Remote Sensing of Environment 237 (2020) 111520

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

# Modelling canopy gap probability, foliage projective cover and crown projective cover from airborne lidar metrics in Australian forests and woodlands

Adrian Fisher[a,b,*], John Armston[c,a], Nicholas Goodwin[d,a], Peter Scarth[a,d]

[a] Joint Remote Sensing Research Program, School of Earth and Environmental Sciences, University of Queensland, Brisbane, QLD, 4072, Australia
[b] Centre for Ecosystem Science, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, 2052, Australia
[c] Department of Geographical Sciences, University of Maryland, 2181 Samuel J. LeFrak Hall, 7251 Preinkert Drive, College Park, MD, 20742, USA
[d] Remote Sensing Centre, Science Delivery, Department of Environment and Science, 41 Boggo Road, QLD, 4102, Australia

ARTICLE INFO

*Keywords:*
Airborne lidar
Vegetation structure
Canopy density

ABSTRACT

Tree canopy density metrics (TCDM) derived from airborne lidar data are used in a range of crucial environmental monitoring, forestry and natural resource management applications. The derivation of spatially and temporally consistent TCDM, however, typically requires field calibration to account for differences in instrument/survey parameters. Lidar surveys with no coincident field measurements consequently will have an unknown error associated with TCDM limiting their application. In this study, we analysed an extensive set of lidar captures with coincident field data to determine the lidar TCDM that best match the canopy gap probability ($P_{gap}$), foliage projective cover (FPC) and crown projective cover (CPC). Furthermore, we developed and evaluated models designed to reduce the bias introduced by variations in lidar instrument and survey acquisition parameters. The dataset incorporated 148 field sites (100 m diameter circular plots) coincident with 13 different lidar surveys between 2008 and 2015, distributed across a range of Australian forests and woodlands. The best lidar metric for $1 - P_{gap}$, achieving a root mean square error (RMSE) of 6.7% with 95% confidence intervals (CI) of 6.1–7.3%, was the proportion of all returns greater than a canopy height threshold ($t_{canopy}$) of 1.5 m above ground ($d_{all}$). The best metric for FPC (RMSE = 6.0%, CI = 5.3–6.7%) used the proportion of returns, weighted as the fraction of the number of returns recorded from each pulse ($d_{weighted}$), with $t_{canopy}$ of 1.7 m. The best metric for CPC (RMSE = 7.0%, CI = 6.4–7.7%) was the proportion of 0.5 m pixels greater than 0.8 m above the ground, for an interpolated canopy height model ($d_{interp}$). Overall bias for these metrics was low (~1%), however, the bias for individual surveys varied significantly. For example, for one survey $d_{all}$ consistently underestimated $1 - P_{gap}$ with a bias of −8.3%, while a different survey consistently overestimated $1 - P_{gap}$ with a bias of 3.8%. Elastic net regression models, using instrument, survey and plot parameters as predictor variables, were unable to consistently remove the bias. No relationships could be discerned between lidar parameters and the bias between lidar metrics and field measurements, potentially due to complex interactions between parameters, the spatial scale of the field plots, and uncertainties in field measurements and lidar attributes. Although the bias could not be modelled, the results provide metrics to derive $P_{gap}$, FPC and CPC with less than 10% error from lidar surveys captured with similar parameters across Australia (> 600,000 km$^2$).

## 1. Introduction

Over the last two decades airborne lidar has been increasingly used to derive quantitative estimates of vegetation structure over large continuous areas of woodland and forest at a scale between field surveys and satellite imagery (Wulder et al., 2012). This has seen the development of new instruments with a variety of survey configurations with implications for deriving vegetation structure and tree canopy density in particular (Disney et al., 2010; Goodwin et al., 2006; Korpela et al., 2012; Næsset, 2009). Instruments vary in their laser wavelength, pulse repetition frequency (PRF), pulse length, pulse shape, pulse energy, signal triggering mechanism, detector sensitivity, beam

* Corresponding author. Joint Remote Sensing Research Program, School of Earth and Environmental Sciences, University of Queensland, Brisbane, QLD, 4072, Australia.
*E-mail addresses:* adrian.fisher@unsw.edu.au (A. Fisher), armston@umd.edu (J. Armston), nicholas.goodwin@des.qld.gov.au (N. Goodwin), p.scarth@uq.edu.au (P. Scarth).

https://doi.org/10.1016/j.rse.2019.111520
Received 3 May 2019; Received in revised form 21 October 2019; Accepted 2 November 2019
Available online 19 November 2019
0034-4257/ 

divergence, vertical discrimination distance, and scan pattern (e.g. Leica Geosystems 2007, 2008). Surveys vary in their flying height above ground, flying speed, maximum scan angle from nadir, scan rate and swath overlap. Data processing is also variable, with providers using different detection thresholds, ground classification algorithms, noise filtering algorithms, and instrument calibration routines. Other commonly reported parameters are conditional on instrument/survey design: swath width is dependent on flying height and scan angle; footprint size at nadir (the diameter of the laser pulses when they interact with vegetation and the ground) is dependent on the flying height and beam divergence; and pulse density (the mean number of laser pulses per unit area) is dependent on the PRF, flying height and speed, scan rate and swath overlap. Instrument/survey parameters are often selected to achieve a desired pulse density and survey cost within return pulse signal-to-noise limits.

Airborne lidar can be used to derive tree canopy density metrics (TCDM) as the laser pulse energy passes through gaps within and between tree crowns and reflects off foliage and woody elements, as well as the ground. In the lidar systems considered here, each return recorded a 3-dimensional position (x, y, z) and an intensity value related to the strength of the backscatter. Dense canopies reflect comparatively more of the laser pulses' energy, while in sparse canopies more pulses are likely to reach the ground due to fewer interactions and reduced energy loss. Lidar canopy density metrics are usually calculated from the fraction of returns detected within the overstorey canopy, although the calculations vary depending on several factors, such as which return number(s) are used and how the overstory canopy is defined.

Calibration of lidar TCDM is usually achieved though empirical relationships with field measurements of density (Armston et al., 2009; Næsset, 2009; Wulder et al., 2008). The plot-based measurements considered here are canopy gap probability ($P_{gap}$), foliage projective cover (FPC) and crown projective cover (CPC). $P_{gap}$ is the probability of a beam of infinitesimal width penetrating through the canopy and intercepting the ground, where the beam zenith angle is usually fixed at nadir. $P_{gap}$ is a unitless proportion with a range of 0–1, and $1 - P_{gap}$ is positively related to other measures of canopy density. The parameter used in many Australian vegetation mapping programs is FPC (Danaher et al., 2010; Fisher et al., 2016). FPC is the proportion of horizontal area covered by the vertical projection of photosynthetic foliage for all vegetation taller than 2 m (Specht, 1983), and can only be derived from $P_{gap}$ through accounting for the proportion of woody material within the canopy ($\alpha$). Most trees in Australia do not seasonally drop their leaves but maintain a relatively stable FPC, which is correlated to the annual water balance of the ecosystem (Specht, 1983). Crown projective cover (CPC) is the proportion of horizontal area covered by the vertical projection of all tree crowns assuming crowns are opaque (Fisher et al., 2018). CPC includes within-crown gaps but excludes between-crown gaps. It has previously been called crown cover (Hnatiuk et al., 2009), vertical canopy cover (Korhonen et al., 2011), or canopy cover (FAO, 2012; Melin et al., 2017), and is the inverse of between-crown gaps (Schliemann and Bockheim, 2011). FPC and CPC are commonly calculated as percentages and are often used in classifying structural classes of woodland and forest. Further information on $P_{gap}$, FPC and CPC, and how they relate to each other was presented in Fisher et al. (2018).

Lidar instrument/survey parameters influence the characterisation of vegetation canopies in several ways, and a great deal of research has been conducted on the issue. TCDM generally overestimate true canopy density, as instruments are blind to canopy gaps smaller than the footprint size (Lovell et al., 2003; Morsdorf et al., 2006). This implies that greater footprint sizes would result in greater TCDM, however, experiments in which flying height and footprint size were increased found negligible change to TCDM (Goodwin et al., 2006), or decreases to TCDM (Armston et al., 2013a; Morsdorf et al., 2008). These studies increased the footprint size by increasing the flying height with a fixed beam divergence, so they also increased the transmission energy losses of the laser pulses, which may have had a greater influence on TCDM (Armston et al., 2013a). Decreasing the PRF, flying height and footprint size all increase the pulse peak energy or energy concentration, which can increase the detectable backscatter for greater distances within the canopy (Chasmer et al., 2006; Hopkinson and Chasmer, 2009; Næsset, 2009). Excluding terrain interactions, higher scan angles increase the footprint size and the time/distance pulses take through the canopy, increasing the likelihood of canopy returns (Disney et al., 2010; Montaghi, 2013; Morsdorf et al., 2008). Several studies have found that changing pulse density does not significantly change TCDM (Hansen et al., 2015; Lim et al., 2008; Wilkes et al., 2015), although even when instrument/survey parameters were kept the same, TCDM were found to vary by up to 4% (Bater et al., 2011). Importantly however, low pulse densities (e.g. $< 1$ pulse/m$^2$) in dense canopies and complex terrain can result in less accurate elevation models of the ground surface, increasing errors in lidar TCDM that depend on accurate vegetation height to define the position of the canopy (Wilkes et al., 2015).

