about the trunks is depicted.

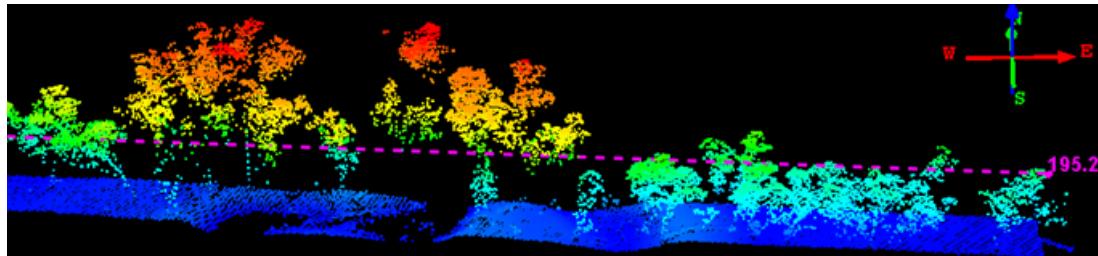


Figure 8-2: LiDAR point cloud showing that there are very limited points reflected from tree trunks.

*****Note read again to make sure it matches OK**

The acquired data are full-waveform LiDAR data. Traditional ways of interpreting FW LiDAR data, suggests extraction of a denser points cloud using Gaussian decomposition [28] [29]. Nevertheless, in this project we uses the open source software DASOS. DASOS was influenced by Persson et al, 2005, who used voxelisation to visualise the waveforms [2]. But, it does not only uses voxelisation for visualisations but also for extracting metrics useful in classification. It further normalises the intensities so that equal pulse length exists inside each voxel, making intensities more meaningful. It is further seems that the literature is moving towards voxelisation with promising results obtained at recent publication on tree species classification [30].

Hollow Owl: <http://www.mariewinn.com/marieblog/uploaded_images/screech2-738532.jpg>

Here, it is introduced an approach for quick dead tree detection derived from the boost cascade approach [31] but extended into 3D. This approach further contains similarities of the 3D tree shape signatures proposed by Dong, 2009, for distinguishing Oaks from Douglas fir tree crowns [32].

8.2 Materials

8.2.1 Study Area

The study area (Figure 8-3) is a native River Red Gum (*Eucalyptus camaldulensis*) forest of size 542km^2 in south-eastern Australia. The regeneration of the eucalyptus is extremely dependant in floods and therefore, their distribution in respect to density, health and age is highly variance [33]. Additionally, the height of *Eucalyptus camaldulensis* reaches up to $30 - 40\text{m}$ and their structural complexity is high with multiple trunk splits [34]. The size and structure of the forest, with a human as reference, is depicted in Figure 8-4, while examples of the variance shape of dead trees is shown in Figure 8-5.

8.2.2 Acquired full-waveform LiDAR data

Multiple-echo, full-waveform (FW) LiDAR data are supplied by RPS Australia East Pty Ltd. The data were acquired from 900m above ground level, using the Trimble AX60 Airborne LiDAR sensor, which was released in October 2013 [35]. The wavelength of the emitted laser was 1062nm, the maximum scan angle was 60 degrees, and the pulse rate was 400kHz. The acquisition was held from the 6th of March till the 31st of March 2015. The collected LiDAR were delivered into 206 flightlines, of which 13 are cross runs used for geometric correction. There is also a 30% of swath overlap. The point spacing along and across the track is 0.48m and the average point spacing is 4.3 points per square meter. Figure 8-6 shows an example of a dead tree in respect to the acquired discrete LiDAR point cloud. Detailed information about FW LiDAR related concepts are given in section 2.

8.2.3 Field Data

The field data were collected in July 2015 during the winter season of Australia and they include tree and canopy related measurements on circular plots. There are 33 plots with radius 35.68m and area 0.4ha allocated randomly inside the study area. On these plots, a total of 2386 trees were individually measured. Tree measurements include the geo-location, the trunk diameter at the standard height of 1.3m (breast height), height,

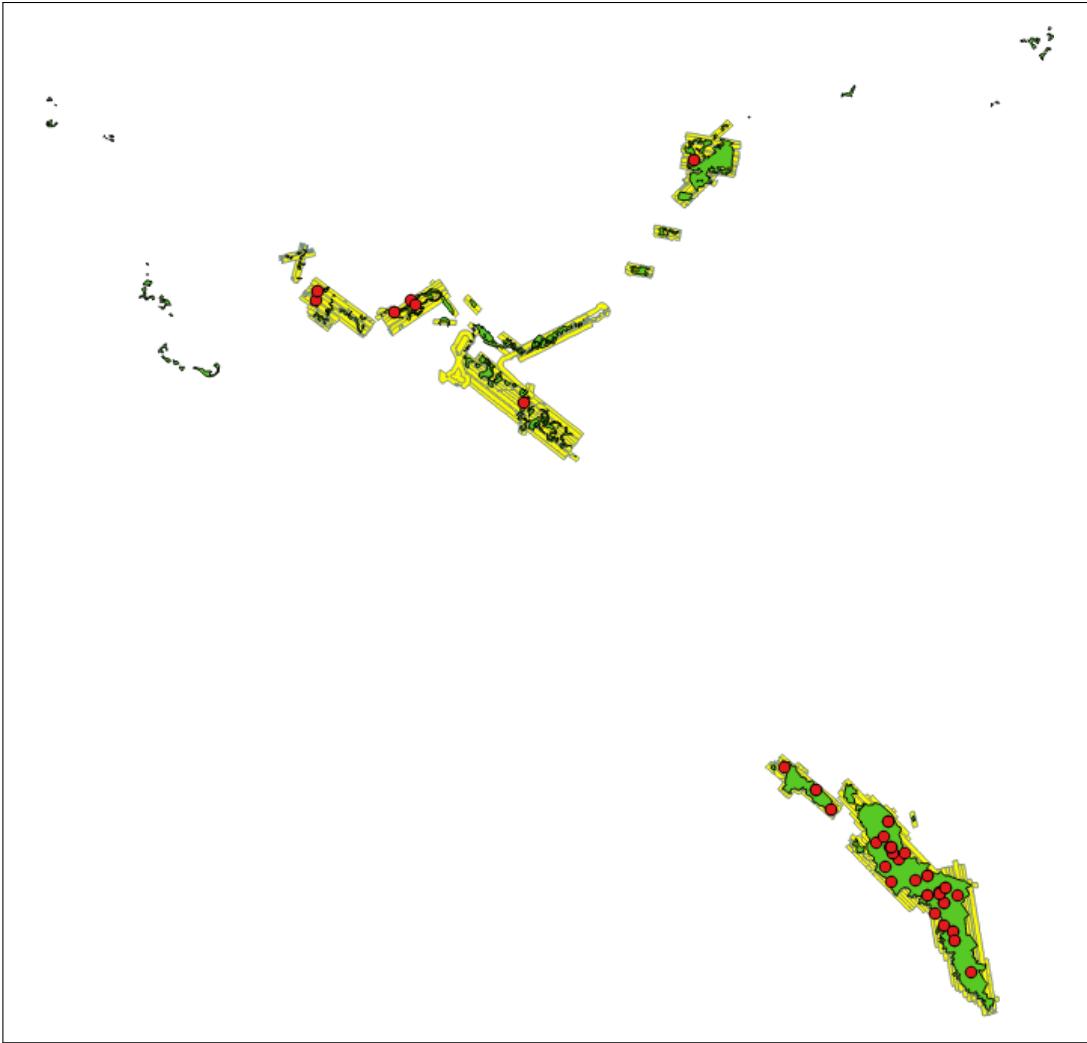


Figure 8-3: The study area is depicted by green (542km^2), the yellow strips are the LiDAR flightlines and the red dots are the position of the field plots. ****Note: this image many need to be removed due to confidentiality of the company. I will talk with them and hopefully it will be ok.**

species and health conditions (i.e. dead or alive). The geo-location of each tree is defined by the magnetic bearing from the centroid of the plot in degrees (range [1, 360]) and the distance from the centroid in meters. The northing and easting coordinates of the geo-location of each tree were calculated in post-processing. Here is worth mentioning that a single tree may be recorded as multiple trees if there is a trunk split bellow the breast height of 1.3m. Furthermore, 91.59% are River Red Gum and the rest are Black Box (*Eucalyptus largiflorens*) and Wattle group (*Acacia* spp.).

Inside the field data, there are 260 dead trees recorded. Nevertheless, not all of

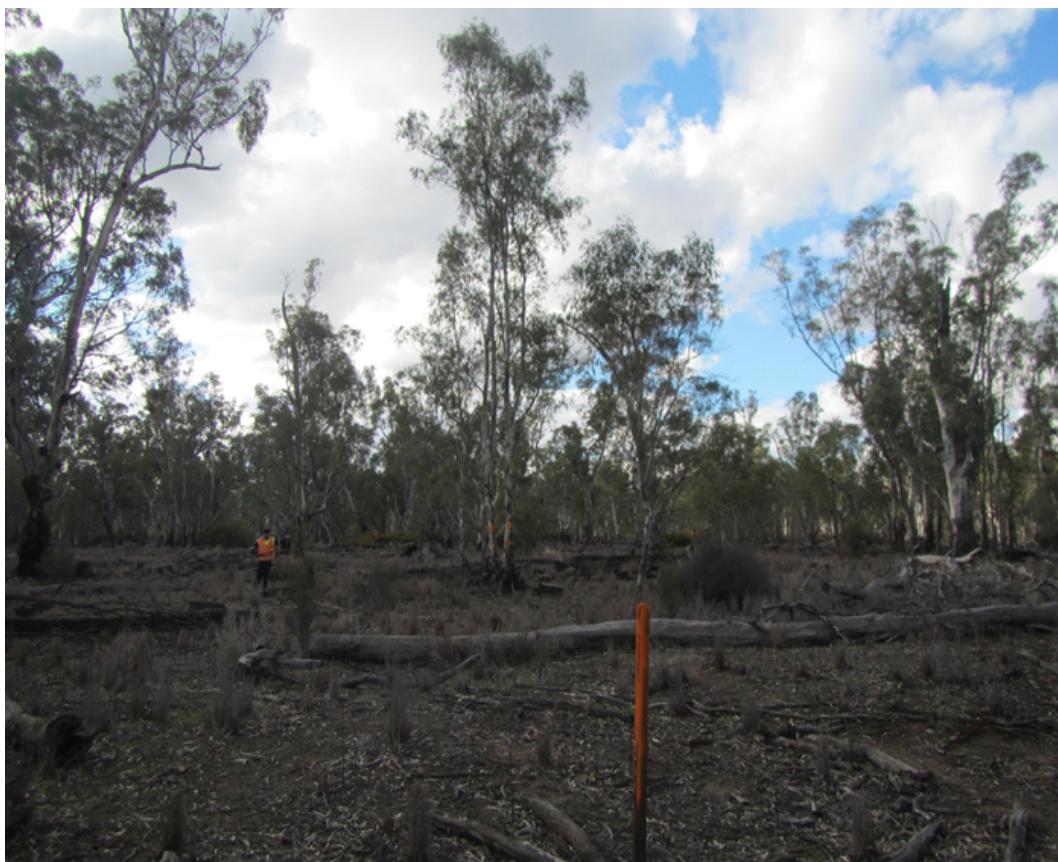


Figure 8-4: Structure of Red Gum Forest in south-eastern Australia.



Figure 8-5: Example of dead trees indicating their variance in shape.

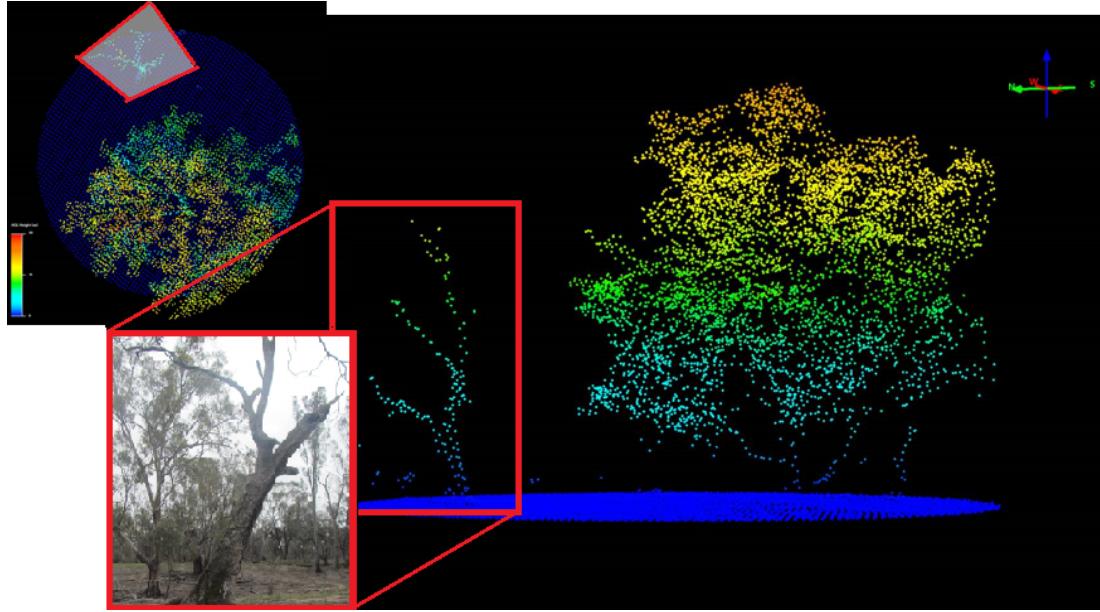


Figure 8-6: Example of a dead tree in relation to the discrete LiDAR point cloud.

those trees 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 data are not identifiable from the FW LiDAR data. Table 8.1 shows the number of dead and alive trees in respect to their DBH.

