The accuracy of image-based individual tree crown detection and delineation across vegetation types

Nicolas Pucino [1], Tim McVicar [2], Shaun Levick[3], Albert van Dijk[4]

[1] Australian National University, Fenner School of Environment and Society, Canberra, ACT, Australia - Nicolas.pucino@anu.edu.au
[2] CSIRO Environment, Canberra, ACT, Australia - tim.mcvicar@csiro.au
[3] CSIRO Space and Astronomy, Urrbrae, SA, Australia - shaun.levick@csiro.au
[4] Fenner School of Environment and Society, Australian National University, Canberra, ACT, Australia - albert.vandijk@anu.edu.au

Keywords: Canopy Height Model Evaluation, Airborne LiDAR Accuracy Assessment, Vegetation Mapping, Forest Analysis, Point Cloud

Abstract

Australia's terrestrial ecosystems are critical to the global carbon cycle, yet they face numerous environmental pressures such as forest degradation and biodiversity loss. Accurate monitoring of vegetation dynamics is crucial to mitigating these challenges and informing sustainable management strategies. Individual Tree Segmentation (ITS) methods, powered by deep learning, enable large-scale mapping of tree crowns, which is vital for assessing above-ground biomass and carbon stocks across vast landscapes. Despite their promise, inconsistencies in algorithmic performance arise due to varying vegetation types, point cloud densities, and dataset-specific characteristics, which limit the generalizability of supervised models.
This study evaluates the performance of different ITS and Canopy Height Model (CHM) algorithms for generating large tree crown datasets using LiDAR-derived data from across Australia. We applied these methods to 37 representative airborne LiDAR point clouds across 15 vegetation classes, representing a range of ecosystems from rangelands to tropical forests.
Our analysis reveals that the effectiveness of tree detection and crown delineation varies significantly across vegetation types and point cloud densities. The Pit-Free CHM algorithm generally outperforms others, yielding higher match rates in the delineation of tree crowns. Additionally, the DalPonte ITS algorithm provides the most accurate results, especially in sparsely vegetated areas such as rangelands, which are critical for mapping and monitoring. In contrast, closed-canopy forests present challenges, particularly due to crown clumping and multi-layered vegetation structures. This study highlights the importance of selecting the appropriate ITS and CHM methods for different vegetation types and emphasizes the need for algorithm optimization in complex environments, such as tropical and eucalypt forests. Ultimately, these findings provide valuable insights into enhancing large-scale vegetation monitoring and improving model generalization for tree crown detection.

1. Introduction

Australia's terrestrial ecosystems play a pivotal role in the global carbon cycle and are facing increasing environmental and anthropogenic pressures, including forest degradation, habitat fragmentation, and biodiversity loss (Lindenmayer, 2023). Accurate monitoring of vegetation dynamics is crucial to mitigate these challenges and inform sustainable management strategies.

Individual Tree Segmentation (ITS) methods enable high-resolution mapping of tree crowns. Deep learning (DL) models trained on individual tree crown delineations can segment billions of crowns from multispectral imagery across entire continents. This facilitates large-scale monitoring of landscape changes and provides critical insights into above-ground biomass and carbon stocks. Such capabilities are essential for evaluating national or continental carbon sequestration rates and assessing the effectiveness of carbon management strategies (Oehmcke et al., 2024). When coupled with canopy height models (CHM), these datasets facilitate the estimations of above-ground biomass and carbon stocks, potentially at regional to continental scales (Oehmcke et al., 2024).

However, the diversity of vegetation types, varying point cloud densities, and algorithmic approaches introduce inconsistencies that limit the generalizability of supervised models built on these delineations. Additionally, selecting the most appropriate ITS method for specific datasets and regions is vital (Wallace et al., 2014), as tree detection and crown delineation accuracy can vary significantly between methods. Even in similar environments with identical data, certain methods—particularly image-based ones—can yield different outcomes (Latella et al., 2021). Segmentation inaccuracies may arise from errors in canopy height models (Goldbergs et al., 2018) or the misidentification of non-tree vertical structures as trees (Weinstein et al., 2019), emphasizing the importance of carefully selecting an appropriate CHM algorithm.