Removing the effects of instrument/survey properties has required lidar TCDM to be calibrated against field measurements. Once calibrated, they can be reliably compared with independent measures, such as those from satellite imagery (Armston et al., 2009; Korhonen et al., 2013). Other approaches, such as using sampling theory to normalise lidar canopy heights for variations in pulse density and footprint size have been investigated (Roussel et al., 2017), but calibration against field measurements using empirical or physically based models (Pearse et al., 2017) is currently the standard approach. Unfortunately, this severely limits the use of most historical lidar data acquired without field measurements of vegetation density, such as data from topographical surveys, which in Australia cover $> 600{,}000$ kms$^2$. It may be possible to calibrate canopy density metrics from these surveys if models could be demonstrated to be robust across surveys with a range of instrument/survey parameters.

## 2. Objectives

The research presented here explored an extensive field and airborne lidar dataset in Australian woodlands and forests to test the calibration of lidar TCDM against three field measurements of canopy density ($P_{gap}$, FPC and CPC). More specifically, the two main objectives were to answer the following questions:

1. Do different instrument/survey parameters bias the relationships between lidar metrics and field measurements, and if so, can models be developed to reduce the bias?
2. Which lidar metrics and models best equate to the field measurements, and how accurate are these for deriving $P_{gap}$, FPC and CPC across a range of lidar surveys?

## 3. Methods

### 3.1. Study area

The study encompasses sites across all states of Australia, with most sites located in the eastern states of Queensland and New South Wales (NSW). Seven lidar datasets and 148 coincident field sites were collected across the region between 2008 and 2015 (Fig. 1, Table 1). The vegetation across Australia varies from shrublands, grasslands, woodlands and forests, whose broad distribution is influenced by climatic patterns. Forests that are present on the humid east coast, and in the far south-west, transition to woodlands and then shrublands and grasslands as aridity increases towards central Australia. Northern Australia has tropical summer rainfall and extensive savannas, while southern Australia has a Mediterranean climate with winter rainfall.

The surveys sample many native vegetation communities present across Australia. In Queensland, a variety of vegetation communities across a very wide area (Fig. 1) were sampled by the QLD survey,

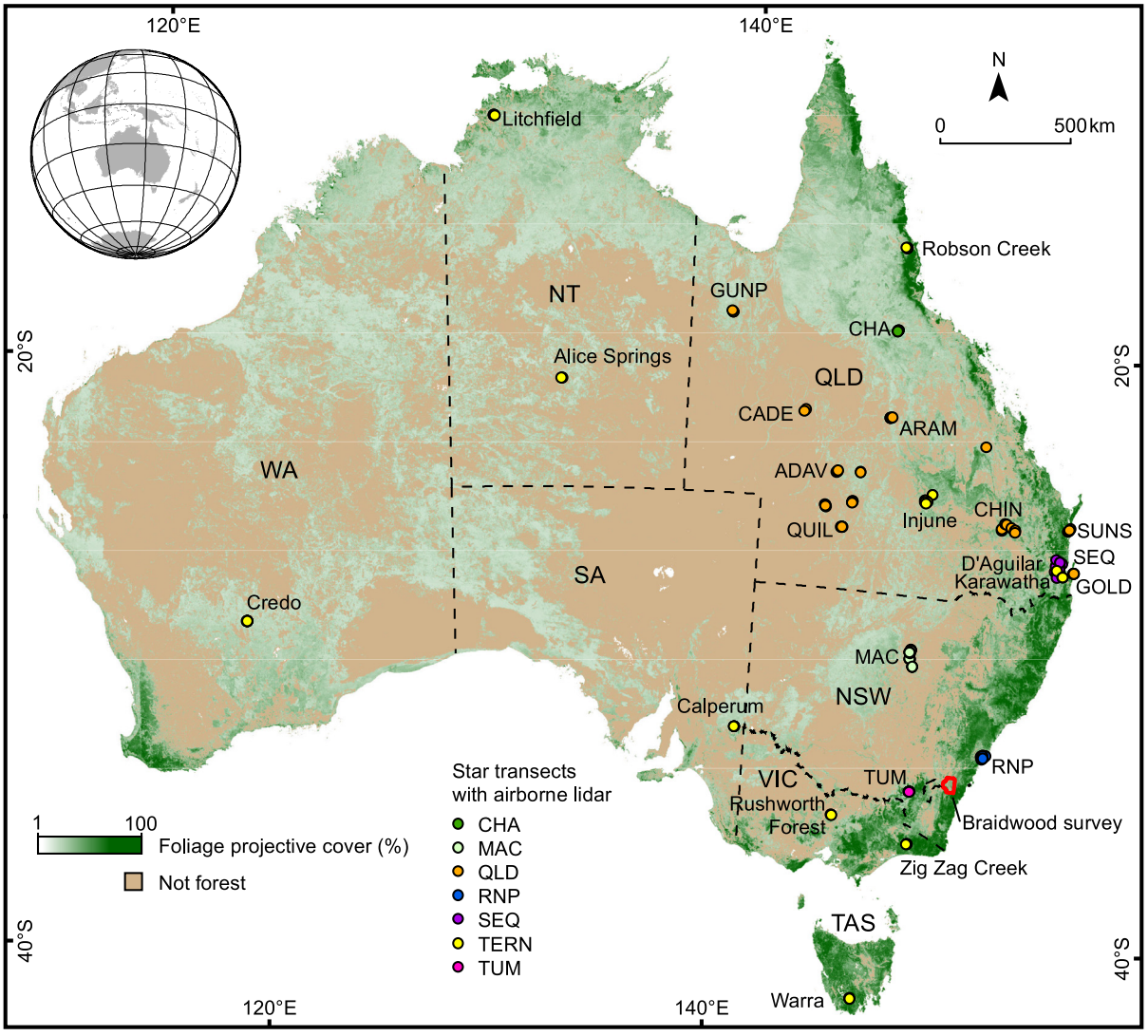



**Fig. 1.** Location of the sites across the Australian states of Queensland (QLD), New South Wales (NSW), Victoria (VIC), Tasmania (TAS), South Australia (SA), Western Australia (WA) and the Northern Territory (NT). The background image is of foliage projective cover (FPC), derived from a time series of Landsat satellite imagery (Gill et al., 2017). FPC is greatest in the humid forests along the east coast and south west coast, and is lowest in the grasslands, shrublands and woodlands in central and southern Australia.

including woodland (*Acacia spp.*), open forest (*Eucalyptus spp.* and *Callitris sp.*), and closed forest (*Eucalyptus spp.* and *Corymbia spp.*), with more detail in Armston et al. (2009). The CHA survey sampled savanna open woodland (*Eucalyptus spp.*, *Coryombia spp.* and *Acacia spp.*), which is described in Armston et al. (2013a). The SEQ survey sampled woodlands and forest close to the city of Brisbane. Several sites were sampled by the TERN surveys, including the notophyll vine rainforest of Robson Creek, the forests of D'Aguilar National Park and Karawatha Forest near Brisbane, and the woodlands at Injune. These woodlands have been intensively researched as part of the Injune Landscape Collaborative Project (Lucas et al. 2006, 2008).

In NSW, the campaigns also sampled a variety of vegetation types. The TUM survey sampled tall wet sclerophyll forest dominated by mature alpine ash (*Eucalyptus delegatensis*) near the town of Tumbarumba, although the area is a managed forest last thinned in 1985 (Hopkinson et al., 2013). The RNP survey sampled woodland (*Eucalyptus spp.*), mallee-heath-woodland (*Banksia spp.*, *Eucalyptus spp.*), and forest (*Angophora spp.*, *Corymbia spp.* and *Eucalyptus spp.*) in the Royal National Park, which may have been logged in the 1920's despite being proclaimed as a national park in 1879 (Adam and King, 2013; Keith, 2013). The MAC survey sampled River Red Gum (*Eucalyptus camaldulensis*) and River Cooba (*Acacia stenophylla*) forest around the Macquarie Marshes wetland, and a variety of the drier woodlands (*Eucalyptus spp.*, *Flindersia spp.*, *Syncarpia spp.* and *Geijera spp.*) and is described in Dawson et al. (2016).

The TERN survey also sampled several sites across Australia. In Victoria short, dry sclerophyll forest was sampled at Rushworth Forest and medium height, dense forest was sampled at Zig Zag Creek. In Tasmania tall, wet sclerophyll forest dominated by *Eucalyptus obliqua* was sampled at Warra. In South Australia mallee woodlands (*Eucalyptus spp*) were sampled at Calpernum. In the Northern Territory tropical savanna was sampled at Litchfield and mulga (*Acacia aneura*) woodlands were sampled at Alice Springs. In Western Australia open woodlands dominated by *Eucalyptus salmonophloia* were sampled at Credo.