DBH	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

Table 8.1: Number of trees according to their DBH. **Note: I think it is in centimeter but I will confirm it with the company and add it afterwards.

Please note that the aforementioned field data were provided by Forestry Corporation of NSW, Wauchope, Australia and Interpine Ltd Group, New Zealand. For this thesis, a case study for collecting field data was conducted in New Forest, UK. This

helped to better understand classification challenges in forestry applications. More information about this study is provided in Appendix B.

8.3 Classification Challenges *** NEILL : everything from here is new ***

This section focuses on the challenges faced while working on the detection of dead standing eucalyptuses. Table 8.2 underlines these challenges, categorised into three groups: the nature of the study area, the acquired data and the field data. All these challenges influence the quality of the classifier and the accuracy of the results.

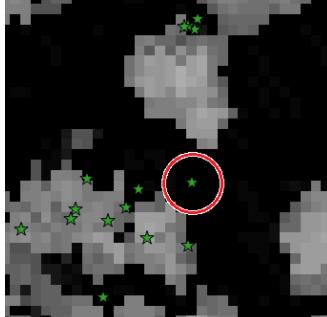
Study Area	Acquired Data	Field Data
<ul style="list-style-type: none"> The study area is a native eucalyptus forest. Native forests contain trees of different ages and heights. The height of a dead tree could be within the range of [1.5,40] meters. The density of the forest is highly variance. Sometimes the testing/training priors of the small dead trees may contain information from either nearby alive trees or ground. A tree may have dead branches but still be alive. Eucalyptus trees have irregular shapes and multiple trunk splits making tree delineation to require very dense acquired data. 	<ul style="list-style-type: none"> The pulse density of the acquired data does not allow bottom to top tree delineation. Crown detection from DEM (top) leads to over-segmentation due to the multiple trunk-splits. We, therefore, investigate the performance of object detection algorithms that do not require tree delineation. An important factor of identifying dead trees is the light reflectance, but for this project this kind of data (i.e. coloured imagery) was not acquired. Therefore, the classifier is only trained on tree shapes. But the shape of the tree is not an independent factor of identifying dead trees, since a tree may not have leaves but still be alive. 	<ul style="list-style-type: none"> If a tree has a trunk split below the 1.3m height, then it is recorded as multiple trees within the field data. This results into an inconsistency of the "one tree" concept. They contain small trees, which are non detectable from the acquired data. The accuracy of the geo-spatial positions is unknown. Even though it is claimed to be within centimetres, there are trees clearing appearing on the ground, once visualised on top of the DEM. An example: 

Table 8.2: The Classification challenges of automated detection of dead eucalyptuses

8.4 Methods and Algorithms

This section provides an deep explanation about the algorithms implemented. An overview of the work flow is given here:

1. Subtraction the Digital Terrain Model (DTM) from the FW LiDAR data
2. Generation 3D priors using DASOS
3. Identification of the most important variables of the 3D priors using random forest
4. Generation a probabilistic field using a weighted k-nearest neighbour (KNN) algorithm.
5. Height histogram and ground pixels removal
6. Thresholding dead pixels from alive, filtering, applying a seed growth algorithm for grouping nearby pixels and assignment of dead trees position.

8.4.1 Subtract DTM from FW LiDAR

DASOS has a feature for subtracting pre-calculated Digital Terrain Model (DTM) saved into .bil files. Generating DTM is beyond the scope of this research and the DTM files used, were provided by Interpine Ltd Group. The provided DTM files were generated using the Quick Terrain Modeller from the discrete LiDAR using the parameters shown in Figure 8-7.

The subtraction of the DTM is done during the voxelisation (Section 4). The terrain height is subtracted from the position of the sample before it is inserted into the volume. Please note that this terrain value is not subtracted from the origin of each pulse but from the position of each sample since the terrain value at the origin and the terrain value at the position of a sample may differ.

Figure 8-8 shows an example of a DEM generated before and after the subtraction using DASOS.

8.4.2 Generating 3D priors using DASOS

The 3D priors is a new feature of DASOS (version 2), which was released on the 20th January 2017 [36]. The dead tree detection is its first application. This feature is useful for characterising object inside the 3D space (e.g. trees). For each column of interest within the voxelised FW LiDAR data, information around its local area are exported as a prior. Multiple priors are listed within .csv files for easy manipulation into software packages specialised in statistical analysis like R and matlab. There are

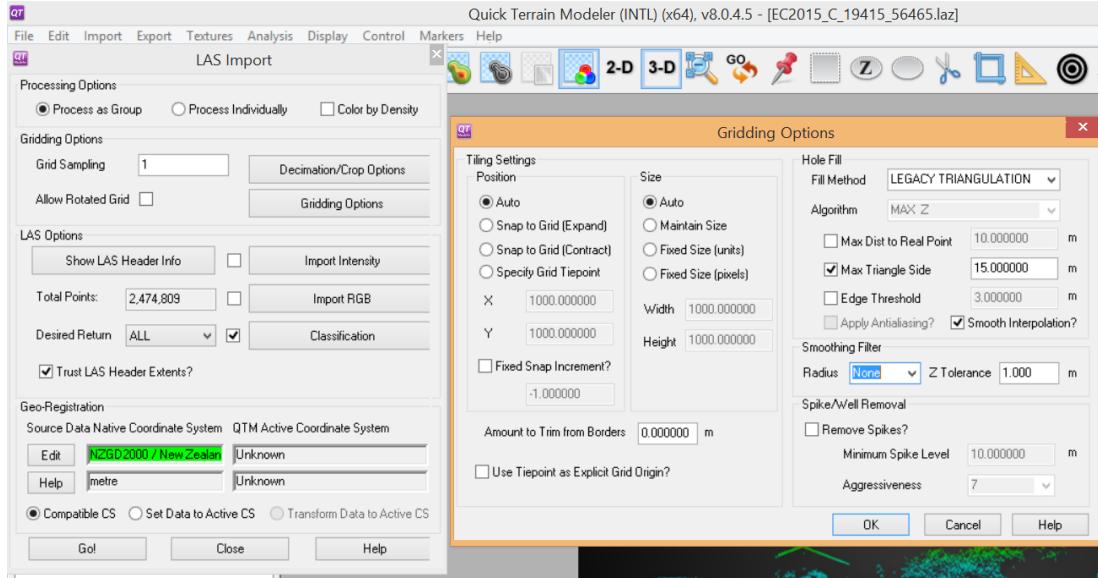
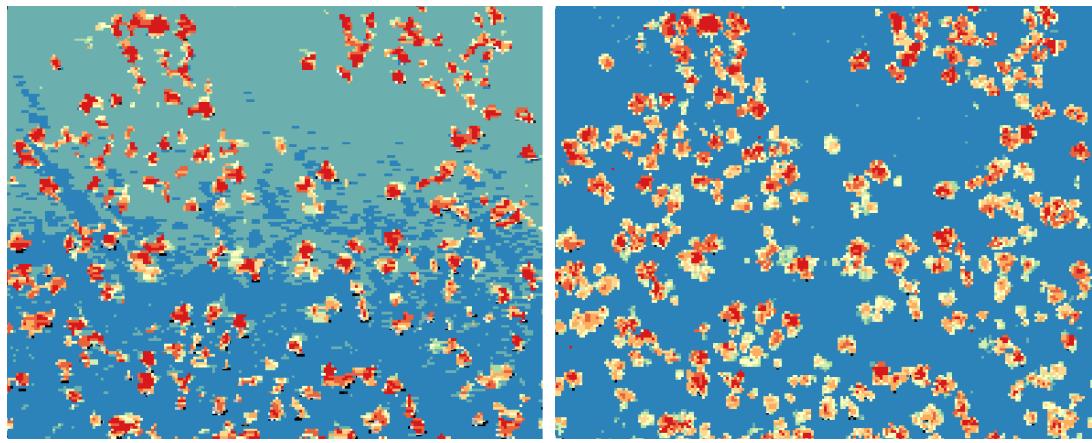


Figure 8-7: Parameters used in Quick Terrain Modeller to obtain the DTM used here.



(a) The DEM before subtracting the DTM (b) The DEM after subtracting the DTM
Figure 8-8: The difference of the DEM before and after subtracting the terrain height. The red indicates big height, while the darker the blue is the lower the DEM is.

two types of exported information from these local areas: processed and raw. If the processed option is chosen, then information like distribution of non-empty voxels and standard deviation of heights are listed. A sample of the exported processed information along with explanations is given in Table 8.3, while the entire list is provided within the Appendix A. If the exported parameters are raw, then the corresponding intensity values of the local area's voxels are exported. Additionally, there are two available shapes of priors (the cuboid and the cylinder). The size of priors is also user defined.

Here, the aforementioned feature of DASOS is used for generating 3D priors used for training and testing the dead tree detection algorithms implemented.

Explanation of some features of DASOS's 3D priors that proved to be useful for building the classifier		
No	Label	Description
1	Height_Middle_Column	The height of the middle column of the prior
	Height_Mean	The Mean height of all the columns included in the template
	Height_Median	The Median height of all the columns included in the template
1	Height_Std	The Standard Deviation of the heights of the columns included in the template
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 voxel that contains an intensity above the isolevel
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 voxel that contains an intensity above the isolevel
8	Dis_Median	Median distance from the central voxel to every voxel that contains an intensity above the isolevel
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

Table 8.3: Explanation of some features of DASOS's 3D priors that proved to be useful for building the classifier. All the features are explained in Appendix A

Within the field data, some plots exist on two flightlines due to the overlapping of the flights. Overlaps happen at the edges of the flightlines and their scan angle

significantly varies. For that reason, each unique set of field plot and corresponding flightline is considered as a test/training plot. This results into 50 plots. These plots were randomly divided into 5 equal training datasets. Another dataset was also created by merging the first, second and third dataset in order to check whether the increased training data improves the classification accuracy.

The priors 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 priors per plot. For each type, three .csv files are generated. The first one contains the priors of the dead trees, the second one contains the priors of the alive trees and the third one contains one prior for each column of the voxelised space. The first two are used for training the classifier and the last one for testing. The dimensions of the priors were chosen to be a bit smaller than the estimated average size of the dead trees to reduce the size of the irrelevant information contained within the priors. Figure 8-9 depicts the divisions of the datasets and the information about the priors generated.

8.4.3 Random Forest

Random Forest is able to identify the importance of predicting variables. At first, it generates multiple regression trees by randomly sampling the data at its nodes and choosing the best predicting variables for each sampled data. The variable importance is then defined according to influence it has to the classification once this variable is modified and the rest remain unchanged [37]. In this project, the R package is used for finding the most relevant feature of the 3D priors (Section 8.4.2 in identifying dead trees).

At this point, it worth highlighting that Random Forest failed to find relation between the 3D priors with the "Raw Intensities" due to the irregular shapes of Eucalyptus trees and the variant scan angle of each field plot. Nevertheless, "Raw Intensities" may be useful for other classification, e.g. pine trees in commercial forest, where their shape variance is smaller.

Regarding the "Processed Intensities", Figure 8-10a shows a list with the variable importance according to Random Forest and Table 8.3 gives the explanation of each important variable identified. The most important one is the standard deviation of height. This is reasonable since the canopy of dead trees has bigger height variance in comparison to alive trees whose canopy is leafy. Please note that in Figure 8-10a the union of all datasets is used and that the significant features slightly vary depending on each sub dataset used.