Leveraging LiDAR-derived data from multiple locations across Australia, this study investigates the efficacy of various ITS methods for generating high-quality and large tree crowns datasets, potentially useful for training DL models that are generalizable and balanced towards the natural vegetation type distribution. We evaluate the detection and delineation accuracies of some of the most used CHM and ITS methods, focusing on their adaptability across diverse vegetation types and point cloud densities.

2. Methods

We analysed 37 representative airborne LiDAR-derived point clouds (see Figure 1) across 15 vegetation classes, selected to replicate the statistical distribution of Australian vegetation as defined by Scarth et al. (2019).

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-1223-2025

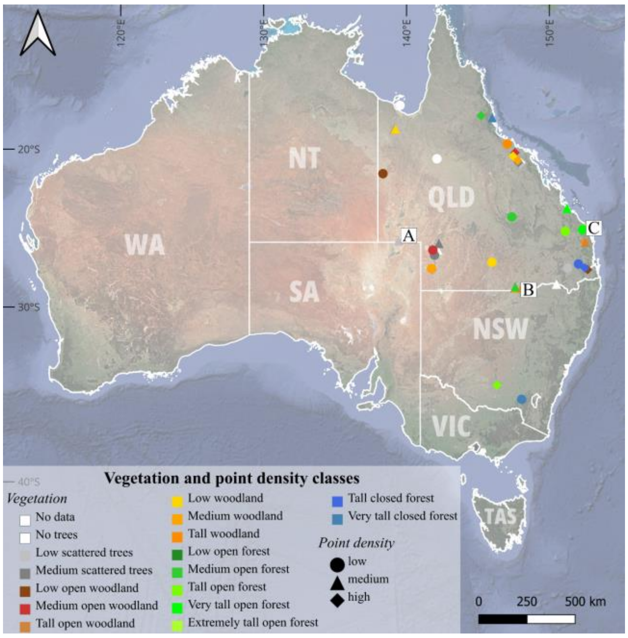

Figure. 1. Distribution of point cloud datasets across Australia, categorized by vegetation class (based on Scarth et al., 2019) and represented by symbols indicating point cloud density classes. Insets include red polygons denoting manually delineated tree crown ground truth, with background imagery sourced from Google and DigitalGlobe.

Covering approximately 7.7 million km², Australia spans latitudes from 10° S (Cape York, Queensland) to 43° S (South East Cape, Tasmania) and hosts a diverse range of biomes, such as grasslands, deserts, temperate forests, subtropical regions, tropical rainforests, and equatorial zones (Stern et al., 2000). Known for its exceptional climate variability, which exceeds that of comparable climatic regions globally (Harris et al., 2018; Ma et al., 2016; Peel et al., 2004), Australia supports ecosystems ranging from tropical savannas and rainforests in the north, semi-arid shrublands and grasslands in the center, to evergreen forests and woodlands in the south (Moore et al., 2016). This ecological diversity makes Australia an ideal location for testing and evaluating vegetation mapping techniques.

The vegetation classes utilised for this study and derived from Scarth et. al (2019) range from low/medium scattered trees to very tall closed/open forests, including low/medium/tall woodlands and medium/tall forests. These study areas exhibit marked differences in the distribution and density of foliage and woody elements from the sub-canopy through to the overstory, including variations in the layering of vegetation (e.g., the presence or absence of a distinct understory), the continuity of foliage, and the vertical gap fraction, providing a complete overview of the variability found across the continent.

We made sure to also take point cloud densities into account, classifying datasets into low (< 3 pts/m2), medium (3 to 16 pts/m2) and high (>16 pts/m^2) point cloud densities datasets. To categorize point cloud density, the ELVIS subset's first-return point cloud densities were grouped into three distinct classes. This classification was performed using the Jenks-Caspall algorithm which optimizes data partitioning into natural breaks.