### 3.2. Field data

The 148 field sites were accumulated from local, regional and national projects, and do not represent a systematic sample across Australia (Table 1). All sites were positioned away from roads, buildings, cliffs and gullies. Using satellite and aerial imagery and existing

**Table 1**
Lidar instrument/survey properties and field data from seven campaigns across eastern Australia (Fig. 1). Mean point and pulse densities and maximum scan angles were calculated from the data captured over the field sites. Footprint diameter was calculated from beam divergence and mean flying height. Note that all sites in the CHA survey and three sites in the TERN survey were acquired with multiple instrument/survey configurations. All sites for the TERN survey were measured with the B2 configuration (*) except for one site (gold0101). Some acquisition dates are not specific as data was collected in several campaigns over a period of time.

| | Survey name | QLD | SEQ | CHA | TUM | TERN | RNP | MAC |
|---|---|---|---|---|---|---|---|---|
| Lidar data | Survey description | Several flights across Queensland sampling remnant vegetation types | Large area survey in southeast Queensland | Several flights near Charters Towers, northern Queensland | Small area survey over the Tumbarumba TERN site in southern NSW | Several flights over TERN Auscover sites around Australia | Large area survey over southern Sydney and the Royal and Heathcote National Parks, NSW | Large area survey over the Macquarie Marshes region, central west NSW |
| | Dates acquired | 2008 | 2009/03/25–2009/09/16 | 2010/06/18 | 2011/01/16 | 2012–2015 | 2013/04/11–2013/04/24 | 2015/03/30–201504/09 |
| | Instrument | Optech ALTM3025 | Leica ALS50-II | RIEGL LMS-Q680i | Leica ALS60 | RIEGL LMS-Q560 | Leica ALS50-II | Leica ALS50-II |
| | Laser wavelength (nm) | 1064 | 1064 | 1550 | 1064 | 1550 | 1064 | 1064 |
| | Codes for multiple flights | – | – | A2/A3/A4 | – | B1/B2*/B3/B4/B5 | – | – |
| | Flying height above ground (m) | 1800 | 1650 | 450/900/1200 | 1350 | 150/300/450/600/800 | 1828 | 1498 |
| | Pulse rate (kHz) | 70 | 126 | 200/200/150 | 162 | 240/240/200/100/70 | 142 | 142 |
| | Beam divergence (mrad) | 0.30 | 0.22 | 0.50 | 0.22 | 0.50 | 0.22 | 0.22 |
| | Footprint diameter, nadir, ground (m) | 0.54 | 0.36 | 0.23/0.45/0.60 | 0.30 | 0.08/0.15/0.23/0.30/0.40 | 0.40 | 0.33 |
| | Maximum returns per pulse | 6 | 4 | 7 | 4 | 7 | 4 | 4 |
| | Vertical discrimination distance (m) | 4.9 | 3.5 | 1.0 | 3.5 | 1.0 | 3.5 | 3.5 |
| | Mean point density (points $m^{-2}$) | 3.4 | 2.0 | 5.6/2.7/1.4 | 6.5 | 40/13/22/12/8 | 2.0 | 2.0 |
| | Mean pulse density (pulses $m^{-2}$) | 2.6 | 1.4 | 4.2/2.3/1.2 | 3.9 | 17/8/11/5/3 | 1.3 | 1.3 |
| | Max. scan angle from nadir (°) | 22 | 17 | 25/30/30 | 6 | 26/29/25/21/21 | 15 | 21 |
| Field data | Number of field sites | 46 | 20 | 3 | 5 | 47 | 12 | 15 |
| | Date field data acquired | 2008 | 2009 | 2010/06 | 2011/01/11-13 | 2012–2014 | 2013/05/15-26 | 2015/05-06 |
| | Vegetation | Woodland and forest | Woodland - forest | Woodland | Forest | Woodland - forest | Woodland - forest | Woodland - forest |
| | FPC range (%) | 0–96 | 23–93 | 5–18 | 37–48 | 0–93 | 3–87 | 18–62 |

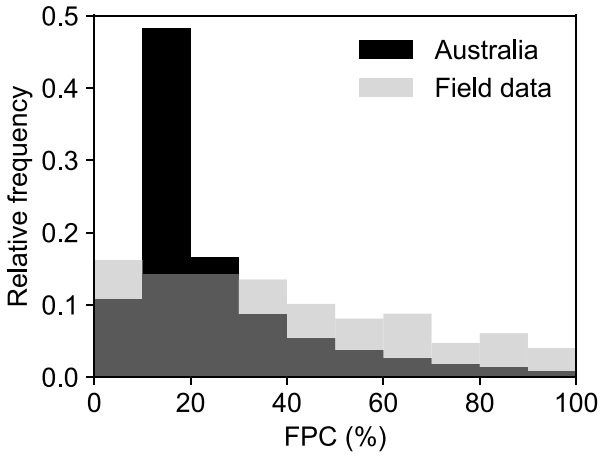



**Fig. 2.** Comparing FPC histograms from Australia (Gill et al., 2017) with the 148 field plot measurements.

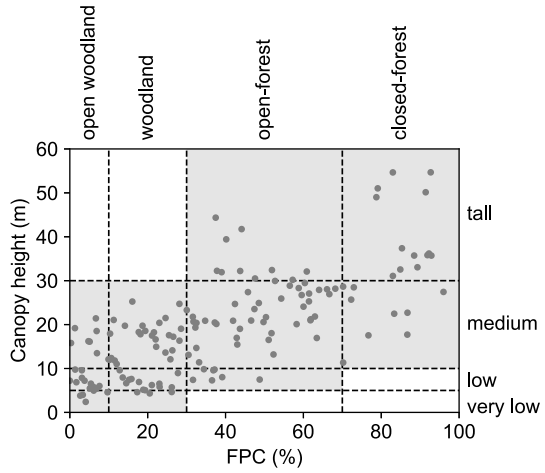



**Fig. 3.** Vegetation structural classes sampled by the 148 field plots (grey dots), using height and FPC thresholds (dashed lines) from Specht and Specht (1999). The shaded area shows the common classes across Australia (Scarth et al., 2019; TERN AusCover, 2018), which all contain field plots. Field measured FPC values are plotted against lidar derived canopy height (99th percentile of heights greater than 0.5 m above the ground, within the 100 m diameter plot).

$$P_{gap} = 1 - (P_{o,g} + P_{o,b}) \quad (1)$$

$$FPC = \frac{P_{o,g}}{(1 - P_{o,b})} \quad (2)$$

vegetation maps, each site was selected in a patch of homogeneous vegetation and canopy cover. The plots sample the full range of FPC values (Fig. 2) and vegetation structural classes present in Australia (Fig. 3), however, they are not a statistically representative sample. Compared to the Australia wide map of FPC derived from Landsat satellite imagery (Gill et al., 2017) the field sampling of FPC is lower between 10-30% and higher between 0-10% and 30–100% (Fig. 2).

At each site the star transect method was used to measure $P_{gap}$, FPC and CPC from 300 vertical tube observations across a 100 m diameter plot (Muir et al., 2011). Overstorey (woody plants ≥ 2 m height) observations were made using a GRS densitometer®. The GRS, a siting tube with a mirror, a centred cross-hair and two bubble-line levels, was mounted on a 1.5 m pole. Midstorey (woody plants < 2 m height) observations were made if plant material was touching the side of the pole. Understorey (herbaceous plants < 2 m height) observations were made using a downward pointing laser attached to the pole. Overstorey observations were also classified as either within or between tree crowns, where crown edges were determined as the convex hull surrounding all living crown components (Muir et al., 2011). The distinction between overstory and midstorey woody plants was based on the maximum height of each plant, not on the height of the GRS. For example, an observation of a green leaf at 1.8 m height from a 15 m tall tree is included as an overstory observation even though it was below 2 m. The location of the centre of each site was determined through global positioning, with 87 sites having sub-metre accuracy through differential corrections and the remaining sites having an assumed accuracy of 5 m. Although the star transect method provided independent estimates of canopy density, the measurement errors are not easily quantified and have a known bias, with more observations in the centre of the star. They also have possible observer bias, especially in tall canopies, and on windy days (Trevithick et al., 2012). From the observations recorded at each star transect $P_{gap}$ was calculated according to equation (1), and FPC was calculated according to equation (2), where $P_{o,g}$ was the proportion of overstorey green foliage observations, and $P_{o,b}$ was the proportion of overstorey branch observations. Equation (2) assumes that overstorey branch observations are likely to occlude green foliage from the observer. CPC was calculated as the proportion of overstorey observations that were within living tree crowns.

### 3.3. Airborne lidar data

The seven lidar surveys were collected over the field sites by different teams, using a variety of equipment, for either research purposes or as part of larger surveys (Table 1). While some data are restricted, access details for the freely available data are described below.

The QLD survey included the revisit of sites used in Armston et al. (2009) and Armston et al. (2012), and were located to sample the main structural formations and remnant vegetation communities observed throughout Queensland, which range from open woodland to closed forest. The airborne surveys used in the SEQ, RNP and MAC datasets were acquired for large-area topographic mapping with opportunistic vegetation field measurements. Access to the data is available through the ELVIS online portal for elevation and depth foundation spatial data (Geoscience Australia, 2018). The TUM and TERN data were acquired at Terrestrial Ecosystem Research Network (TERN) sites, and are available through TERN AusCover (Trevithick, 2017). The CHA data were acquired for an experiment with full waveform data (Armston et al., 2013a), while the repeat flights acquired for three of the TERN sites were acquired during an experiment comparing terrestrial laser scanners within woodland and forest (Armston et al., 2013b) and are available through TERN AusCover. Field transects for most surveys were made within a month of the airborne lidar acquisition to minimise seasonal changes to the vegetation. Field measurements made for the TERN survey at Alice Springs, Rushworth Forest and Warra were different from the airborne acquisition by 2, 7 and 8 months respectively.