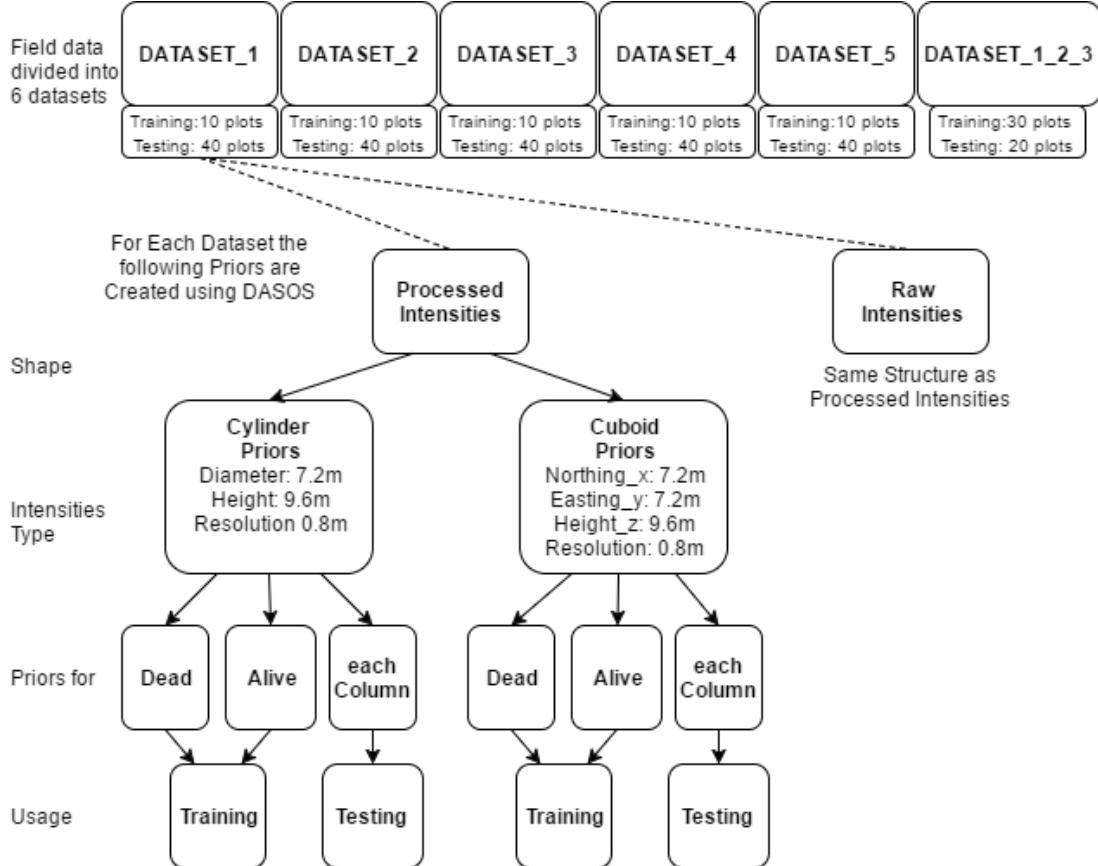


Figure 8-9: This figure shows what priors were created for testing and how they are divided for cross validation.

8.4.4 K-Nearest Neighbour, Filtering and Smoothing

Please skip kNN chapter for now, since I got stuck on writing the equations

Once the ten most significant variables are identified using the Random Forest, the k -nearest neighbour algorithm is applied to generate a probabilistic field. Within the algorithm, each selected variable is associated with a weight according to its importance ($W = w_1, w_2, \dots, w_{10}$). Additionally the number k is set to 5, because the training datasets are relatively small and they may contain wrong data due to the low accuracy of the field data (Section 8.3). There are also positive training datasets (dead) and negative (alive).

The assigned pixel value calculated as follow:

$$P = P(\text{dead}) / (P(\text{dead}) + P(\text{alive})) \quad (8.1)$$

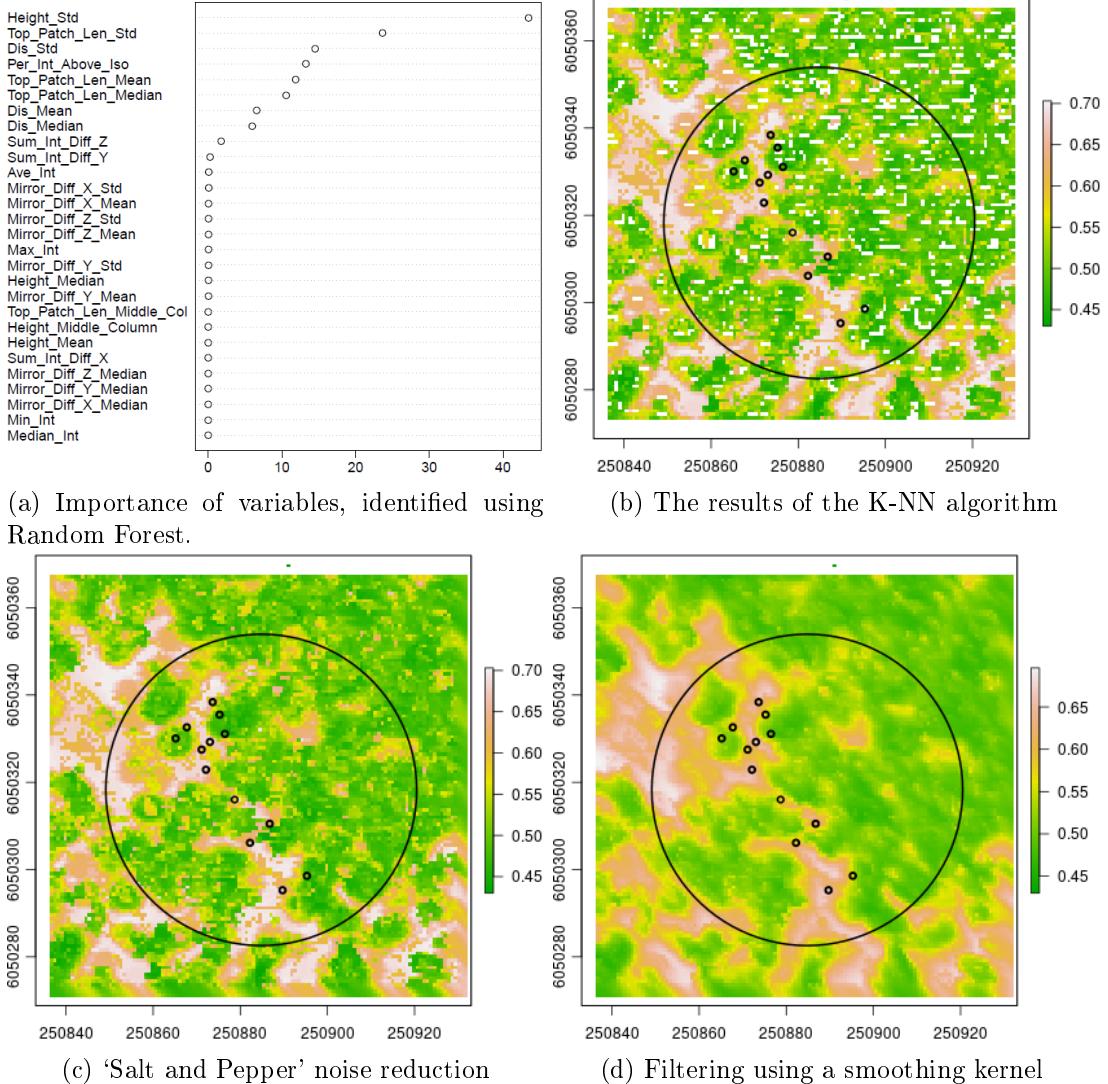


Figure 8-10: Application of K-NN using the most significant features identified by Random Forest and filtering.

$$\sum_{i=1}^{10} w_i \times (P_i - D_i)^2 \quad (8.2)$$

$$P_i = P_{i1}, P_{i2}, \dots, P_{i10} \quad D_i = D_{i1}, D_{i2}, \dots, D_{i10}$$

$$i = \arg \max |\mathbf{v} - \mathbf{g}_i|^2. \quad (8.3)$$

Salt and Pepper noise appears on the results (Figure 8-10b) because of empty columns existing within the volume. This is removed (Figure 8-10c) using a median filter which assigns to every empty pixel the median value of its non-empty neighbouring pixels. By the end a smoothing filter is applying for further noise reduction (Figure 8-10d).

8.4.5 Removing Ground Pixels

Removing the ground pixels is a trivial task because the DTM has already subtracted from the data and therefore the height of the ground is approximately constant. A histogram of the height values was generated. As shown in Figure 8-11b, there are the three well-defined classes (ground, trees and noise). The ground and noise are removed using simple thresholding. This processed is illustrated in Figure 8-11.

8.4.6 Threshold, Filtering, Segmentation and Position Assignment

In order to obtain the positions of the dead trees, there are four steps left:

1. Thresholding
2. Filtering
3. Segmentation
4. Position assignment

Up to this stage, we have an image of the probabilistic field and the ground has been removed (Figure 8-11d). After that a threshold for separating dead and alive pixels is chosen using the training data and the alive pixels are removed (Figure 8-12a). The output image contains pixels which are classified as dead by are away from the rest. To reduce over-detection of dead trees, these pixels are filtered out (Figure 8-12b). Afterwards, the pixels are grouped into trees relatively to their neighbouring pixels using a seed growth segmentation algorithm (Figure 8-12c). By the end, it is assumed that each segment is a dead tree and its position is calculated by the average of geo-spatial location of its pixels (Figure 8-12d).

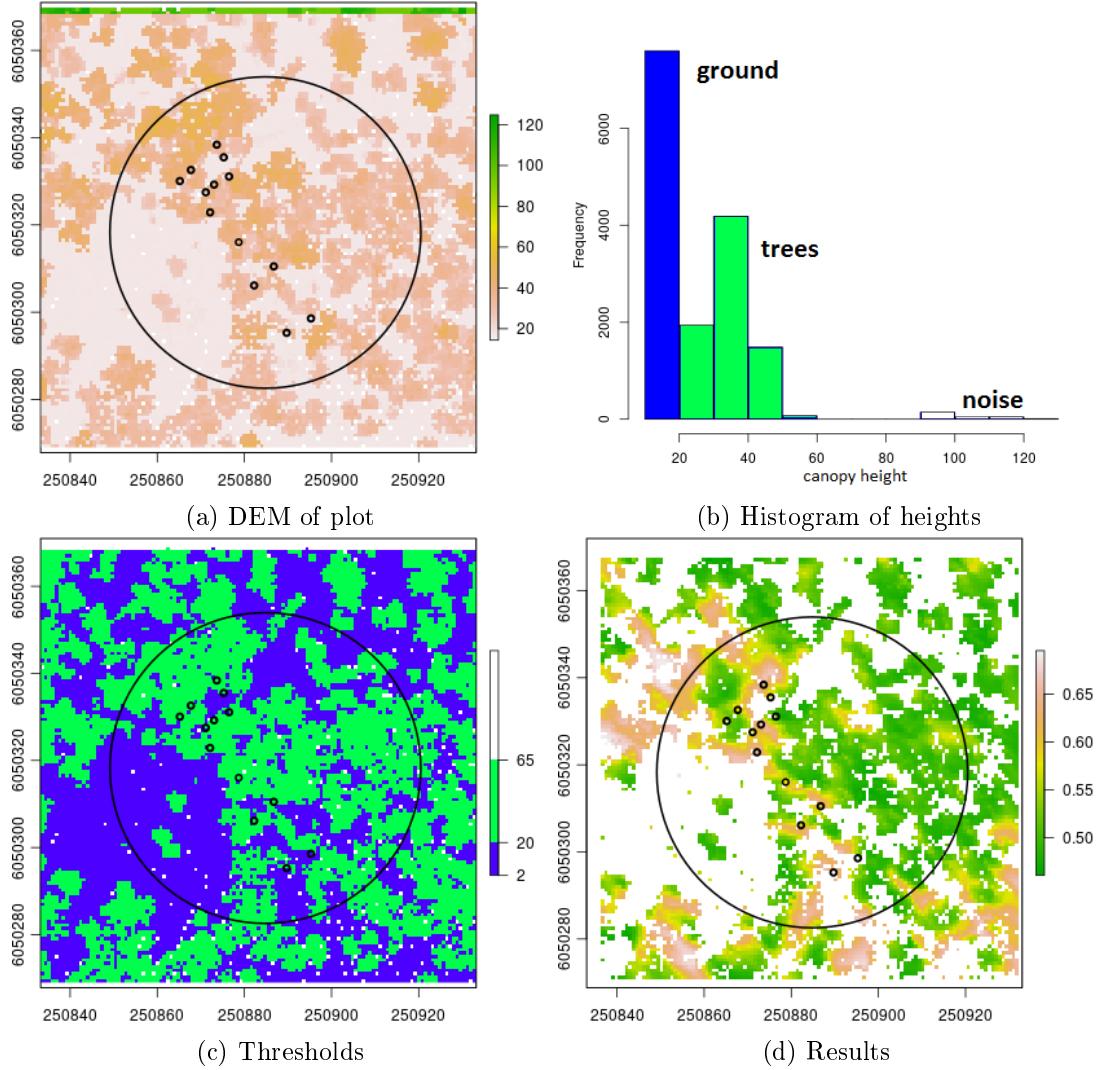


Figure 8-11: Removing the ground pixels

8.5 Evaluation *** The following are not ready yet

8.5.1 Distance Related Evaluation

8.5.2 Pixelwise Evaluation

8.6 Discussion

Dead tree detection is a difficult task due to the irregular shapes of the trees and different sizes. Here we produced this algorithm (pla pla) which is new because it doesn't need tree segmentation but has a lot of room for improvement.