These datasets were processed using several combinations of three Canopy Height Model (CHM) algorithms (Point to Raster [P2R], Pit-Free (Khosravipour et al., 2014), and Triangulated Irregular Network [TIN]), and four ITS algorithms, namely DalPonte (Dalponte et al., 2016), Watershed, Silva (Silva et al., 2016), and Li (Li at al. (2012). This generated 444 segmentation results (Figure 2), validated against 3,387 manually delineated tree crowns to evaluate detection and delineation accuracy across vegetation types and point cloud densities.

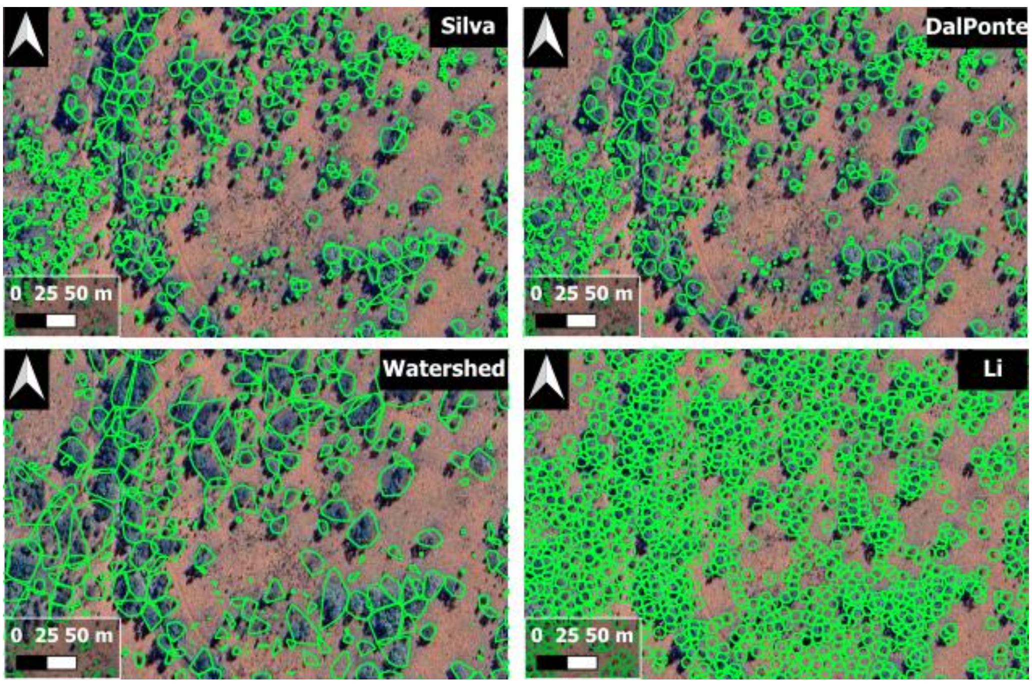

Figure 2. Example results of the four Individual Tree Segmentation (ITS) algorithms run in Wyandra, QLD, Australia.

To assess detection accuracy, each delineated crown was labelled as oversegmented, where multiple crowns overlap one groundtruth, undersegmented, where one crown spans multiple groundtruths, or a match, where one crown aligns with one groundtruth (Figure 3).

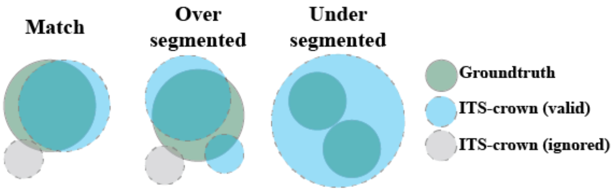

Figure 3. The figure highlights matches, oversegmentation, and undersegmentation cases used to evaluate detection accuracy.