Five different lidar instruments were used across the seven surveys (Table 1), each with different instrument/survey parameters, including a variety of flying heights, pulse rates, pulse densities, laser footprints and scan angles (Table 1). The total number of unique surveys with different instrument/survey parameters was 13. The Optech ALTM3025 instrument recorded up to two returns per pulse, and had a dead time equivalent to a minimum vertical discrimination distance of 4.9 m between first and last returns (Goodwin et al., 2006). The Leica ALS50-II and ALS60 instruments recorded up to four returns per pulse, and had a vertical discrimination distance of 3.5 m (Leica Geosystems, 2007; 2008). The Riegl LMS-Q680i and LMS-Q560 instruments were able to record digitised full waveform signals with up to seven returns and have a much smaller vertical discrimination distance (~1 m).

#### 3.3.1. Lidar processing

Each dataset was acquired, processed and delivered as a point cloud in the lidar exchange format (LAS) with ground returns classified. Where the lidar surveys consisted of multiple flights, each flight's data

was separated so the influence of changes in scan angle and pulse density between flights on vegetation metrics could be examined. This resulted in 265 lidar subsets covering the 148 field sites. All processing of LAS files was performed using the open source software LAStools (Isenburg, 2015) and PyLidar (Armston et al., 2015). The ground height for each return location was calculated using the natural neighbour interpolator (Bunting et al., 2015; Sibson, 1981) and subtracted from the z-value to produce heights above ground level.

For each lidar survey over the sampled plots, we calculated a set of common forest metrics that describe the structural variations between plots. These include the number of points, pulses and ground points; the median scan angle; and the 99th percentile of heights above ground for all returns greater than 0.5 m above the ground. We also calculated ground slope in degrees by fitting a plane through the ground classified points, and ground relief as the vertical root mean square error (RMSE) of the ground points from the plane.

### 3.3.2. Lidar canopy density metrics

Calculations for most lidar canopy density metrics ($d$) involve defining a height above ground threshold ($t_{canopy}$) that separates the tree canopy from the understory and determining the fraction of lidar returns from above this threshold within an area (Table 2, Fig. 4). The simplest approach is to use all returns and calculate $d_{all}$, which provide the greatest sample of returns (Hopkinson and Chasmer, 2009; Riaño et al., 2004). As $d_{all}$ is influenced by the vertical discrimination distance of the instrument, it may be preferable to only use first returns to calculate $d_{first}$ (Armston et al., 2009; Lovell et al., 2003; Morsdorf et al., 2006). Both these metrics assume that returns are equal, while in reality they represent partial reflections of varying intensity. To account for this, the metrics $d_{intensity}$ (Gill et al., 2009; Lovell et al., 2003; Ni-Meister et al., 2001) and $d_{weighted}$ (Lovell et al., 2003) were developed. The calculation of $d_{intensity}$ assumes a linear intensity response of the lidar instrument, while $d_{weighted}$ assumes that multiple returns from a single pulse intercept targets with the same cross-sectional area. Although many other metrics have also been developed, such as the last echo cover index (LCI) (Korhonen et al., 2011), the Solberg Cover Index (SCI) (Solberg et al., 2009), or a modified version of $d_{first}$ where the canopy threshold ($t_{canopy}$) is defined as the mean height of all vegetation returns (Pearse et al., 2017), they were not considered in the present study.

The metrics defined in Table 2 were calculated for each lidar flight from all lidar pulses within each 100 m diameter circular plot, using a range of canopy height thresholds ($t_{canopy}$) between 0.5-2.0 m incremented in 0.1 m heights. Although trees were defined as > 2 m tall in the field, foliage, branches and trunks extend below 2 m. Testing different values of $t_{canopy}$ allowed the optimum threshold to be determined from the lidar data. The maximum height above ground threshold for ground returns ($t_{ground}$) was fixed at 0.5 m.

**Table 2**
The area based lidar tree canopy density metrics (TCDM) tested across Australia.

| Metric | Formula | Sources |
|---|---|---|
| $d_{first}$ | $\frac{\sum_{z=t_{canopy}}^{z=\max(z)} C_{first}}{\sum_{z=0}^{z=\max(z)} C_{first}}$ | Lovell et al. (2003)<br>Morsdorf et al. (2006)<br>Armston et al. (2009) |
| $d_{all}$ | $\frac{\sum_{z=t_{canopy}}^{z=\max(z)} C_{all}}{\sum_{z=0}^{z=\max(z)} C_{all}}$ | Riaño et al. (2004)<br>Hopkinson and Chasmer (2009) |
| $d_{weighted}$ | $\frac{\sum_{z=t_{canopy}}^{z=\max(z)} \frac{C_{all}}{n}}{\sum_{z=0}^{z=\max(z)} C_{first}}$ | Lovell et al. (2003)<br>Armston et al. (2012)<br>Armston et al. (2013a) |
| $d_{intensity}$ | $\frac{\sum_{z=t_{canopy}}^{z=\max(z)} I}{\sum_{z=t_{ground}}^{z=\max(z)} I + \sum_{z=0}^{z=t_{ground}} I\frac{\rho_v}{\rho_g}}$ | Ni-Meister et al. (2001)<br>Lovell et al. (2003)<br>Gill et al. (2009) |

| Term | Definition |
|---|---|
| $t_{canopy}$ | minimum height above ground for canopy returns |
| $t_{ground}$ | maximum height above ground for ground returns |
| $C_{first}$ | first returns within the area ($C_{first}$ = 1) |
| $C_{all}$ | all returns within the area ($C_{all}$ = 1) |
| $n$ | number of returns for each pulse |
| $I$ | intensity of each return within the area |
| $\rho_v$ | mean intensity for all returns above $t_{ground}$ within the area |
| $\rho_g$ | mean intensity for all returns below $t_{ground}$ within the area |

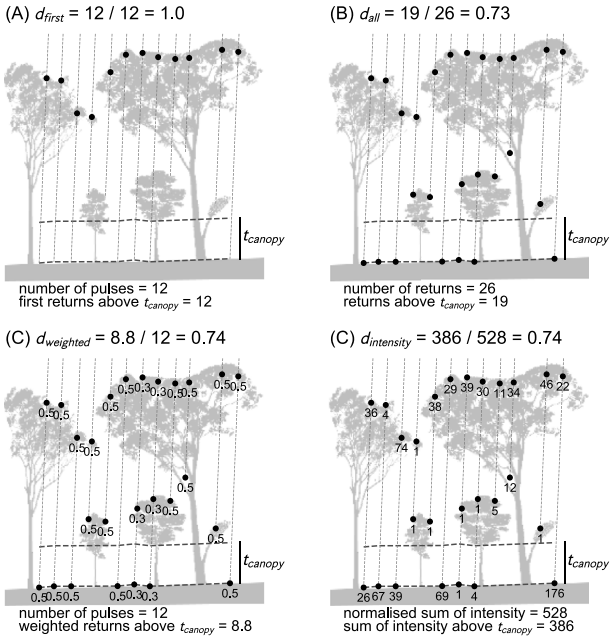



**Fig. 4.** An example of how the four tree canopy density metrics are calculated from a 10 m strip of lidar data over four trees. Lidar pulses are shown with dashed grey lines, while interpolated ground and minimum canopy height surfaces are shown with thicker dark grey dashed lines. (A) $d_{first}$ is calculated using only the first returns of each pulse (black dots). (B) $d_{all}$ is calculated from all returns (black dots). (C) $d_{weighted}$ is calculated from all returns (black dots) with the weights shown calculated by the number of returns in each pulse. (D) $d_{intensity}$ is calculated from all returns (black dots) normalised using the intensity values shown and the formula in Table 2.

### 3.3.3. Gridded tree crown models

An alternative to the lidar canopy density metrics listed in Table 2 is to create gridded tree crown models from the lidar data. Interpolation of the height above ground of lidar returns into a gridded canopy height model (CHM) at various resolutions is a very common practice. Vepakomma et al. (2008) classified a CHM using a canopy threshold ($t_{canopy}$) and then calculated the fraction of the CHM above this threshold within an area, deriving a metric which we term $d_{interp}$. We calculated $d_{interp}$ within each 100 m diameter circular plot using three grid sizes (0.1, 0.5 and 1 m) and a range of $t_{canopy}$ values (0.5–2.0 m in 0.1 m increments). The maximum lidar return height above ground was assigned to each grid cell, with the height of empty cells interpolated using the natural neighbour algorithm. Pixels where the CHM was > $t_{canopy}$ were 'in-crown', while those ≤ $t_{canopy}$ were considered 'between-crown' gaps, with $d_{interp}$ calculated as the proportion of the in-crown area within each circular plot area. The $d_{interp}$ metric is thought to relate to CPC, where it can be assumed that any canopy pits resulting from the interpolation are between-crown gaps.

Wang et al. (2008) and Korhonen et al. (2011) developed methods using morphological filters on CHMs, rather than interpolation, to fill no-data pixels, remove outlier pixels, and smooth crown edges. We implemented a simplified morphological method to derive the metric $d_{morph}$ from lidar data across each 100 m diameter plot. The method used four steps: 1) a binary canopy model from all returns above $t_{canopy}$ was gridded, 2) morphological closing filled gaps in the canopy model,

3) we assumed that within-crown gaps were smaller than the lidar pulse spacing, and between-crown gaps were sampled as cells that contained a return $< t_{ground}$ but that had no returns between $t_{ground}$ and $t_{canopy}$ (by re-defining the between-crown gap pixels in the model after morphological closing, we limited the gap filling to within-crown gaps without smoothing crown edges with further filtering), 4) $d_{morph}$ was then calculated as the proportion of the crown area within each circular plot area. The method required three parameters to be optimised: $t_{canopy}$; the size of the grid; and the size of the structuring element used by the morphological closing. We tested the method using a range of $t_{canopy}$ values (0.5–2.0 m in 0.1 m increments), three grid sizes (0.1, 0.5 and 1 m), and several sizes of disk-shaped structuring element (0–19 pixels in 2 pixel increments).