Also don't know the accuracy of the tree position and as we can see at some height

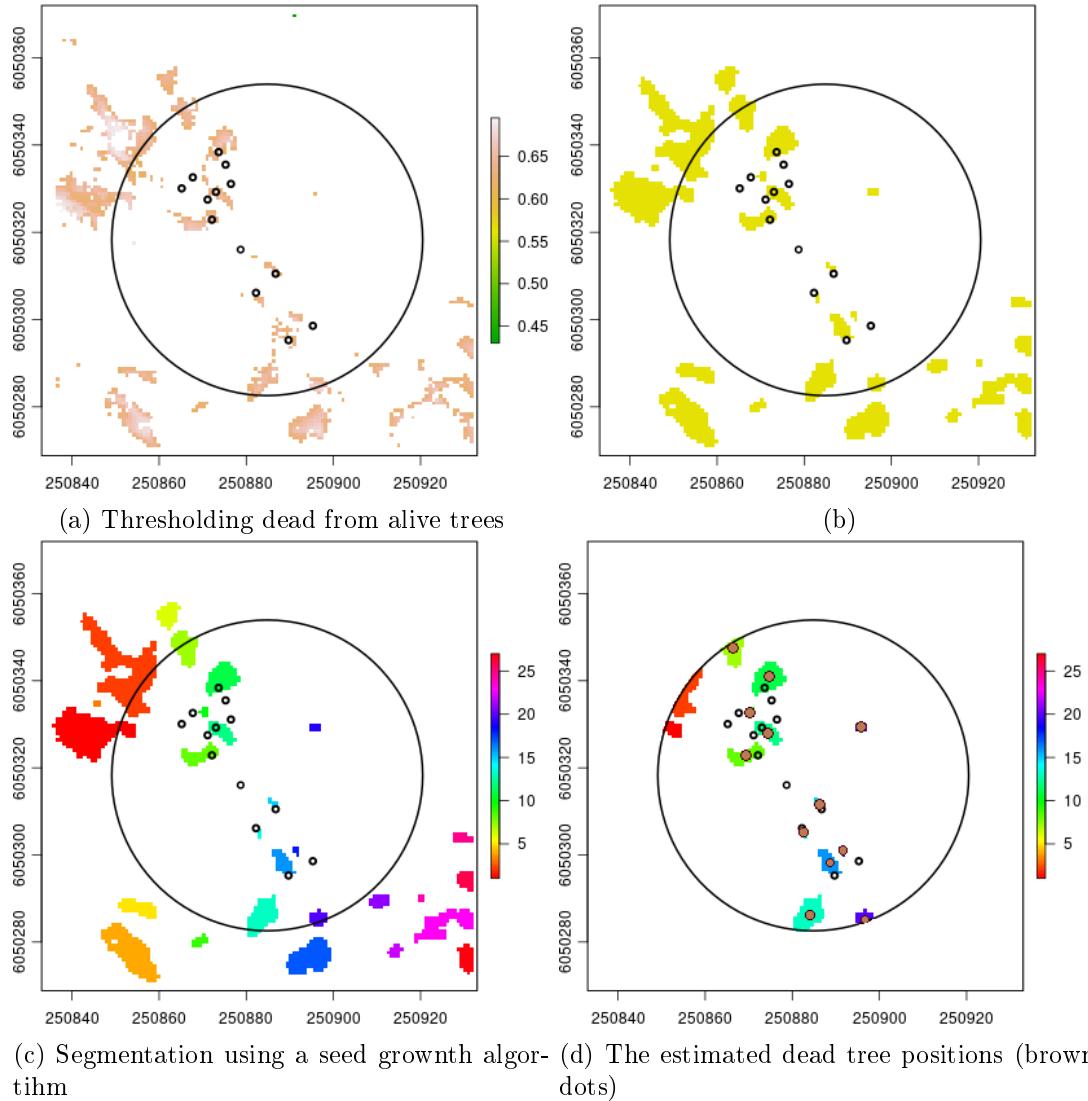


Figure 8-12: Thresholding, filtering, segmentation and calculating the dead trees' position.

maps there are places where there are trees according to the fieldplots but the data clearly shown that there are not trees

Here it is worth mentioning that the dead tree detection is the first application of DASOS's 3D priors.

8.7 Future Work

- Manually check and improve position of dead trees using visualisations of the data. In order to improve accuracy of test and evaluating data

- Separate trees from field data according to their height because tree with different heights have different shape properties and the priors used had constant size
- Create priors that have adjustable size according to the height of the tree
- After the seed growth algorithm, check the size of the segments and look into the possibility of merging two segments into one or dividing a segments into multiple sub-segments.
- Test the results when only using dead trees for training data and not alive
- The system is usually confused at the edges of the alive trees. Research on how this could be improved.

Chapter 9

Overall Results

Chapter 10

Conclusions

10.1 Contributions

Bibliography

- [1] M. Isenburg, “Pulsewaves: An open, vendor-neutral, stand-alone, las-compatible full waveform lidar standard.,” 2012.
- [2] A. Persson, U. Soderman, J. Topel, and S. Ahlberg, *Visualisation and Analysis of full-waveform airborne laser scanner data*. V/3 Workshop, Laser scanning 2005, 2005.
- [3] M. Miltiadou, M. A. Warren, M. Grant, and M. Brown, “Alignment of hyperspectral imagery and full-waveform lidar data for visualisation and classification purposes,” *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, no. 7, p. 1257, 2015.
- [4] J. F. Franklin, H. H. Shugart, and M. E. Harmon, “Tree death as an ecological process,” *BioScience*, vol. 17, no. 8, pp. 550–556, 1987.
- [5] J. Siitonen, “Forest management, coarse woody debris and saprophytic organisms: Fennoscandian boreal forests as an example,” *Ecological bulletins*, pp. 11–41, 2001.
- [6] I. Hanski, “Extinction debt and species credit in boreal forests: modelling the consequences of different approaches to biodiversity conservation,” *Annales Zoologici Fennici*, pp. 271–280, 2000.
- [7] G. Peterson, C. R. Allen, and C. S. Holling, “Ecological resilience, biodiversity, and scale.,” *Ecosystems*, vol. 1, no. 1, pp. 6–18, 1998.
- [8] N. Abrego and I. Salcedo, “How does fungal diversity change based on woody debris type? a case study in northern spain,” *Ekologija*, vol. 57, no. 3, 2011.
- [9] J. N. Stokland and K. H. Larsson, “Legacies from natural forest dynamics: Different effects of forest management on wood-inhabiting fungi in pine and spruce forests,” *Forest Ecology and Management*, vol. 261, no. 11, pp. 1707–1721, 2011.

- [10] D. Lonsdale, M. Pautasso, and O. Holdenrieder, “Wood-decaying fungi in the forest: conservation needs and management options,” *European Journal of Forest Research*, vol. 127, no. 1, pp. 1–22, 2008.
- [11] P. Gibbons and D. Lindenmayer, *Tree Hollows and Wildlife Conservation in Australia*. CSIRO Publishing, 2002.
- [12] D. B. Lindenmayer and J. T. Wood, “Long-term patterns in the decay, collapse, and abundance of trees with hollows in the mountain ash (eucalyptus regnans) forests of victoria, southeastern australia,” *Canadian Journal of Forest Research*, vol. 40, no. 1, pp. 48–54, 2010.
- [13] R. L. Goldingay, “Characteristics of tree hollows used by australian birds and bats,” *Wildlife Research*, vol. 36, no. 5, pp. 394–409, 2009.
- [14] “List of extinct, threatened and near threatened australian birds,” *Environment Protection and Biodiversity Conservation Act*, 1999.
- [15] Government of Western Australia, “Oarks and wildlife the list of threatened and priority fauna list,” tech. rep., November 2015.
- [16] Y. Kim, Z. Yang, W. B. Cohen, D. Pflugmacher, C. L. Lauver, and J. L. Vankat, “Distinguishing between live and dead standing tree biomass on the north rim of grand canyon national park, usa using small-footprint lidar data.,” *Remote Sensing of Environment*, vol. 113, no. 11, pp. 2499–2510, 2009.
- [17] P. Polewski, W. Yao, M. Heurich, P. Krzystek, and U. Stilla, “Detection of fallen trees in als point clouds using a normalized cut approach trained by simulation,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 105, pp. 252–271, 2015.
- [18] W. Mücke, B. Deák, H. M. Schroiff, A., and N. Pfeifer, “Detection of fallen trees in forested areas using small footprint airborne laser scanning data,” *Canadian Journal of Remote Sensing*, vol. 139, no. s1, pp. S32–S40, 2013.
- [19] W. Yao, P. Krzystek, and M. Heurich, “Identifying standing dead trees in forest areas based on 3d single tree detection from full-waveform lidar data,” *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. I-7, pp. 359–364, 2012.
- [20] J. Pasher and D. J. King, “Mapping dead wood distribution in a temperate hardwood forest using high resolution airborne imagery,” *Forest Ecology and Management*, vol. 258, no. 7, pp. 1536–1548, 2009.

- [21] I. Shendryk, M. Broich, M. G. Tulbure, A. McGrath, D. Keith, and S. V. Alexandrov, "Mapping individual tree health using full-waveform airborne laser scans and imaging spectroscopy: A case study for a floodplain eucalypt forest," *Remote Sensing of Environment*, vol. 187, pp. 202–217, 2016.
- [22] S. C. Popescu, R. H. Wynne, and R. F. Nelson, "Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass," *Canadian journal of remote sensing*, vol. 29, no. 5, pp. 564–577, 2003.
- [23] L. Jing, B. Hu, J. Li, and T. Noland, "Automated delineation of individual tree crowns from lidar data by multi-scale analysis and segmentation," *Photogrammetric Engineering & Remote Sensing*, vol. 78, no. 12, pp. 1275–1284, 2012.
- [24] B. Hu, J. Li, L. Jing, and A. Judah, "Improving the efficiency and accuracy of individual tree crown delineation from high-density lidar data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 26, pp. 145–15, 2014.
- [25] S. C. Popescu and K. Zhao, "A voxel-based lidar method for estimating crown base height for deciduous and pine trees," *Remote sensing of environment*, vol. 112, no. 3, pp. 767–781, 2008.
- [26] I. Shendryk, M. Broich, M. G. Tulbure, and S. V. Alexandrov, "Bottom-up delineation of individual trees from full-waveform airborne laser scans in a structurally complex eucalypt forest," *Remote Sensing of Environment*, vol. 173, pp. 69–83, 2016.
- [27] J. L. Lovell, D. L. B. Jupp, G. J. Newnham, N. C. Coops, and D. S. Culvenor, "Simulation study for finding optimal lidar acquisition parameters for forest height retrieval.,," *Forest Ecology and Management*, vol. 214, no. 1.
- [28] A. Neuenschwander, L. Magruder, and M. Tyler, "Landcover classification of small-footprint full-waveform lidar data," *Jounal of Applied Remote Sensing*, vol. 3, no. 1, pp. 033544–033544.
- [29] J. Reitberger, P. Krzystek, and U. Stilla, "Analysis of full waveform LiDAR data for tree species classification," *International Journal of Remote Sensing*, vol. 29, no. 5, pp. 1407–1431, 2008.
- [30] L. Cao, N. Coops, L. Innes, J. Dai, and H. Ruan, "Tree species classification in subtropical forests using small-footprint full-waveform lidar data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 49, pp. 39–51, 2016.

- [31] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” *Computer Vision and Pattern Recognition*, vol. 1, 2001.
- [32] P. Dong, “Characterization of individual tree crowns using three-dimensional space signatures derived from lidar data.,” *International Journal of Remote Sensing*, vol. 30, no. 24, pp. 6621–6628, 2009.
- [33] J. Kerle, “Collation and review of stem density data and thinning prescriptions for the vegetation communities of new south wales,” 2005.
- [34] N. Wilson and N. C. C. of N.S.W, “The flooded gum trees : land use and management of river red gums in new south wales,” *The Council, Sydney*, 1995.
- [35] E. van Rees, “Trimble’s ax60i and ax80,” *GeoInformatics*, vol. 7, no. 5.
- [36] M. Miltiadou, N. D. F. Campbell, M. Brown, A. S. G., M. Warren, D. Clewley, and M. Grant, “User guide of the 2nd version o dasos,” 2017.
- [37] A. Liaw and M. Wiener, “Classification and regression by randomforest,” *R news*, vol. 2, no. 3.
- [38] R. Crippen, “Calculating the vegetation index faster,” *Remote Sensing of Environment*, vol. 34, no. 1, pp. 71–73, 1990.
- [39] R. B. Smith, *Introduction to Hyperspectral Imaging*. MicroImages, 2014.
- [40] M. Sumnall, *Assessment of habitat condition and conservation status for lowland British woodland using earth observation techniques*. Bournemouth University: Unpublished PhD thesis, 2013.