To quantify matched and undersegmented crowns delineation accuracy, we used the Intersection over Union (IoU) as defined as:

$$IoU_i = \frac{A_{\cap_i}}{A_{\cup_i}} \tag{1}$$

where i represents a valid pairing between a grountruth (gt) and its ITS-crown while $A_{\cap_i}$ and $A_{\cup_i}$their spatial intersection and union respectively. For oversegmented crowns, calculating individual IoU for each fragment (*frag*) would lead to inaccurate and overestimated IoU results. Therefore, we adjusted the IoU (adj_IoU) to account for fragmentation of the ITS-crowns as follows:

$$adj\_IoU_i^{frag} = IoU_i \times \frac{A_{\cap_i}^{frag}}{A_{gt_i}} \tag{2}$$

where IoU_i is the IoU calculated for each fragment *frag* of the oversegmented ITS-crown in pairing *i*, $A_{\cap_i}^{frag}$ is the area of a fragment of the pairing *i* and A_{gt_i} is the total area of the ground truth of the pairing *i*.

3. Preliminary results and discussion

Regarding detection accuracies, among the algorithms tested, the Pit-Free CHM algorithm obtained the highest match ratio (29% matches) whereas the TIN the lowest (25%), despite differences are minimal (Figure 4A). The Pit-Free (pit) algorithm's effectiveness lies in its ability to eliminate pits, which occur when fine grid resolutions exceed the point density, leaving certain grid cells without any points and resulting in undefined values or "pits." These empty pixels can negatively impact the accuracy of subsequent analyses. By addressing this issue, the Pit-Free algorithm ensures a smoother and more continuous representation of canopy surfaces. This capability not only improves the visual quality of the data but also significantly enhances the accuracy of downstream applications, such as tree detection and delineation, when compared to traditional methods (Khosravipour et al., 2014). The reduction of pits allows for more reliable segmentation of tree crowns, contributing to better overall performance in vegetation mapping and monitoring.

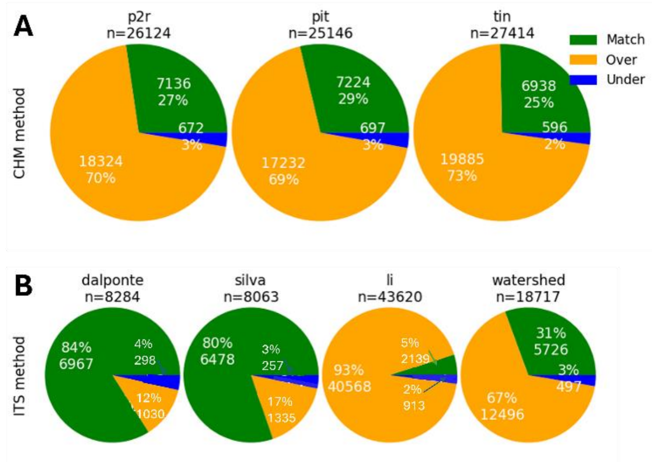

Figure 4. Detection accuracies per CHM (A) and ITS (B) algorithms.

Regarding ITS algorithms, the best detections are obtained using DalPonte (84% match) whereas the worst with Li (5%), which is very prone to oversegmentation (Figure. 4B). Missed trees are an uncommon occurrence, with each approach typically failing to detect only one groundtruth tree in their delineations, on average.

In terms of vegetation type, the best detections are obtained in low scattered trees environments (54% match), typical of rangelands. These ecosystems play a vital role in accurate tree mapping, given that rangelands account for 75% of Australia's land area and hold approximately two-thirds of its carbon reserves (Donohue et al., 2012). Despite their significance, field data from these regions remain scarce and challenging to obtain due to their remote and vast nature (Porfirio et al., 2020). As a result, large-scale ITS methods are indispensable for effective monitoring and analysis of these critical landscapes. Conversely, the worst detection rates are found in tall open woodlands (15% match, Figure. 5) and generally in forests. In forested areas, the reduced delineation accuracies were primarily attributed to the phenomenon of clumping, or crown clustering, which led to higher undersegmentation ratios. This clustering effect occurs when tree crowns are closely spaced together, making it difficult for segmentation algorithms to accurately distinguish individual crowns, resulting in undersegmentation of tree crowns.