### 3.4. Statistical methods

In previous research, single lidar metrics were used as proxies for $P_{gap}$, FPC or CPC, and regression models between lidar metrics and field measurements determined calibration coefficients that allowed $P_{gap}$, FPC or CPC to be predicted from the lidar metrics. For example, FPC has been calibrated with linear models (Gill et al., 2009; Tickle et al., 2006; Weller et al., 2003) and physically based power models (Armston et al. 2009, 2012; Ediriweera et al., 2013). Others have assumed that instrument/survey effects are negligible, and have compared metrics and field measurements directly (Korhonen et al., 2011). Our approach was to conduct three separate comparisons. Firstly, the TCDM and field measurements were compared directly to one another. Secondly, regression models were developed between TCDM and measurements, to assess whether the models reduced RMSE and survey specific bias compared to the first direct comparison(s). Thirdly, multiple-output regression models were developed that modelled all three measurements ($P_{gap}$, FPC and CPC) from the TCDM simultaneously, which may have an advantage over single output models in that the three measurements are closely related. Each of these three methods are described in more detail below.

#### 3.4.1. Comparing metrics and measurements

The direct comparison was conducted between all lidar metrics ($d_{first}$, $d_{all}$, $d_{intensity}$, $d_{weighted}$, $d_{interp}$ and $d_{morph}$) calculated with the range of $t_{canopy}$ values, grid sizes and structure element sizes, and the three types of field measurements ($P_{gap}$, FPC or CPC). The comparisons were assessed by calculating RMSE (4) and bias (5), where the error was calculated as the orthogonal distance (6), which allowed for error in the field measurements (Boggs et al., 1988). RMSE and bias were calculated for all samples, and for samples from each separate survey. 95% confidence intervals for RMSE and bias were calculated using a 1000 sample with replacement bootstrap on the error distribution, to show how the different sample numbers for each survey affected the variability in values.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} \varepsilon^2}{n}} \tag{4}$$

$$\text{bias} = \frac{\sum_{i=1}^{n} \varepsilon}{n} \tag{5}$$

$$\varepsilon = \sqrt{\frac{(y_i - \hat{y}_i)^2}{2}} \tag{6}$$

where $y_i$ is the $i$th metric, and $\hat{y}_i$ is the $i$th measurement, from $n$ data points

RMSE was calculated in the same units as the measurements of $P_{gap}$, FPC and CPC, which are proportions ($m^2/m^2$). As previous studies often report RMSE as a percentage, we also convert our RMSE values through dividing by the range of values: because the data ranges from close to zero up to nearly one, the conversion to a percentage is equivalent to multiplying by 100. These percentage values are used in the discussion and conclusion, as unitless proportions often cause confusion when not in the context of presenting results.

#### 3.4.2. Elastic-net regression models

For each comparison between lidar metric and field measurement we also developed a model to assess whether the RMSE and bias could be reduced by accounting for the differences between surveys. The models assumed that the lidar canopy density metrics underestimated or overestimated the field measurements by a proportion ($m$), which was a function of the instrument, survey and plot characteristics calculated from the lidar data. For example, we include the plot 99th percentile height as a predictor to capture the possible bias in TCDM between plots with tall trees and those with shorter trees. Values for $m$ were modelled using elastic net regression, which can deal well with a large number of predictor variables that may be highly correlated (Pearse et al., 2017; Zou and Hastie, 2005). We constructed models with the python implementation of elastic net regression from the scikit-learn module (Friedman et al., 2010; Kim et al., 2007) using the following steps.

1. A set of 14 predictor variables were determined for each value of $m$ that related to the instruments, surveys or were calculated from the lidar data (Table 3). Extra predictors were created through multiplying pairs of predictors to allow for non-linear interactions between predictors. Adding the 105 unique combinations of variables to the 14 individual variables resulted in 119 predictor variables. All predictor variables were centred on zero and standardised.
2. For each of the 265 pairs of metric and field measurements the optimum value of $m$ was determined as the field measurement divided by the metric. The distribution of $m$ values was then centred.
3. Five-fold cross validation was used to select the optimum L1-ratio and alpha parameters required for elastic net regression, where the $m$ values were modelled as the linear combination of the predictor variables.
4. The best model coefficients were determined using all data ($n = 265$).
5. Leave-one-out cross validation (LOOCV) was used to calculate RMSE and bias, representing model prediction error.

The modelling focused on selecting a single lidar TCDM, and a set of instrument/survey/site specific $m$ values that minimised the overall error with field measurements. The alternative, to include all TCDM as predictors and model $P_{gap}$, FPC or CPC directly, was not considered due to the risk of overfitting. The approach also allowed TCDM to be compared, optimum values of $t_{canopy}$ to be determined, and the modelled $m$ values for sites from the 13 different surveys to be compared.

**Table 3**
The 14 instrument, survey and plot parameters, and their ranges, used as predictor variables in the elastic net regression.

| Parameter | Range |
|---|---|
| Laser wavelength (nm) | 1064–1550 |
| Beam divergence (mrad) | 0.22–0.50 |
| Maximum returns per pulse | 4–7 |
| Vertical discrimination distance (m) | 1.0–4.9 |
| Pulse rate (kHz) | 70–240 |
| Flying height above ground (m) | 150–1828 |
| Footprint diameter, nadir, ground (m) | 0.08–0.60 |
| Point density per plot ($m^{-2}$) | 0.91–49.92 |
| Pulse density per plot ($m^{-2}$) | 0.78–23.50 |
| Ground point density per plot ($m^{-2}$) | 0.01–14.83 |
| Median scan angle per plot (°) | 0.9–27.5 |
| 99th percentile of height per plot (m) | 2.4–55.1 |
| Ground slope per plot (°) | 0.0–26.9 |
| Ground relief per plot (m) | 0.02–4.92 |

#### 3.4.3. Multiple output elastic-net regression models

$P_{gap}$, FPC and CPC are related to each other through a set of equations using several other tree stand attributes, such as the proportion of woody material in the canopy ($\alpha$), the plant area index, the stand foliage clumping factor and the leaf angle distribution (Fisher et al., 2018). They are also constrained by simple inequalities: CPC represents solid crowns and must be greater than $1 - P_{gap}$, which does not include within-crown gaps; and $1 - P_{gap}$ must be greater than FPC, which does not have within-crown gaps or woody canopy elements.

Multiple output elastic-net regression from the python scikit-learn module (Friedman et al., 2010; Kim et al., 2007) was used to investigate if the relationships between measurements could improve the models. Rather than model each of the three measurements separately, they were modelled simultaneously from the same set of predictor variables, using the same set of steps as for the single output models. Separate models were run for each of the six lidar TCDM calculated with the range of $t_{canopy}$ values, and grid and structuring element sizes. For each model, a vector of three $m$ values was modelled as the linear combination of the 119 predictor variables, which when multiplied with the lidar TCDM produced the modelled vector of $P_{gap}$, FPC and CPC. We did not consider models that used multiple lidar TCDM as predictor variables due to the risk of over fitting.

### 3.5. Applying the methods to lidar survey data

Determining the best methods for deriving $P_{gap}$, FPC and CPC from airborne lidar was motivated by the need to generate gridded spatial products of tree canopy density from the large archive of lidar data available across Australia. To demonstrate these products, a processing method was developed of the best method for lidar derived FPC and applied to a large ($> 2000\,km^2$) lidar survey over Braidwood, in south eastern NSW (Fig. 1). It was deemed unnecessary to also demonstrate $P_{gap}$ and CPC spatial products, as they are highly correlated to FPC. This survey was acquired between January and November in 2015 with similar specifications as the MAC survey (Table 1) except flying height was 2700 m. The processing method followed the same procedure as described earlier (section 3.3.1), except the TCDM were calculated from lidar returns within square grid cells, to allow the output of raster images. Unlike some vegetation measurements, such as maximum height, FPC is spatially scalable and should not be influenced by the resolution of the grid. A preliminary test of this was conducted by deriving FPC in 100 m and 10 m grid cells, and then comparing the 100 m values to those calculated as the average of the nested 10 m values.

## 4. Results

### 4.1. Lidar derived $P_{gap}$

The lidar TCDM that best fit field measured $1 - P_{gap}$ was $d_{all}$ calculated using $t_{canopy} = 1.5$ m (Fig. 5A). This metric achieved an RMSE with 95% confidence intervals of 0.067 (0.061–0.073) and a bias with 95% confidence intervals of −0.003 (−0.011–0.005). The $d_{intensity}$ metric achieved almost the same RMSE and bias across the values of $t_{canopy}$ as $d_{all}$, but due to its complexities and the assumptions made in its calculations, $d_{all}$ is the preferred metric. The $d_{first}$ metric tended to overestimate $1 - P_{gap}$ with a consistent positive bias across all values of $t_{canopy}$, while $d_{weighted}$ tended to underestimate $1 - P_{gap}$, with a consistent negative bias (Fig. 5A).