Appendices

Appendix A

DASOS's user guide, released on the 20th of January 2017

A.1 Introduction

FW LiDAR systems have been available for a number of years but there still very few uses of FW LiDAR data. NERC-ARF has been acquiring airborne data for the UK and overseas since 2010 and it has more than 100 clients of new and archived data. Many clients request FW LiDAR data to be acquired, but despite the significant number of requests, the majority of research still only uses discrete LIDAR. Some of the factors regarding this slow takeup are:

- Typically FW datasets are 5 – 10 times larger than discrete data, with data sizes in the range of 50GB – 2.5TB GB for a single area of interest. NERC-ARF's datasets are up to 100GB each because most clients are research institutes but for commercial purposes each FW dataset is a couple of TB.
- Existing workflows are only able to work with the discrete data since the increased amount of information recorded within the FW LiDAR makes handling the quantity of data very challenging.

The open source software DASOS was developed to encourage foresters to use the FW LiDAR data. DASOS was named after the Greek word "*δάσος*" (=forest) and it was firstly presented at the 36th International Symposium of Remote Sensing of the Environment, 2015. The main way of interpreting FW LiDAR data in DASOS is fundamentally different from the state-of-art available software packages. In a few words, the FW LiDAR data are voxelised by inserting the wave samples into a 3D discrete density volume. It accumulates intensity values from multiple shots and stores

them into a 3D regular grid, resolving this way the problem with the sinusoidal footprints pattern of the Leica system. For more information please refer to the related paper at <https://www.researchgate.net/publication/277347868_Alignment_of_hyperspectral_imagery_and_full-waveform_LIDAR_data_for_visualisation_and_classification_purposes>

This user guide aims to give an in depth understanding of DASOS's functionalities. In a few words there are three functionalities of DASOS:

1. the construction of 3D polygon meshes;
2. the generation of 2D metrics aligned with hyperspectral images and
3. the characterisation of objects using feature vectors.

A.2 License

DASOS is released under the GNU General Public Licence, Version 3. The full description of the usage licence is available here:<<https://github.com/Art-n-MathS/DASOS/blob/master/License.txt>>

The following paper is the paper that introduced DASOS and it must be cited in any publications, software or other media using DASOS:

Miltiadou M., Warren M. A., Grant M., Brown M., 2015, Alignment of Hyperspectral Imagery and full-waveform LiDAR data for visualisation and classification purposes, ISPRS Archives 36th International Symposium of Remote Sensing of the environment. [3]

Full paper available here: at:<https://www.researchgate.net/publication/277347868_Alignment_of_hyperspectral_imagery_and_full-waveform_LIDAR_data_for_visualisation_and_classification_purposes>

The 1st sample dataset provided for testing was collected by the NERC Airborne Research and Survey Facility (ARSF). Copyright is held by the UK Natural Environment Research Council (NERC). The data are free for non-commercial use, but NERC-ARSF must be acknowledged in any publications, software or other media that make use of these data.

The 2nd sample dataset provided by Interpine Group Ltd. Copyright is held by Interpine Group Ltd and the data are free for non-commercial use, but Interpine Group Ltd must be acknowledged for any publications, software or other media that make use of these data.

A.3 Installation Guide

A.3.1 Windows

The windows executables are available at:

<https://github.com/Art-n-MathS/DASOS/tree/master/DASOS_win>

DASOS is a command line program and a terminal is required. For Windows XP, Vista and 7, Command Prompt is the default terminal and it can be found from the search tab on the start menu. If you are using Windows 8, then right click on the start icon and choose 'Command Prompt'. Find the directory where DASOS is saved (the command 'dir' shows the files inside the current directory and the command 'cd' to open folder). Once you are in the correct directory, execute the following command to test that the program works fine.

```
$: DASOS --help1
```

Information of all the available commands should be printed. If an error is occur, then either a .dll file is missing or DASOS is not supported at your computer.

A.3.2 Linux

The source code is available at: <<https://github.com/Art-n-MathS/DASOS>>.

For compiling DASOS on Linux, there are three major dependencies:

1. qmake-qt4 (or later release) / qtcreator
2. gmtl library - please update the .pro file to point to the correct directory
3. -std=c++11

Once those are installed compile DASOS as shown below:

```
$: qmake-qt4
```

```
$: make
```

To test it, write the following command:

```
$: DASOS --help
```

Information of all the available commands should be printed.

¹the '\$:' is not included in the command. It just illustrate the start of a command in the terminal

A.4 Instructions

A.4.1 Overview

DASOS is a command line program and can either be used in Command Prompt on Windows or a Linux shell.

At first change directory (cd) to go to the directory DASOS is saved or compiled in. Then for Windows run:

```
$: DASOS <arg1> <arg2> ... <argN>  
or $: DASOS.exe <arg1> <arg2> ... <argN>
```

On Linux run:

```
$: ./DASOS <ar1> <ar2> ... <argN>
```

For consistency this guide uses the 1st Windows example since all of the inputs, parameters and output arguments are the same for both Windows and Linux.

The tags of DASOS are divided into three groups: Inputs, Parameters and Outputs.

1. Inputs (Section A.4.2)
2. Parameters (Section A.4.3)
3. Outputs (Section A.4.4)

Even though many tags are optional or contain default values, it's essential to follow the order <inputs> <parameters> <outputs> because if the outputs are defined first unexpected results may occur, due to adding outputs to the stack before parameters are initialised. The aforementioned Sections give an explanation of all the possible tags of DASOS.

Before proceeding to the explanation, it worth highlighting and numbering the three main outputs of DASOS. The corresponding sub-sections do not only explain the output but also the parameters that are only specific to the corresponding output. The user guide refers to those outputs using their numbers:

1. The generation of 3D polygonal meshes (Section A.4.4.1)
2. The 2D metrics aligned with Hyperspectral Imagery (Section A.4.4.2)
3. The feature vectors for local inspection of data (Sub-section A.4.4.3)

At the end of this guide, there is a list of limitations and the dependencies.

A.4.2 Inputs

The inputs are divided into FW LiDAR, Hyperspectral and fieldplots. The FW LiDAR are compulsory for all the functionalities of DASOS. The Hyperpsectral Inputs are optional for the 1st and 2nd output of DASOS (3D polygonal meshes and 2D metrics), while the fieldplots option is compulsory for the 3rd option, the feature vectors.

Table A.1 explains the tags for loading the FW LiDAR files files. Please note that it is compulsory to load one of those options. If more than one FW LiDAR files are loaded then it is essential to keep consistency between projects; load only one type of full-waveform LiDAR data simultaneously. Table A.2 outlines how the hyperspectral imagery is loaded, how to subtract a pre-calculated DTM and also how the file with fieldplots is loaded if the 3rd output option is chosen (feature vectors).

Tags	Description
-las <file1> <file2> ... <fileN>	The name/directory of a number of LAS files (i.e. "C:\Dir Las\LAS1.las"). It is further suggested to manually define the boundaries of the area of interest when multiples input files are loaded (use command -userLimits explained at Table A.3). Otherwise the boundaries of the first LAS file loaded are used, which may lose data from subsequent input files. Furthermore DASOS only supports LAS1.3 with waveform packet format 4.
-pw <file1> <file2> ... <fileN>	loads a number of pulsedwave files (*.pls). Same rules apply as the -las tag
-volume <file>	loads an exported volume, generated using DASOS, instead of reading a LAS or pulsedwave file.
-vols <dir>	loads all the exported volume that are inside the given directory "dir". This option must and only be used for generating feature vectors (Section A.4.4.3).

Table A.1: DASOS fundamental file inputs

Tags	Description
-igm <igmFileName>	The name/directory of the .igm file that defines the geolocation of the hyperspectral pixels. This file is an ENVI bil with latitude and longitude per pixel.
-bil <bilFileName>	The name/directory of the .bil file that contains the hyperspectral cube.
-fodis <fodisFile>	The name/directory of the fodis (upward looking illumination sensor) .bil file for hyperspectral imagery
-dtm <dtmFileName>	loads a pre-calculated DTM and subtracts it from the position of each waveform sample before importing it to the volume. Please note that the DTM file format must be .bil and saved into float pointing numbers. Potential further file format limitations may exist. This is optional.
-csv <field-plots.csv>	The input csv file that lists all the trees from a number of field-plot. This is a compulsory input for generating feature vectors;

Table A.2: Optional or output request dependant.

A.4.3 General Parameters

All the general parameters has been pre-defined and they are therefore optional. Nevertheless, parameters are advised to be adjusted for each project. Table A.3 contains information about all the parameters and how to modify the voxelised FW LiDAR data during construction. Figures A-1 and A-2 show how the results are affected when these parameters are modified.

Tags	Description
-userLimits <maxNorthY> <minNorthY> <maxEastX> <minEastX>	User define boundaries of the area of interest. If not defined then the boundaries of the first file loaded are used (as defined in the header).
-vl <voxel- Length>	The voxel length controls the resolution of the output; the bigger the voxel length is the lower the resolution and the number of cubes are. Default value is 2.5m
-nl <noise- Level>	The noise level is the threshold of the low level filtering applied during voxelisation. Default value is 25. Please note that the intensity of each wave sample is not transformed to volts. Additionally, it is recommended to use the -exportPulses tag to export the amplitude of a few pulses and use those as sample data to define an appropriate noise threshold.
-iso <isolevel>	The iso-level is the intensity boundary that defines whether a voxel is empty or not. This is mostly used during polygonisation. The default value is zero. Please note that noise level and isosurface level are closely related but only the isolevel can be modified from an exported volume.

Table A.3: Description of the all the available tags that customises voxelisation of the FW LiDAR data.

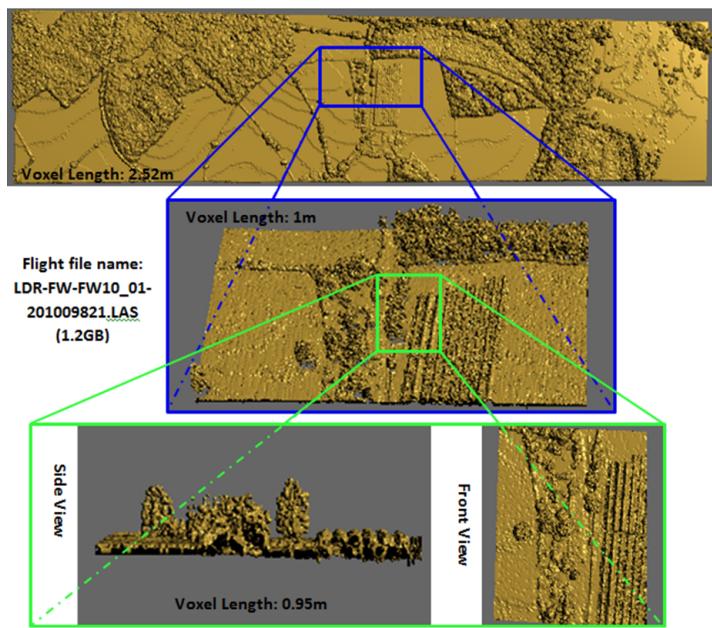


Figure A-1: Selecting Region of Interest

Voxel Length	Visualisation with different voxel lengths	Isolevel	Visualisations with various isolevels	Noise Level	Visualisations with various noise levels
10.0 m		45		5	
5.7 m		15		15	
4.44 m		-45		17	
1.43 m		-60		30	
1.0m		-85		75	
0.67 m		-100		135	

Figure A-2: Effect of modifying the user defined parameters; voxel length, isolevel and noise level.

A.4.4 Outputs

DASOS has three main outputs and three supplementary. At least one of them must be requested for the program to run.

The main outputs are the following:

1. **Polygonal Meshes**: exported into an .obj which is a standard graphics format that stores the vertices, edges and faces of the polygon. An image is also exported if hyperspectral data are loaded. (Section A.4.4.1)
2. **2D Metrics**: information about the scanned area in .asc format. If hyperspectral Images are loaded then aligned metrics from both datasets are available. (Section A.4.4.2)
3. **List of Feature Vectors**: exported into .csv files. Each row of the spreadsheet contains information about a local 3D cylideral or cubic area. (Section A.4.4.3)

The three suplementary outputs are explained in Table A.4 while the main outputs are explained in Sections A.4.4.1, A.4.4.2 and A.4.4.3 respectively.

Tags	Description
--help	It prints a list with all the available commands along with their description.
-exportPulses <noOfPulses> <fileName.csv>	Method that exports a number of pulses into a .csv file. The input <noOfPulses> the number of sample pulses to be exported into the <fileName.csv> file. It is used for deciding the noise level threshold for each project.
-exportVolume c <volumeFileName>	Exports the volume into an ASCII file to speed up future interpolation of the data. 'c' refers to compressed and it's an implicit functionality. If 'c' is not included then a non compressed file is exported, which sometimes is too big to be read back into DASOS. Therefore 'c' should always be included.

Table A.4: The suplementary ouput options of DASOS.