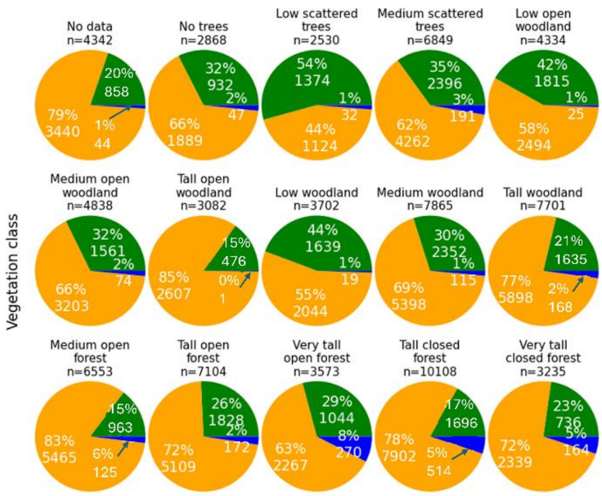

Figure 5. Detection accuracies per vegetation type.

https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-1223-2025

Regarding delineation accuracies, reported as median *adj_IoU (unitless)*, the best performing CHM algorithms, independently from segmentation methods, is Pit-Free (0.16), followed by p2r (0.14) and lastly Tin (0.12). This corroborates our previous observations that Pit-Free algorithm generated CHM that are suitable to obtain the best delineations overall. Note that these seemingly low values are attributable to the combined performances of each CHM the Li ITS algorithm.

In fact, Li segmentation algorithm resulted in very poor delineation accuracies (0.07) across all CHM methods, whereas the best ITS method was DalPonte (0.66), followed by Silva (0.43) and watershed (0.19). Once again, DalPonte algorithm superior performances also in delineating crowns result in the combination Pit-Free & DalPonte algorithms being the most generalisable and accurate of our study.

In terms of vegetation type, the highest median *adj_IoU* are obtained in high point density point clouds representing medium scattered trees (0.3, Figure 6) whereas the lowest in low points density tall woodlands (0.03, Figure 6). The superior performances of high point cloud density datasets was expected, however also low density point cloud classes also returned satisfactory results.

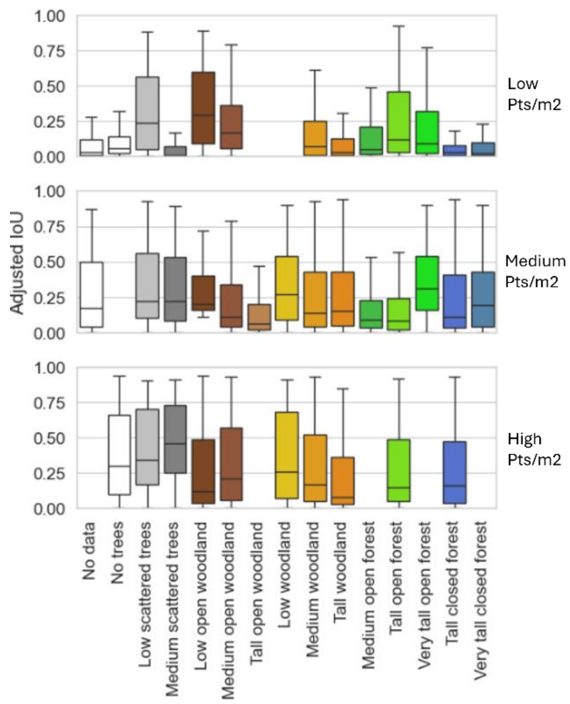

Figure 6. Delineation accuracies per vegetation type and point cloud density. Note some classes are missing due to unavailable data.