The best elastic net regression model for $1 - P_{gap}$ used $d_{all}$ calculated using $t_{canopy}$ of 1.8 m, with an RMSE of 0.065 (0.060–0.071) and bias of 0.002 (−0.060–0.010), although several of the metrics achieved almost the same range of RMSE and bias (Fig. 5B). This also occurred for the multi-output elastic net regression models, where the best model used $d_{first}$ calculated using $t_{canopy}$ of 1.8 m, with an RMSE of 0.063 (0.057–0.069) and bias of 0.003 (−0.004–0.011) (Fig. 5C). An obvious feature of these results is that the models did not achieve lower RMSE and bias values than the best comparison. Nor were they consistently able to reduce the survey specific bias that was observed when the data from the 13 different surveys are tested separately (Fig. 5D). The best comparison between $1 - P_{gap}$ and $d_{all}$ ($t_{canopy} = 1.5$ m) had no bias, but some surveys, such as MAC, had a negative bias −0.083 (−0.110-−0.058) and others, such as SEQ had a positive bias 0.038 (0.015–0.061). Although the elastic net models reduced bias for some surveys, for other surveys it increased.

### 4.2. Lidar derived FPC

The best TCDM compared to FPC was $d_{weighted}$ calculated using $t_{canopy} = 1.7$ m, with an RMSE of 0.060 (0.053–0.067) and a bias of −0.009 (−0.017-−0.003) (Fig. 6A). All the other metrics had a consistent positive bias compared to FPC, across all values of $t_{canopy}$. The best elastic net regression model for FPC used $d_{interp}$, calculated using $t_{canopy} = 2.0$ m and grid = 0.5 m, with an RMSE of 0.057 (0.052–0.062) and bias of 0.005 (−0.002–0.012) (Fig. 6B), while the best multi-output elastic net regression used $d_{first}$ calculated using $t_{canopy} = 1.8$ m, with an RMSE of 0.058 (0.052–0.065) and bias of −0.006 (−0.013–0.001) (Fig. 6C). Although the other metrics achieved very similar results for both the single and multi-output models, the models using $d_{first}$ were more stable than those of $d_{interp}$ across different values of $t_{canopy}$. The models did not lower the RMSE or bias compared to using the metric, nor did they consistently improve on the survey specific bias (Fig. 6D).

### 4.4. Lidar derived CPC

The best TCDM compared to CPC was $d_{first}$ calculated with $t_{canopy} = 0.8$ m, achieving an RMSE of 0.067 (0.060–0.074) and bias of −0.022 (−0.030-−0.014), although $d_{interp}$ calculated with $t_{canopy} = 0.8$ m and grid = 0.5 m achieved a very similar RMSE of 0.070 (0.064–0.077) and had lower bias of −0.011 (−0.019-−0.003) (Fig. 7A). The best elastic net regression for CPC used $d_{first}$ calculated with $t_{canopy} = 1.4$ m, achieving an RMSE of 0.064 (0.056–0.072) and bias of 0.003 (−0.004–0.012) (Fig. 7B), while the best multi-output elastic net regression for CPC used $d_{interp}$ calculated with $t_{canopy} = 1.8$ m and grid = 1.0 m, achieving an RMSE of 0.072 (0.064–0.079) and bias of 0.005 (−0.003–0.013) (Fig. 7C). The models did not improve on the comparison, even when survey specific bias was examined (Fig. 7D).

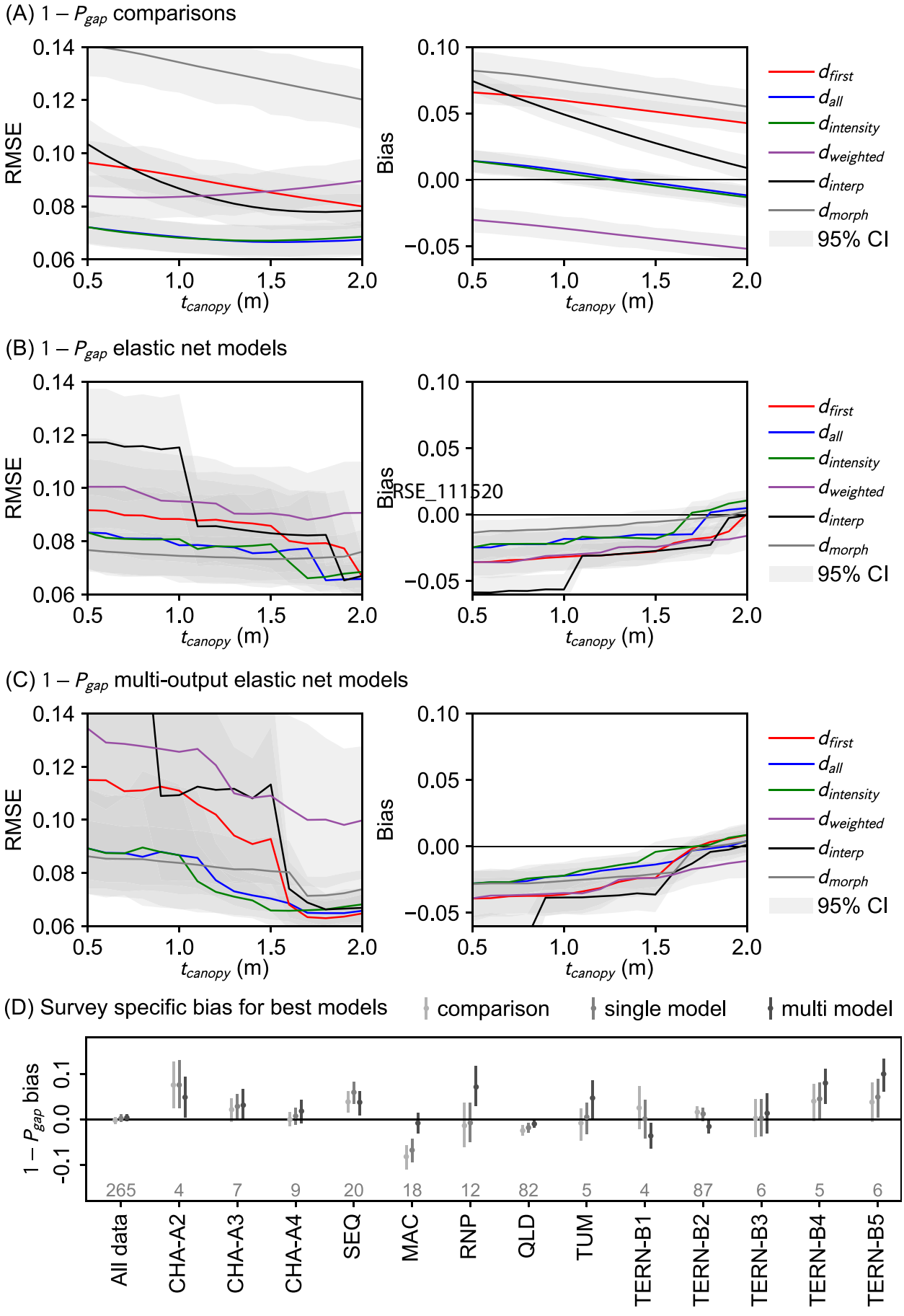
### 4.5. Lidar derived FPC over large areas

An example of lidar derived FPC using $d_{weighted}$ ($t_{canopy} = 1.7$ m) over the Braidwood survey area in south eastern NSW is shown in Fig. 8. This area covers a variety of vegetation types, from woodland with FPC 20–30% (Fig. 8A) to forest with FPC 85–95% (Fig. 8D). As was expected, lidar derived FPC was not sensitive to scaling. FPC calculated within 100 m grid cells was almost identical to the mean of all nested 10 m FPC cells, with an RMSE of only 1% (Fig. 9). Areas where scaling effects were worst contained 10 m pixels with no lidar data, such as water bodies, adjacent to pixels containing vegetation. This effect, and other possible scaling issues that may result from calculating FPC in smaller grid cell sizes and from datasets with lower pulse densities will be examined in future research. Furthermore, it will also be possible to split data into separate flights within a survey and derive FPC for each unique flight date. This will allow an examination of FPC changes over time, and better comparison of lidar FPC with satellite derived products.

## 5. Discussion

### 5.1. Instrument/survey effects on lidar tree canopy density metrics

The first objective was to examine whether different instrument/




*(caption on next page)*

**Fig. 5.** Results of testing different methods of deriving $1 - P_{gap}$ from lidar metrics using comparisons (A), elastic net regression models (B) and multi-output elastic net regression models (C). Survey specific bias for the best models (D) includes sample numbers in grey.