A.4.4.1 Polygonal Meshes

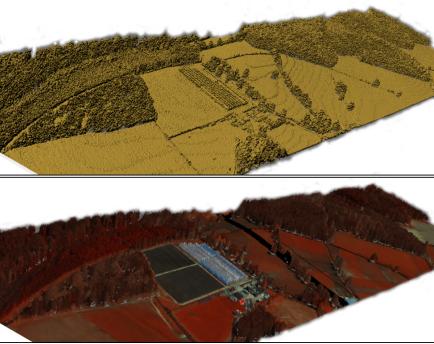
1st Main Output: 3D Polygon Mesh Generation		
Tags	Description	Output Example
-obj <objFileName>	The input <objFileName> is the name of the .obj file where the polygon representation of the LiDAR file will be exported to. A texture is also exported when hyperspectral images are loaded.	

Table A.5: Description of generating polygonal meshes and example outputs

A.4.4.2 2D Metrics Aligned with Hyperspectral Imagery

2nd Main Output: Generation of 2D metrics aligned with hyperspectral imagery	
Tags	Description
-map <type> <outputName>	<p>The available types are the following. Full description of each option is given in Table A.7 along with output examples.</p> <ul style="list-style-type: none"> • HEIGHT • THICKNESS • DENSITY • FIRST_PATCH • LAST_PATCH • AVERAGE_HEIGHT_DIFFERENCE • LOWEST_RETURN • INTENSITY_MAX • INTENSITY_AVG • HYPERSPECTRAL_MEAN • NDVI • ALL_FW <p>All the maps are exported into .asc format and can be loaded into QGIS and other software packages. The ALL_FW option generates one metric for each available full-waveform LiDAR related metric and their names are: outputName+metricsType+.asc Table A.7 explains what each metric is and gives output examples.</p>
-map HYPERSPECTRAL <band> <outputName>	The hyperspectral map needs an extra parameter defining which band will be output.

Table A.6: DASOS ouput options

Metric Description	Example
HEIGHT (DEM): The distance between the top non-empty voxel and the lower boundaries of the volume.	
THICKNESS: The distance between the first and last non empty voxels in every column of the 3D volume.	
DENSITY: Number of non-empty voxel over all voxels within the range from the first to last non-empty voxels.	
FIRST_PATCH: The number of non-empty adjacent voxels, starting from the first/top non-empty voxel in that column.	
LAST_PATCH: The number of non-empty adjacent voxels, starting from the last/lower non-empty voxel in that column.	
AVERAGE_HEIGHT_DIFFERENCE: An edge detection algorithm.	
LOWEST_RETURN The height of the lowest non empty voxel (the actual heights are very low and close to each other but the example image has been scaled and the difference seems bigger)	

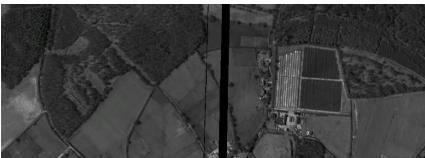
INTENSITY_MAX The maximum intensity of each column	
INTENSITY_AVG The average intensity per column	
HYPERSPECTRAL_MEAN The mean of the hyperspectral spectrum.	
NDVI The Normalised Difference Vegetation Index indicates whether green vegetation exists or not and it is derived from the electromagnetic spectrum as follow:	
$NDVI = \frac{NIR - VIS}{NIR + VIS} \quad (A.1)$ <p>where the NIR is the near-infrared region of the spectrum (700-2500nm) and VIS is the Visible/Visual spectrum (430-770) [38].</p>	
HYPERSPECTRAL A single user defined hyperspectral band.	

Table A.7: Description of generating polygonal meshes and example outputs

A.4.4.3 List of Feature Vectors

This is useful for characterising object inside the 3D space (e.g. trees). For each column of the voxelised FW LiDAR, information around its local area are exported.

Similar to the previous functionalities of DASOS, the program requires <inputs> <parameters> and <outputs>. Those requirements are described in Tables A.8, A.9 and A.10 respectively. Please note that these inputs are also described with the rest of the inputs in Section A.4.2.

3rd Main Output: List of Feature Vectors - Inputs	
Tags	Description
-vols <volDir>	the directory of the volume of interest generated beforehand.
-icsv <fieldplots.csv>	the input csv file that contains all information about the field-plots.

Table A.8: Explanation of how to define the two compulsory inputs to get the 3rd main output of DASOS

Figure A-3 shows an example of a file with fieldplots. A file may contain multiple fieldplots, but it has to have at least 6 columns: the 3 columns define the fieldplot (northing, easting and radius) and the rest give information about the trees (northing, easting and class). The order of the columns has no significance. Figure A-3 shows an example. The labels of the those columns could vary and can be defined as explained in Table A.9.

IsDead	Northing	Easting	X	Y	RADIUS
Live	60	70	55	75	40
Live	60	70	75	85	40
Dead	60	70	65	55	40
Live	60	60	20	60	40

Figure A-3: Example of fieldplot input

Additionally, the size and shape of the investigated area from where the features are extracted is user defined and Table A.9 lists all the related, modifiable parameters.

3rd Main Output: List of Feature Vectors - Parameters	
Tags	Description
-column <label>	the label of the column that defines the class of each entry (e.g. <label> = isDead)
-class <className or ALL>	the name of the class (e.g. dead or alive) of interest or ALL. If a class is chosen, then only the columns that contain a tree of that class are taken into consideration; a feature list is exported for each tree that belongs to this class only. The ALL option is the area of interest and generates a template for each column that lies inside the voxelised space.
-ttype square <x> <y> <z>	generates a feature vector derived from a cuboid area of size x, y, z voxels. The systems finds the first non empty voxel starting from the top of the column. By default it moves one voxel upwards and sets that to be the top of the cuboid/cylinder. It is highly recommended to use odd numbers, otherwise the centre of the cuboid/cylinder will be wrongly set and unpredicted output values may occur.
-ttype cylinder <h> <r>	generates a cylindrical template with height h and diameter $(2 \times r + 1)$ voxels and height h. The systems finds the first non empty voxel starting from the top of the column. By default it moves one voxel upwards and sets that to be the top of the cuboid/cylinder.
-mheight <n>	moves the template into the y-axis n voxels upwards instead of one which is the default. The value n must be a positive number.
-eparameters <raw or processed>	the ‘raw’ option saves all the intensity values of the template and the ‘processed’ option saves parameters derived from the raw intensities. Table A.11 explains how each processed parameters is derived.

Table A.9: Explanation on how to Modify the Parameters of the 3rd Main Output of DASOS

3rd Main Output: List of Feature Vectors - Outputs	
Tags	Description
-ocsv <nameStart>	For each .vol file found in the given directory (using -vols), a csv file is exported. The name of each file exported is: <nameStart> + <volFileName> + ".csv" and it contains the list of the feature vectors generated from the corresponding volume

Table A.10: Explanation of the tag that exports the list of feature vectors

Figure A-4 shows examples of two exported list of feature vectors: one with processed parameters and one with raw intensities. In each .csv file exported, each line is a feature vector. The first column is its ID as it defined during run time. The second and third columns define the centroids of each investigated local area (cuboid/cylinder). The other columns contain either processed or raw parameters. If they are processed, then information like mean height and standard deviation of heights are listed. Table A.11 is a full list of all the proccesed parameters. If the parameters are raw, then the corresponding voxel intensity values are exported. The label of each voxel is "v_x_y_z", where "v_0_0_0" is the lower voxel of the cuboid/cylinder and it has the minimum easting and northing it as well.

Index	centroid_x	centroid_y	Height_Middle_Column	Height_Mean	Height_Median	Height_Std	Sum_Int	Diff_X	...
0	251836.109	6048994.5	36	35.5	36	0.943	95.125	...	
1	251843.906	6048980.5	19.8	20.1	20.4	0.671	0	...	
2	251846.312	6048979	16.8	16	15.6	1.02	169.167	...	
3	251849.312	6049022.5	36	35.7	36.6	0.964	169.278	...	
4	251851.703	6048988	17.4	16.2	16.2	0.346	408.065	...	
5	251852.906	6048975	27	26.4	26.4	0.917	68.537	...	
6	251857.109	6048974	17.4	17.4	18	0.849	162.25	...	
7	251858.312	6049010.5	40.8	40	39.6	1.02	251.36	...	
8	251860.703	6048984	17.4	16.6	16.2	0.663	67.883	...	
9	251861.312	6049000	19.8	20.1	20.4	0.671	0	...	

Index	centroid_x	centroid_y	v0_0_0	v0_0_1	v0_0_2	v0_0_3	v0_0_4	v0_1_0	v0_1_1	v0_1_2	v0_1_3	...
0	251836.109	6048994.5	7	14	10	26	0	0	9	10.25	11.875	...
1	251843.906	6048980.5	0	0	0	0	0	0	0	0	0	...
2	251846.312	6048979	9	60.75	70.75	13	8	0	0	0	7.667	...
3	251849.312	6049022.5	48.556	93.222	20.5	0	7	0	0	0	0	...
4	251851.703	6048988	100.2	53.222	10.5	7.143	0	0	0	0	47.25	...
5	251852.906	6048975	0	0	0	0	0	26.875	0	10.444	13.182	...
6	251857.109	6048974	0	0	0	0	0	45.667	93	16.333	7.25	...
7	251858.312	6049010.5	0	0	0	0	0	0	0	8	6	...
8	251860.703	6048984	0	45.75	8	7.333	0	0	0	0	6.8	...
9	251861.312	6049000	0	0	0	0	0	0	0	0	0	...

Figure A-4: Example of .csv files with a list of feature vectors exported.

Explanation of the List of Feature Vectors Output with the Processed Intensities	
Label	Description
Height_Middle_Column	The height of the middle column of the cuboid/cylinder
Height_Mean	The Mean height of all the columns included in the template
Height_Median	The Median height of all the columns included in the template
Height_Std	The Standard Deviation of the heights of the columns included in the template
Sum_Int_Diff_X	The Mirror Summed Difference of the intensities using the middle column in the x-axis as the axis of symmetry
Sum_Int_Diff_Y	The Mirror Summed Difference of the intensities using the middle column in the y-axis as the axis of symmetry
Sum_Int_Diff_Z	The Mirror Summed Difference of the intensities using the middle column in the z-axis as the axis of symmetry
Max_Int	The maximum intensity found inside the cuboid/cylinder
Min_Int	The minimum intensity found inside the cuboid/cylinder
Ave_Int	The average intensity of the voxels that contain an intensity above the isolevel
Median_Int	The median intensity of the voxels
Per_Int_Above_Iso	Percentage of voxels that contain an intensity above the isolevel
Dis_Mean	Mean distance from the central voxel to every voxel that contain san intensity above the isolevel

Dis_Median	Median distance from the central voxel to every voxel that contains an intensity above the isolevel
Dis_Std	The Standard Deviation of the distances between the central voxel and every voxel that contains an intensity above the isolevel
Top_Patch_Len_Middle_Co	The length of the top patch of the middle column of the cuboid/cylinder
Top_Patch_Len_Mean	The Mean length of all the top patches
Top_Patch_Len_Median	The Median length of all the top patches
Top_Patch_Len_Std	The Standard Deviation of all the top patches
Mirror_Diff_X_Mean	The Mean Mirror Difference of the voxel intensities with the middle column of the x-axis as the symmetric axis
Mirror_Diff_X_Median	The Median Mirror Difference of the voxel intensities with the middle column of the x-axis as the symmetric axis
Mirror_Diff_X_Std	The Standard Deviation Mirror Difference of the voxel intensities with the middle column of the x-axis as the symmetric axis
Mirror_Diff_Y_Mean	The Mean Mirror Difference of the voxel intensities with the middle column of the y-axis as the symmetric axis
Mirror_Diff_Y_Median	The Median Mirror Difference of the voxel intensities with the middle column of the y-axis as the symmetric axis
Mirror_Diff_Y_Std	The Standard Deviation Mirror Difference of the voxel intensities with the middle column of the y-axis as the symmetric axis
Mirror_Diff_Z_Mean	The Mean Mirror Difference of the voxel intensities with the middle column of the z-axis as the symmetric axis

Mirror_Diff_Z_Median	The Median Mirror Difference of the voxel intensities with the middle column of the z-axis as the symmetric axis
Mirror_Diff_Z_Std	The Standard Deviation of the Mirror Difference of the voxel intensities with the middle column of the z-axis as the symmetric axis

Table A.11: Explanation of the processed parameter exported within a feature vector

A.5 Exercises

A.5.1 Sample Data

These exercises will give you an in depth understanding of DASOS, while working with real examples. At first, copy the folder "DASOS_userGuide" into your C:\ drive. This folder is available to download from <https://github.com/Art-n-MathS/DASOS/tree/master/DASOS_win>. To ease typing, all the example commands are given in the ExerciseCommands.bat file, which can be opened in a text editor.