Overall, our analysis revealed that the performance of tree crown delineation algorithms varied significantly across different vegetation types. The highest accuracy rates were observed in sparsely vegetated areas such as rangelands and woodlands, which are particularly important for mapping Australia's vast and ecologically significant rangeland ecosystems. These regions, which cover a substantial portion of the country's landmass, benefit from the relatively open canopy structures that allow for more precise crown detection. In contrast, the delineation accuracy was lower in closed-canopy forests. This reduction in performance was primarily driven by the complex vegetation structures found in these areas, including the presence of dense crown clumping and multi-layered canopies. These characteristics make it challenging for segmentation algorithms to differentiate between individual tree crowns, leading to increased undersegmentation and reduced overall delineation accuracy.

These findings emphasize the need for algorithm optimization tailored to tropical and eucalypt-dominated forests. Moreover, oversegmentation emerged as the predominant challenge across all ITS methods, particularly in tropical savannas and rangelands with multi-peak trees. Oversegmentation inflates error rates and complicates the development of robust and accurate individual tree crown dataset, potentially useful for DL training datasets.

Retaining noisy labels may be beneficial, as overly stringent quality control risks underrepresenting complex canopy environments.

Acknowledgements

We extend our gratitude to Geoscience Australia for hosting and maintaining the publicly available Elevation and Depth (ELVIS) cloud-based platform, in collaboration with agencies under the Australian Intergovernmental Committee on Surveying and Mapping.

We also thank the Commonwealth Scientific and Industrial Research Organization (CSIRO) for granting access to their supercomputing resources, which enabled the extensive data processing and analyses conducted for this research. Their support has been invaluable in advancing this work.

References

Dalponte, M., Coomes, D.A., 2016. Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. *Methods Ecol. Evol*. 7, 1236–1245. https://doi.org/10.1111/2041-210X.12575.

Donohue, R., McVicar, T., Li, L., 2012. Modelling current and future carbon and water balances for Australia’s forests, woodlands and rangelands. CSIRO. https://doi.org/10.4225/08/5852DB73EE65B

Goldbergs, G., Maier, S.W., Levick, S.R., Edwards, A., 2018. Efficiency of Individual Tree Detection Approaches Based on Light-Weight and Low-Cost UAS Imagery in Australian Savannas. *Remote Sensing* 10, 161. https://doi.org/10.3390/rs10020161

Harris, R.M.B., Beaumont, L.J., Vance, T.R., Tozer, C.R., Remenyi, T.A., Perkins-Kirkpatrick, S.E., Mitchell, P.J., Nicotra, A.B., McGregor, S., Andrew, N.R., Letnic, M., Kearney, M.R., Wernberg, T., Hutley, L.B., Chambers, L.E., Fletcher, M.-S., Keatley, M.R., Woodward, C.A., Williamson, G., Duke, N.C., Bowman, D.M.J.S., 2018. Biological responses to the press and pulse of climate trends and extreme events. *Nat. Clim. Chang*. 8, 579–587. https://doi.org/10.1038/s41558-018-0187-9

Khosravipour, A., Skidmore, A.K., Isenburg, M., Wang, T., Hussin, Y.A., 2014. Generating Pit-free Canopy Height Models from Airborne Lidar. *Photogrammetric Engineering & Remote Sensing* 80, 863–872. https://doi.org/10.14358/PERS.80.9.863.

Latella, M., Sola, F., Camporeale, C., 2021. A Density-Based Algorithm for the Detection of Individual Trees from LiDAR Data. Remote Sensing 13, 322. https://doi.org/10.3390/rs13020322

Li, W., Guo, Q., Jakubowski, M.K., Kelly, M., Others, 2012. A new method for segmenting individual trees from the lidar point cloud. Photogramm. Eng. *Remote Sens*. 78, 75–84.