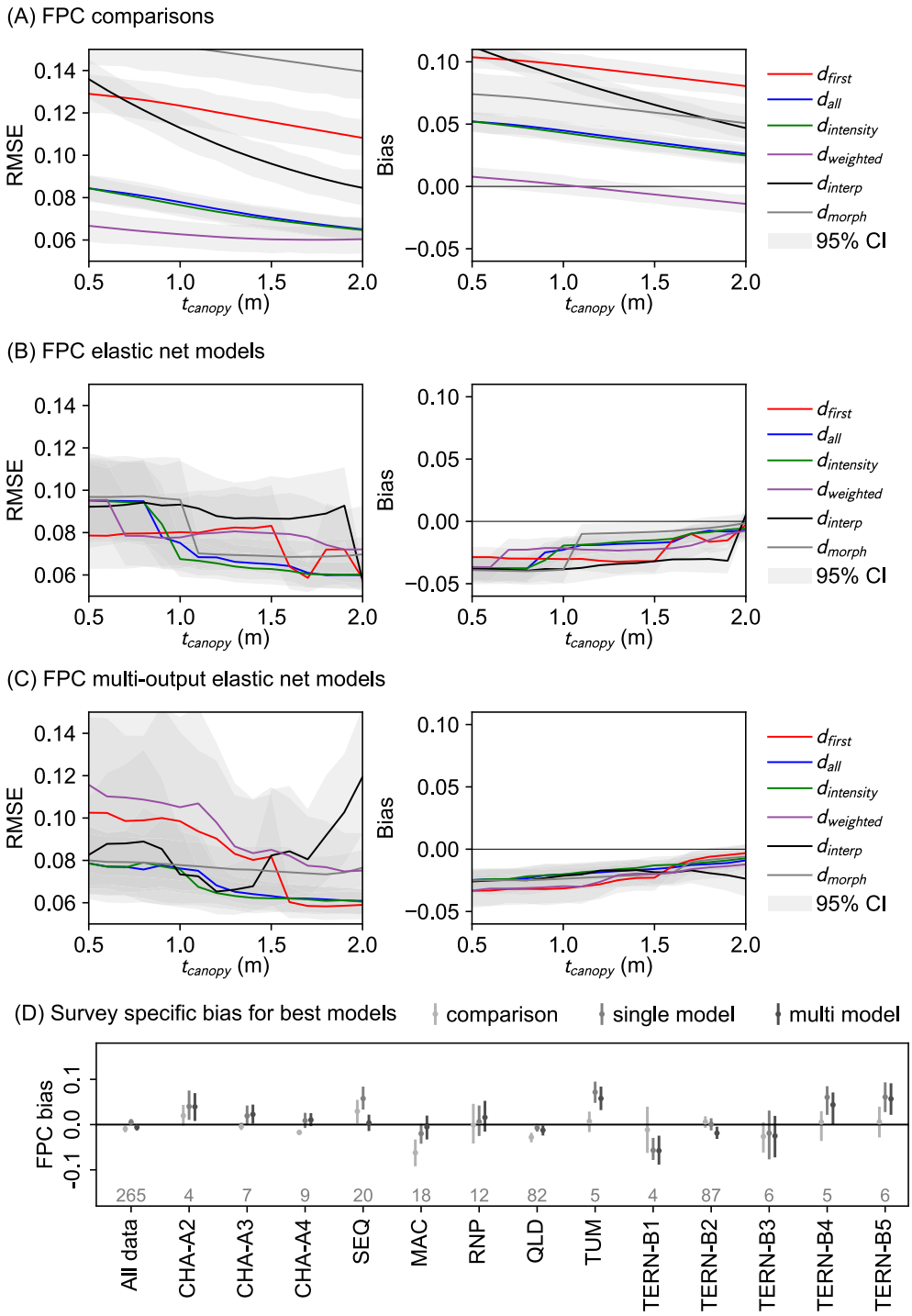
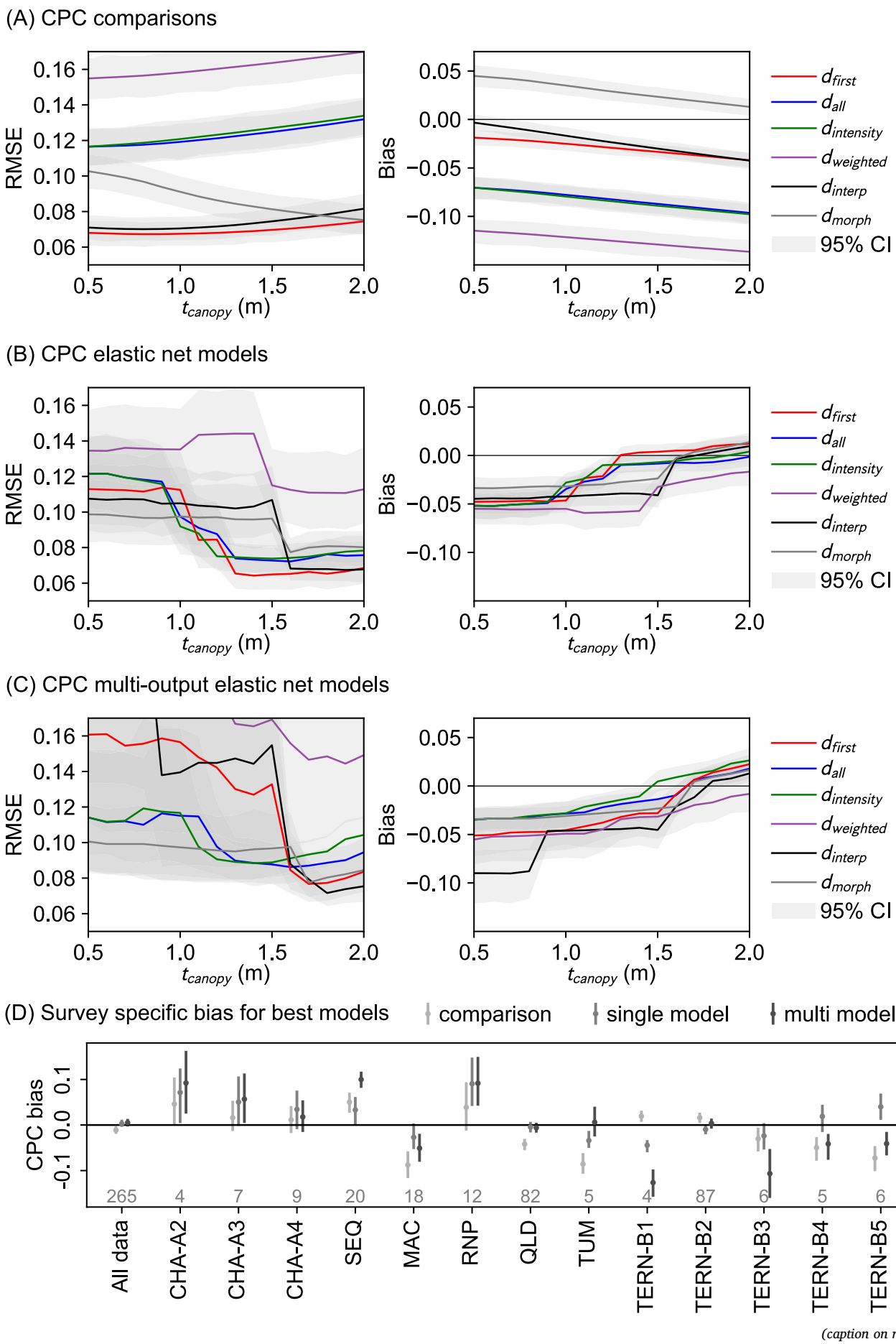



**Fig. 6.** Results of testing different methods of deriving FPC from lidar metrics using comparisons (A), elastic net regression models (B) and multi-output elastic net regression models (C). Survey specific bias for the best models (D) includes sample numbers in grey.

survey parameters biased the relationships between lidar TCDM and field measurements. Figs. 5D, 6D and 7D demonstrate that although the TCDM with the lowest RMSE have very low bias, they have much greater bias when the 13 different surveys are considered separately. This increase in bias may be partly caused by low sample numbers in many of the surveys, but it is also apparent in the surveys with larger sample numbers and must partly be caused by different instrument/survey parameters. The results for the CHA and TERN surveys, flown under several instrument/survey parameters, demonstrate the complexity of the relationships between instrument/survey effects and TCDM bias. The CHA survey metric $d_{all}$ had decreasing bias compared to $1 - P_{gap}$ as flying height and footprint size increased from the A2 to A3 and A4 surveys (Fig. 4D), which was similar to previous comparisons between different metrics and these field measurements (Armston




(caption on next page)

**Fig. 7.** Results of testing different methods of deriving CPC from lidar metrics using comparisons (A), elastic net regression models (B) and multi-output elastic net regression models (C). Survey specific bias for the best models (D) includes sample numbers in grey.

et al., 2013a). This pattern was also seen in the bias between $d_{weighted}$ and FPC (Fig. 5D), and between $d_{interp}$ and CPC (Fig. 6D), and may relate to the increase in the transmission energy losses of the laser pulses (Armston et al., 2013a). The TERN survey TCDM, however, did not show any clear pattern in bias as flying height and footprint size increased from the B1 to B5 surveys (Figs. 4D, 5D and 6D). These contrasting relationships interpreted from the CHA and TERN surveys are difficult to interpret, as they were undertaken using different lidar instruments and only three field sites were measured for each set of varying instrument/survey parameters. Furthermore, while the CHA sites were in woodland with low FPC (5–18%) the TERN sites were in forest with much greater FPC (42–58%).