There are three datasets provided for the following exercises and they are available at: <<https://www.dropbox.com/sh/hzpl16gue5xvjmb/AADQsJ0sqKkx01CX4mJjvBPVa?dl=0>> and <<https://plymouthmarinelaboratory.webex.com/plymouthmarinelaboratory/j.php?MTID=m305f59dda16e653b2946c6a3b00e93f4>>. Please copy the data inside the directory <DASOS/DASOS_win/SampleDATA> check that the following files are included:

1. 1st sample dataset inside < C:\DASOS_userGuide\SampleDATA\DATASET_1>:
 - (a) LDR-FW-FW10_01-201009821.LAS
 - (b) e098211b_FODIS.bil
 - (c) e098211b_FODIS.bil.hdr
 - (d) e098211b_masked.bil
 - (e) e098211b_masked.bil.hdr
 - (f) e098211b_osgn.igm
 - (g) e098211b_osgn.igm.hdr
 - (h) Readme.txt

2. 2nd sample dataset inside < C:\DASOS_userGuide\SampleDATA\DATASET_2>

- (a) *Australia_1.pls*
- (b) *Australia_1.wvs*
- (c) *Australia_1_dtm.bil*
- (d) *Australia_1_dtm.hdr*
- (e) *Australia_2.las*
- (f) *Australia_2.wdp*
- (g) *Australia_2_dtm.bil*
- (h) *Australia_2_dtm.hdr*
- (i) *Australia_3.las*
- (j) *Australia_3.wdp*

3. 3rd sample dataset inside < C:\DASOS_userGuide\SampleDATA\DATASET_3>

- (a) *myTestVol_.vol*
- (b) *myTestVol_flat.vol*
- (c) *testFieldplot.csv*

Information about data usage and related license are given in Section A.2

Once all the files are copied across, open the command Prompt and type:

```
$: cd C:\DASOS_userGuide\DATASOS
```

This will bring you to our working directory. In case you are using a different directory then go to your work directory inside the folder DATASOS and the rest of the commands should work OK.

A full guide of all the available tags is given with the following command.

```
$: DATASOS --help
```

The same information can be found inside the Readme.txt file and this User Guide (Section A.4).

A.5.2 Exercises

A.5.2.1 Deciding Noise Threshold

The following examples export the amplitudes of 12 pulses into a .csv file to help us decide what noise threshold to use.

```
$:DATASOS -las ..\SampleDATA\DATASET_1\LDR-FW-FW10_01-201009821.LAS
```

```
-exportPulses 12 ..\LAS21pulsesSamples.csv

$: DASOS -las ..\SampleDATA\DATASET_2\Australia_2.las -exportPulses 12
..\Australia_2_pulsesSamples.csv
```

A.5.2.2 Exporting metrics from DASOS

The following commands export a height map into .asc files. These files can be used in QGIS. This will give us the location of the flightlines and the relation between them.

```
$: DASOS -las ..\SampleDATA\DATASET_2\Australia_2.las -nl 6 -vl 2 -map
height ..\Australia_2_vl2_height

$: DASOS -las ..\SampleDATA\DATASET_2\Australia_3.las -nl 6 -vl 2 -map
height ..\Australia_3_vl2_height
```

Generating a single map at the beginning is useful for deciding which flightlines lie inside the area of interest.

A.5.2.3 Loading Multiple Flightlines

As mentioned before, for loading multiple flightlines it is suggested to manually define the boundaries of the area of interest. The following command loads two flightlines, generates a volume from the area of interest and exports it into the Australia2-3.vol file.

```
$: DASOS -las ..\SampleDATA\DATASET_2\Australia_3.las
..\SampleDATA\DATASET_2\Australia_2.las -nl 6 -vl 2 -iso 4 -userLimits
6199990 6199639 762405 761951 -exportVolume c ..\Australia2-3.vol
```

A.5.2.4 Exporting Metrics

The following command loads the pre-computed volume and creates a height map and all the FW related metrics. Please note that height is also a FW related metric, therefore it will be created twice.

```
$: DASOS -volume ..\Australia2-3.vol -map height ..\Australia2-3 -map
all_fw ..\Australia2-3
```

A.5.2.5 Subtracting Pre-computed Digital Terrain Model

The next command loads two LAS files, a pre-computed DTM file is subtracted from the wave samples' positions while the volume is created, the volume is exported into the *Australia2-3_dtm.vol* file and finally it exports a height metric.

Please note that when a DTM is introduced, a new volume must be created. Since the volumetric files are raster data and contain no information about pulses.

```
$: DASOS -las ..\SampleDATA\DATASET_2\Australia_3.las
.. \SampleDATA\DATASET_2\Australia_2.las -dtm
..\SampleDATA\DATASET_2\Australia_2_dtm.bil -nl 6 -vl 2 -iso 4
-userLimits 6199990 6199639 762405 761951 -exportVolume
c ..\Australia2-3_dtm.vol -map height ..\Australia2-3_vl2_dtm_height
```

You may then use the same volume to export more metrics:

```
$: DASOS -volume ..\Australia2-3_dtm.vol -map AVERAGE_HEIGHT_DIFFERENCE
..\Australia2-3_dtm_AVG_height_diff
```

A.5.2.6 Pulsewave Data

As mentioned before, it is suggested to first export the amplitudes of a few pulses to decide on an appropriate noise threshold.

```
$: DASOS -pw ..\SampleDATA\DATASET_2\Australia_1.pls -exportPulses 15
..\PLS_amplitudeSamples.csv
```

And then you can generate the desired metrics:

```
$: DASOS -pw ..\SampleDATA\DATASET_2\Australia_1.pls -nl 5 -dtm
..\SampleDATA\DATASET_2\Australia_1_DTM_1m.bil -vl 3 -map thickness
PLS_vl3_thickness -exportVolume ..\Australia_1_vl3_dtm.vol
```

A.5.2.7 Polygon Representation

DASOS create 3D polygon representation using the '-obj' tag. The 3D polygon representations are exported into .obj format, which can be visualised using animation software packages. For this workshop we are using Meshlab because it is a free tool and it can handle millions of triangles.

Meshlab is available to download from here: <<http://meshlab.sourceforge.net/>> and it is also included into our working directory "DASOS_userGuide".

An example of generating polygons is given below:

```
$: DASOS -las ..\SampleDATA\DATASET_1\LDR-FW-FW10_01-201009821.LAS -nl 20  
-vl 1.7 -obj ..\LAS21.obj -exportVolume c ..\LAS21_vl1.7.vol
```

The generated volume is also saved because we need it for the following exercises.

A.5.2.8 Hyperspectral Imagery

One of the key functionalities of DASOS is the alignment with the hyperspectral imagery. DASOS can export 3D coloured polygon representations and aligned metrics between FW LiDAR and hyperspectral data.

For the 3D coloured polygon representations you must not use any directory for the exported .obj file Analysisname because the link between the texture and the .obj file will not work. Here is an example:

```
$: DASOS -volume ..\LAS21_vl1.7.vol -bil  
..\SampleDATA\DATASET_1\e098211b_masked.bil -igm  
..\SampleDATA\DATASET_1\e098211b_osgn.igm -fodis  
..\SampleDATA\DATASET_1\e098211b_FODIS.bil -rgb 240 78 23 -obj  
LAS21_coloured.obj
```

The LAS21.obj file will be saved into the current directory, which in our case is:
C:\DASOS_userGuide\Dasos.

Please note that the following command should give the same results, but as mentioned before importing an exported volume is faster than generating from scratch.

```
$: DASOS -las ..\SampleDATA\DATASET_1\LDR-FW-FW10_01-201009821.LAS -nl 20  
-vl 1.7 -bil ..\SampleDATA\DATASET_1\e098211b_masked.bil -igm  
..\SampleDATA\DATASET_1\e098211b_osgn.igm -fodis  
..\SampleDATA\DATASET_1\e098211b_FODIS.bil -rgb 240 78 23 -obj  
LAS21_coloured.obj
```

An example of generating aligned metrics is given below. The NDVI map is quite slow, so we may need to wait a bit for that.

```
$: DASOS -volume ..\LAS21_vl1.7.vol -bil  
..\SampleDATA\DATASET_1\e098211b_masked.bil -igm  
..\SampleDATA\DATASET_1\e098211b_osgn.igm -fodis  
..\SampleDATA\DATASET_1\e098211b_FODIS.bil -map hyperspectral 140  
..\LAS21_band140 -map height ..\LAS21_height -map NDVI ..\LAS21_ndvi
```

A.5.2.9 All Commands Together

Of course, we are able to use multiple outputs into a single command, even though that's not recommended due to the long processing time. An example of merging previous commands into one is given below:

```
$: DASOS -las ..\SampleDATA\DATASET_1\LDR-FW-FW10_01-201009821.LAS -nl 20  
-vl 1.7 -bil ..\SampleDATA\DATASET_1\e098211b_masked.bil -igm  
..\SampleDATA\DATASET_1\e098211b_osgn.igm -fodis  
..\SampleDATA\DATASET_1\e098211b_FODIS.bil -rgb 240 78 23 -obj  
LAS21_coloured.obj -map  
hyperspectral 140 ..\LAS21_band140 -map height ..\LAS21_height -map  
NDVI ..\LAS21_ndvi -exportVolume ..\LAS21_v11.7.vol
```

A.5.3 Exporting feature vectors from exported voxelised FW LiDAR

This examples takes as input two test .vol files and a fieldplot file. The file named "myTestVol_flat.vol" contains a flat surface, while inside the "myTestVol_.vol" the middle column of the first dead tree that is defined inside the "testFieldplot.csv" is one voxel higher. The covered area of the two .vol files is identical and for that reason the fieldplot circle lies inside both files. The following command produces a list of vectors with features derived after processing the voxel intensities of the cuboids that contain dead trees according to the input field data:

```
$: DASOS -vols ..\SampleDATA\DATASET\_3 -ic平  
..\SampleDATA\DATASET_3\testFieldplot.csv -eparameters processed -column  
isDead -class dead -ttype square 3 3 5 -oc平 templatesProcessedCuboid
```

The following command produces a list of vectors with the voxel intensities of cylinders that contain dead trees according to the input field data:

```
$: DASOS -vols ..\SampleDATA\DATASET_3 -ic平  
..\SampleDATA\DATASET_3\testFieldplot.csv -eparameters raw -column  
isDead -class ALL -ttype cylinder 5 3 -oc平 templatesALLRawCylinder
```

A.6 Limitations

Limitation and bugs have been reported throughout the report, but here is a short summary of them.

- Exporting polygon representation could end up generating a bunch of cones instead of a nice smooth surface.
- Subtracting DTM depends on the input file format and, by subtracting the height, the input data may end up outside the boundary of the volume.
- DASOS may be not be perfectly portable to all systems as development and testing was done on two computers only.
- The raw waveform amplitude is used as intensity and it hasn't been converted to an absolute digitizer voltage, for the LiDAR systems where these raw values are scaled.
- Intensities also have not been calibrated.
- Sometimes memory allocation exceptions occur.

For full bug reports and under development improvements please check the following link:

https://docs.google.com/spreadsheets/d/10yE5p463cLA_GtKkyiaWEzScW7N9cVxbPs5y0muXuZY/edit?usp=sharing

A.7 Related Forums and Social Media

Online social media are used for sharing DASOS updates and discussing issues or potential improvements. Information about DASOS can be found in the following:

- Google Groups: DASOS - the native full-waveform (FW) LiDAR software
<https://groups.google.com/forum/#forum/dasos---the-native-full-waveform-fw-lidar>
This group is used for bringing potential issues and possible improvements up in discussion.
- Blogger: ART & M@thS
<http://miltomiltiadou.blogspot.co.nz/2015/03/las13vis.html>
This blog is more general. The blog contains updates and explanation of DASOS but usually the code used in DASOS is broken down into small projects and explained how they can be used in other applications.
- Twitter: @Miltomiltiadou
Milto Miltiadou's twitter, where all the updates and news of DASOS are published.

Appendix B

Case Study: Field Work in New Forest

B.1 Introduction

This section is a case study containing field work from a [non-forester perspective](#) to better understand the challenges of working remotely with forests. Remotely sensed data contain a great amount of information but in order to build a good system for identifying trees and materials, an in depth knowledge of them is required [39]. For that reason, this case study was created; information about the New Forest, which is a forest in the south of United Kingdom, were collected and a small validation dataset was created. The dataset created includes the tree species and approximate heights of the trees in two areas of interest.

Before travelling to the New Forest, two areas of interest were selected. These areas were selected according to the following criteria:

- There were LiDAR data of the selected area to be able to compare what we can see on the ground with the scanned data
- Areas that had a variation of tree species were selected. This was done according to the (non-validated) results of a thesis of Bournemouth University that classified the tree species of the New Forest [40]. This helped get a broader range of tree species.

The following sections give a detailed description of the information gathered during the trip. This includes the species and height maps generated, the different types of landscapes found and the challenges faced.

B.2 Validation Data Collected

The tree classes were initially defined by the provided Bournemouth thesis [40]. A colour was chosen for each tree class and, while being in the New Forest, the aim was to mark each tree on the paper map with the corresponding colour. Using QGIS (Quantum Geographic Information System) the classification results of the forest assessment, undertaken by Sumnall in 2013 [40], were coloured with the same colours to ease comparison.

At the aforementioned forest assessment, there were 26 classes from 14 different species; the remaining 12 classes were young versions of the 14 species. Here the classes are reduced to 14 by merging all the young trees into the tree species classes (in the 4 years gap between the 2010 assessment and the visit to new Forest in 2014, the young trees would have aged). See table B.1 for the initial 14 classes. Nevertheless, more tree species existed in the areas of interest in New Forest than those 14 classes. The colours and symbols of the extra tree classes are shown on table B.2.