Lindenmayer, D.B., 2023. Forest Biodiversity Declines and Extinctions Linked with Forest Degradation: A Case Study from Australian Tall, Wet Forests. *Land* 12, 528. https://doi.org/10.3390/land12030528.

Ma, X., Huete, A., Cleverly, J., Eamus, D., Chevallier, F., Joiner, J., Poulter, B., Zhang, Y., Guanter, L., Meyer, W., Xie, Z., Ponce-Campos, G., 2016. Drought rapidly diminishes the large net CO2 uptake in 2011 over semi-arid Australia. *Sci. Rep*. 6, 37747. https://doi.org/10.1038/srep37747

Moore, C.E., Brown, T., Keenan, T.F., Duursma, R.A., van Dijk, A.I.J.M., Beringer, J., Culvenor, D., Evans, B., Huete, A., Hutley, L.B., Maier, S., Restrepo-Coupe, N., Sonnentag, O., Specht, A., Taylor, J.R., van Gorsel, E., Liddell, M.J., 2016. Reviews and syntheses: Australian vegetation phenology: new insights from satellite remote sensing and digital repeat photography. *Biogeosciences* 13, 5085–5102. https://doi.org/10.5194/bg-13-5085-2016

Oehmcke, S., Li, L., Trepekli, K., Revenga, J.C., Nord-Larsen, T., Gieseke, F., Igel, C., 2024. Deep point cloud regression for above-ground forest biomass estimation from airborne LiDAR. *Remote Sens. Environ*. 302, 113968. https://doi.org/10.1016/j.rse.2023.113968.

Peel, M.C., McMahon, T.A., Finlayson, B.L., 2004. Continental differences in the variability of annual runoff-update and reassessment. *J. Hydrol*. 295, 185–197. https://doi.org/10.1016/j.jhydrol.2004.03.004

Porfirio, L.L., Antille, D., Watson, I., Abbott, B., Bowman, D., Briggs, P., Canadell, J., Churchill, A., Donohue, R., Guerschman, J., Haverd, V., Hill, M., Knauer, J., Murphy, B., Paget, M., Prior, L., Roxburgh, S., Williamson, G., 2020. Productivity and biomass of Australia's rangelands: Towards a national database. *Proc. R. Soc. Queensl*. https://doi.org/10.5962/p.357887

Silva, C.A., Hudak, A.T., Vierling, L.A., Loudermilk, E.L., O'Brien, J.J., Hiers, J.K., Jack, S.B., Gonzalez-Benecke, C., Lee, H., Falkowski, M.J., Khosravipour, A., 2016. Imputation of Individual Longleaf Pine (Pinus palustris Mill.) Tree Attributes from Field and LiDAR Data. *Can. J. Remote Sens*. 42, 554–573. https://doi.org/10.1080/07038992.2016.1196582.

Scarth, P., Armston, J., Lucas, R., Bunting, P., 2019. A Structural Classification of Australian Vegetation Using ICESat/GLAS, ALOS PALSAR, and Landsat Sensor Data. *Remote Sensing* 11, 147. https://doi.org/10.3390/rs11020147.

Stern, H., de Hoedt, G., Ernst, J., 2000. Objective classification of Australian climates. *Australian Meteorological Magazine* 49, 87–96.

Wallace, L., Lucieer, A., Watson, C.S., 2014. Evaluating Tree Detection and Segmentation Routines on Very High Resolution UAV LiDAR Data. IEEE Trans. Geosci. *Remote Sens*. 52, 7619–7628. https://doi.org/10.1109/TGRS.2014.2315649

Weinstein, B.G., Marconi, S., Bohlman, S., Zare, A., White, E., 2019. Individual Tree-Crown Detection in RGB Imagery Using Semi-Supervised Deep Learning Neural Networks. *Remote Sensing* 11, 1309. https://doi.org/10.3390/rs11111309