Tree	Colour
1. Beech	Yellow
2. Oak	Orange
3. Silver Birch	Light Brown
4. Sweet Chestnut	Brown
5. Corsican Pine	Red
6. Coast Redwood	Pink
7. Douglas Fir	Purple
8. Grand Fir	Light Purple
9. Japanese Larch	Cyan
10. Lawson Cypress	Grey
11. Norway Spruce	Blue
12. Scots Pine	Green
13. Western Hemlock	Brown
14. Common Adler	Dark Brown

Table B.1: Colours of the initial 14 classes

During the visit, tree species maps were generated for a few square meters. The position of the trees were found relative to easily-spotted reference points (e.g. road crossing) that were marked in advance. That was done because, according to Dr. Ross Hill, no GPS can be accurate enough when trees are around since the satellite signal bounces off the leaves and reduces the positioning accuracy. In professional fieldwork, a total station is used but, for the purposes of this visit, it was not considered necessary. By the end of the case study, ground maps were coloured according to the tree species

Tree	Colour / Symbols
15. Ash	A
16. Hawthorn	Blue pen colour
17. <u>Malus (Crabapple)</u>	Highlighter
18. Holly Tree	
19. Trees that have been cut down	x
20. Trees that are mixed together	// (added on top of the normal colour)

Table B.2: Classes that were added during the trip

identified and estimates of the approximate heights of the trees were also noted down.

The following four maps were created for each selected area. The first two maps were created before the trip during preparation, while the last two contain the information collected during the field work.

- a screen shot of the area from Google map,
- the classification results from the forest assessment [40],
- the coloured tree species map and
- the approximated height map.

Comparing the validation dataset created with the classifications done at Bournemouth University (which were not validated), it is clear there are misclassifications. This is shown in Figures B-1 and B-2 and it is likely that occurs due to the over-segmentation of trees. Those wrong classifications justify that validation and field work data are essential for building a good classifier.

The first area is included in the LAS file named LDR-FW-FW10_01-201018715.LAS and it lies inside the limits: X = (433453 - 433761), Y = (102193 - 102405) [British National Grid coordinates]. The four maps that relate to these areas are shown in Figure B-1.

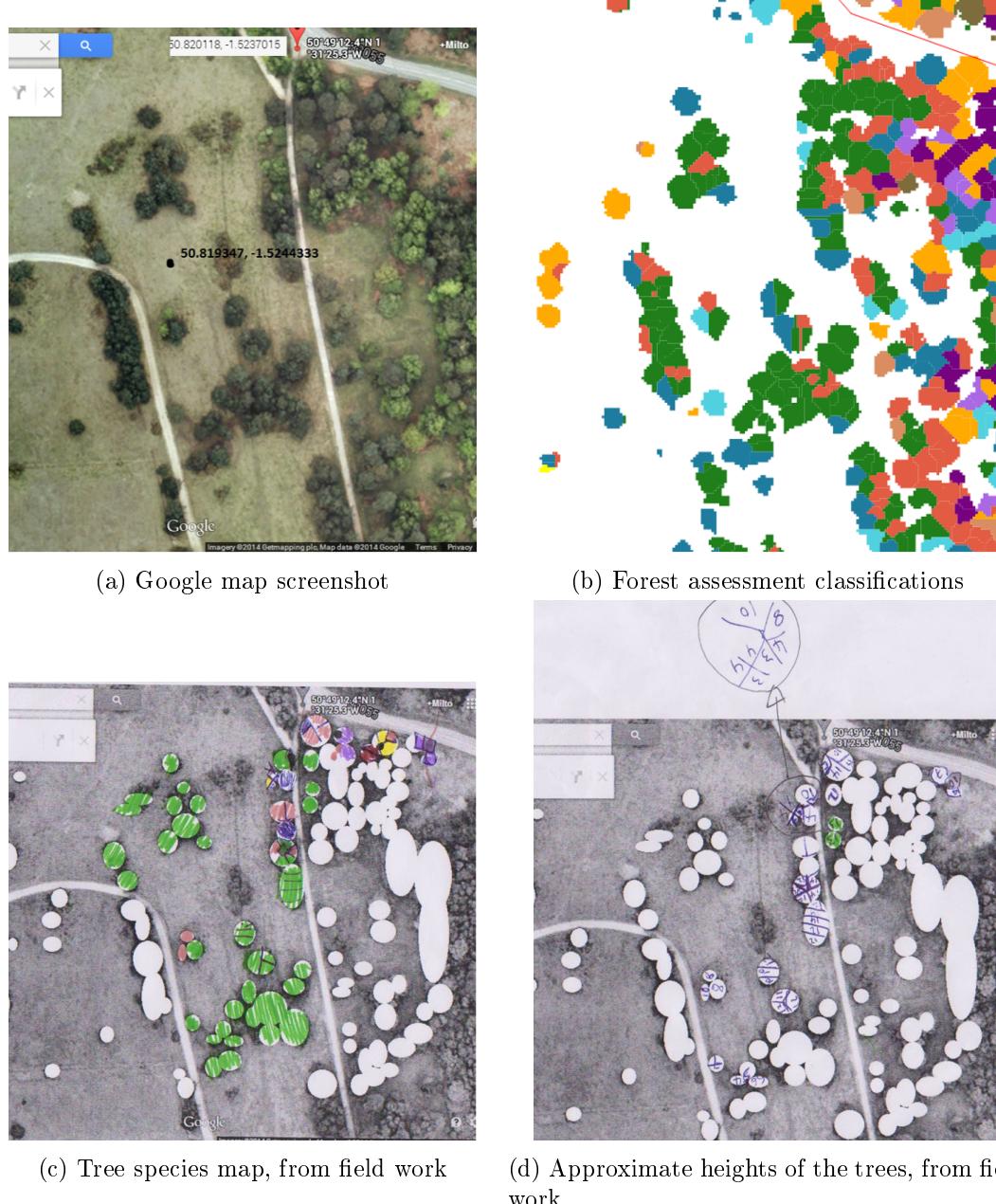
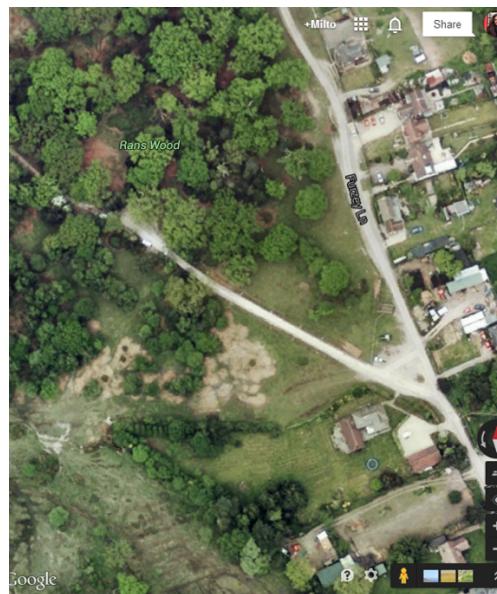
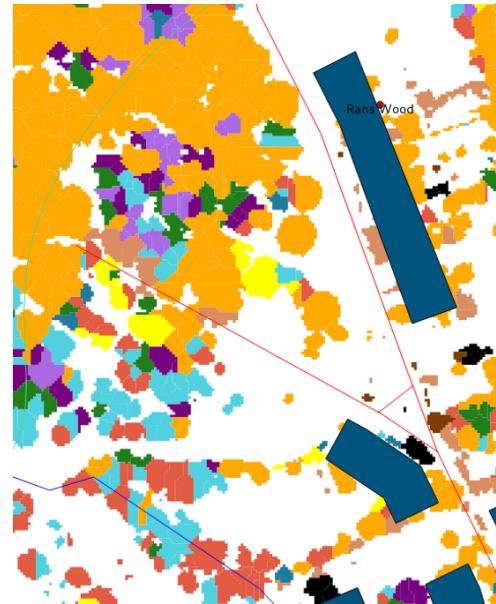


Figure B-1: The first area of interest and the related maps.

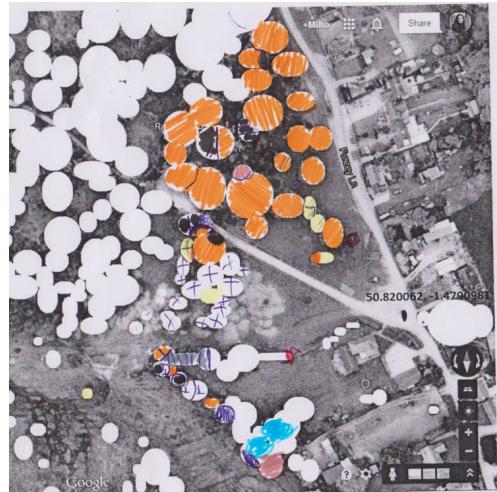
The second area is included in the LAS files named LDR-FW-FW10_01-201018719.LAS and LDR-FW-FW10_01-201018718.LAS and it lies inside the limits: X = (436442 - 436835), Y = (102334 - 102585) [British National Grid coordinates]. The four maps created for these areas are shown in Figure B-2.



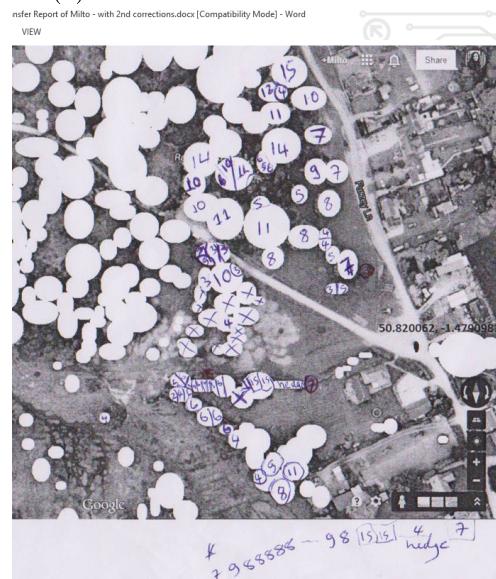
(a) Google map screenshot



(b) Forest assessment classifications



(c) Tree species map, from field work



(d) Approximate heights of the trees, from field work

Figure B-2: The second area of interest and the related maps.

B.3 Landscape types

During the forest assessment in New Forest, not only validation data were collected, but also useful information about classifying the data. The following images show examples of the five landscape types that were found in New Forest:

1. Heather fields:



Figure B-3: Trees that have been cut down

2. Grass with a few scattered trees:



Figure B-4: Grass with a few scattered trees

3. Dense Forest:



Figure B-5: Dense forest

4. Bushes and Shrubs



Figure B-6: Trees that have been cut down

5. Lakes and rivers, which are more rarely found



Figure B-7: Lakes and rivers

Please note that the landscape types could significantly differ according to the scanned area. For example, the landscape of New Forest is flat while the landscape of Eaves Wood (another scanned forest in UK) is hilly. The landscape type should be taken into consideration during classifications.

B.3.1 Classification challenges

This case study brought further understanding of the challenges of creating validation data and writing a tree species classifier. These challenges are listed and explained below with some photos taken during field work:

1. Field work and remotely sensed data collection should happen around the same time to avoid changes that happens over time. In the New Forest case, the airborne data were collected in 2010 and many changes occurred in the intervening time - in the most extreme cases, some trees had been cut down.
2. Machine learning becomes more time consuming as the number of classes increases. Regarding tree species classes, it is unrealistic to expect that all tree species will be identified. This point is underlined by the fact that the list of tree species used



Figure B-8: Trees that have been cut down

in the tree assessment held by Sumnall [40] didn't include a number of trees (e.g. holly trees and crabapple) that were widespread in New Forest.

3. There is much more than just trees in the forest, including mobile animals, that may confuse a classification if LiDAR returns hit rocks, animals, vehicles or buildings instead of branches, leaves and trunks. **Any classification must account for inevitable errors due to background clutter that need to be invariant to.**



Figure B-9: Animals in New Forest

4. Large validation datasets from a single area will not be sufficient, because trees of the same species are usually gathered together. For instance, the first selected area has many pine trees while the second one has many oak trees. Therefore, it is important to have many field plots spread well within the area of interest.

5. Further, some trees are entwined together which makes it difficult to identify from the data whether they are one or two trees. Examples are shown in Figure B-10; in the left image, the trunks of the two trees are very close to each other and, in the right image, a crabapple and an oak tree have grown together.



Figure B-10: Trees, which are mixed together

B.4 Conclusions and Discussion

To sum up, the trip to the New Forest was essential for better understanding the challenges of remote monitoring of forests. During the visit, a small validation dataset was generated; the species and height of trees that are inside the two areas of interest were noted down. Field work is a time consuming task and weeks are required for generating a big enough validation dataset, but it is essential for understanding the object of interest (trees) in relation to the scanned data. Challenges identified were also explained and this increased knowledge about forests should lead to implementing a better classifier.