Elastic net regression models were developed to reduce the bias between field measurements and lidar TCDM, using instrument/survey parameters and plot-based lidar derived characteristics, and their interactive terms, as predictor variables. The models did not effectively decrease the RMSE or bias across all the data, when confidence intervals are considered. Nor were they able to consistently predict the bias across the different surveys, with models decreasing bias in some surveys but increasing bias in others. The inability of the models to remove the bias introduced by different instrument/survey parameters is difficult to explain. It may be that some of the important survey parameters (e.g. flying height) were not recorded accurately across all the surveys, or that important differences in instrument hardware and software that impact the probability of detecting returns were not captured in the parameters. It is also possible that variations in the sampling of the canopy and ground by the lidar pulses due to variations in instrument/sensor parameters occur at a higher frequency than the 100 m diameter plots. For example, scan angle varied by up to 19° within a plot but was only characterised by the median value in the models. Furthermore, it

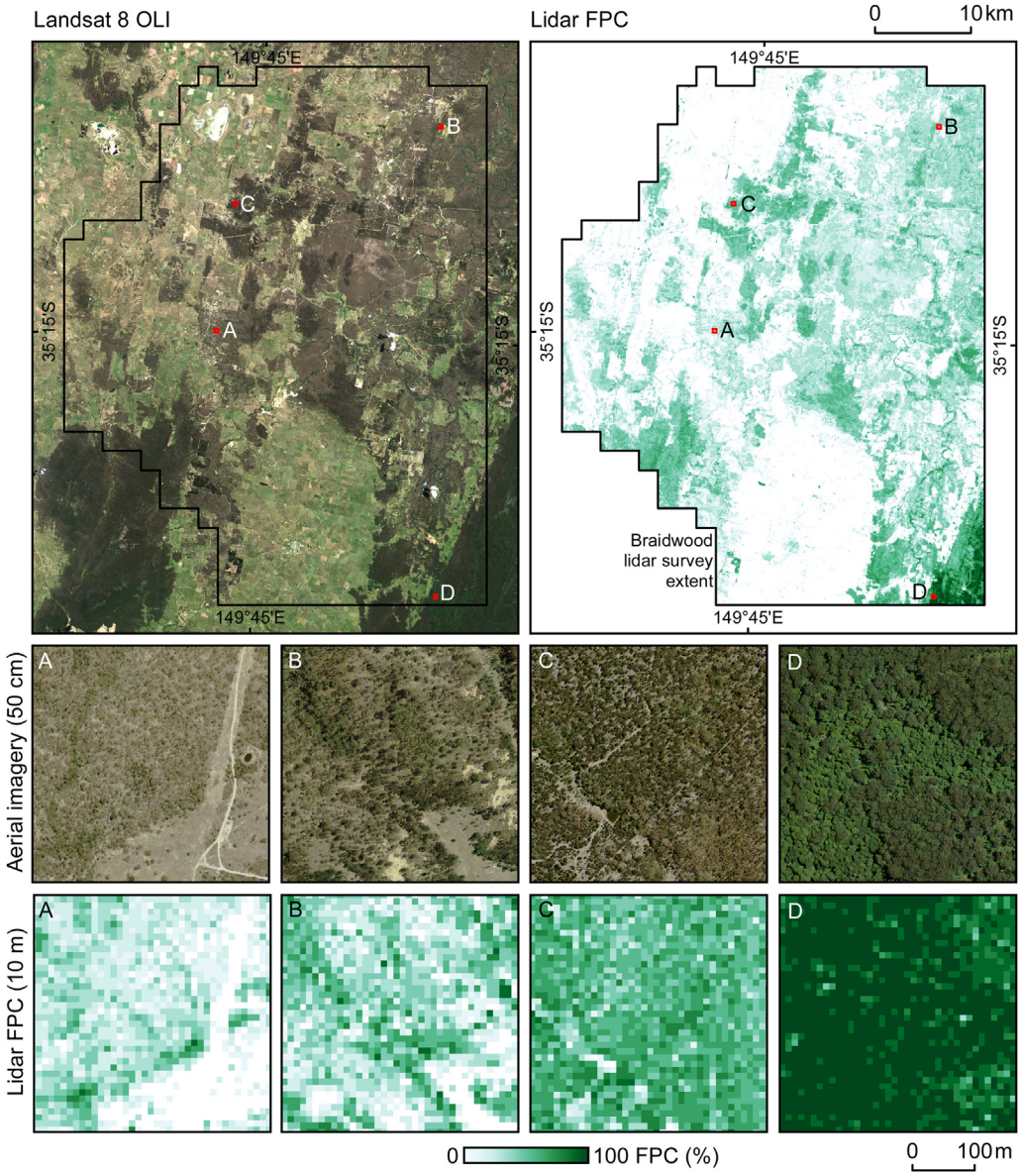



**Fig. 8.** An example of lidar derived FPC using $d_{weighted}$ ($t_{canopy}$ = 1.7 m) for the Braidwood survey in south eastern New South Wales (Fig. 1) acquired between January and November 2015 compared to natural colour Landsat 8 OLI imagery from 2/1/2015 and airborne imagery from 28/1/2014. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

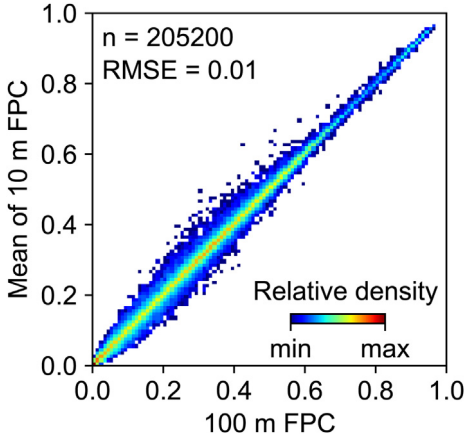



**Fig. 9.** Lidar derived FPC using $d_{weighted}$ ($t_{canopy}$ = 1.7 m) was not sensitive to scaling between 100 and 10 m pixels. The data was from the Braidwood survey (Figs. 1 and 9).

may not be possible to correct for the bias, as the interactions between instrument/survey differences and their sampling of the vegetation and the ground surface are extremely complex.

The accuracy of the field measurements may have contributed to the bias, especially if measurements had bias due to different observers and/or field conditions (Trevithick et al., 2012). The star transect measurements are biased towards the centre of the plot, while the lidar TCDM are calculated from returns spread throughout the plot. Some simulations of this difference in sampling using terrestrial laser scanning data resulted in most modelled FPC values being within 5% of the plot FPC. As this error is very close to the 6–7% RMSE achieved by the comparisons between field measurements and lidar metrics, it highlights the difficulty in decreasing the RMSE through modelling with instrument/survey predictor variables.

The data used to test the relationships between lidar TCDM and field measurements represents many field surveys and airborne campaigns acquired over many years. Analyses of similar datasets of this size, complexity and geographic range have not previously been reported on, and no other published results exist that allow comparison with those presented here. Despite the size of the dataset, the analyses have not provided definitive answers to the questions posed in the objectives. Uncertainties in lidar instrument/survey properties and field measurements have limited the results, and there is no adequate method to remove survey bias from lidar TCDM.

### 5.2. Deriving $P_{gap}$, FPC and CPC from lidar tree canopy density metrics

The second objective was to determine which lidar TCDM or models best equate to the field measurements. Figs. 4–6 show that relatively simple lidar TCDM, calculated with specific canopy height thresholds, offer the best methods of deriving $P_{gap}$, FPC and CPC. The more complex models that tried to correct the bias introduced by different instrument/survey parameters did not reduce the overall RMSE and bias values or consistently reduce the bias observed for the separate surveys. The best TCDM for $1 - P_{gap}$ was $d_{all}$ ($t_{canopy}$ = 1.5 m), the best TCDM for FPC was $d_{weighted}$ ($t_{canopy}$ = 1.7 m), and the best TCDM for CPC was $d_{interp}$ ($t_{canopy}$ = 0.8 m, grid = 0.5 m) (Fig. 10). Each TCDM achieved the lowest RMSE and bias at a unique $t_{canopy}$ value, but the RMSE and bias values were not sensitive to changes in $t_{canopy}$ around these optimum values as shown by the flat curves for the best metrics in Figs. 5A, 6A and 7A. These were all lower than the 2 m used to define trees in the field, which allowed for tree branches and foliage to extend below 2 m, while excluding understorey and mid-storey vegetation.

Using all returns equally for $1 - P_{gap}$ means that $d_{all}$ best approximates the density of all canopy material. Weighting the multiple returns for FPC works as $d_{weighted}$ is lower than $d_{all}$, and it best approximates the foliage canopy elements. Using $d_{first}$ for CPC means that any pulse whose first return is below the canopy is in a gap between trees, and that gaps within crowns are too small to be penetrated. Others have also found that $d_{first}$ was a good proxy for CPC (Fisher et al., 2018; Korhonen et al., 2011). The gridded TCDM of $d_{interp}$ was designed to mimic CPC, and it achieved very similar results to $d_{first}$. As CHMs are often produced from airborne lidar data, they may be a preferable source for deriving CPC.

The TCDM $d_{all}$ and $d_{weighted}$ were found to be the best match for field measured $1 - P_{gap}$ and FPC. In the past, these metrics, which are theoretically affected by the vertical discrimination distances and maximum numbers of returns recorded, have been used with caution (Armston et al., 2013a). However, it appears that for the datasets presented here, these parameters did not influence the TCDM greatly, and the use of all returns in calculating $d_{all}$ and $d_{weighted}$ resulted in better estimates of $1 - P_{gap}$ and FPC than when only first returns were used.

## 6. Conclusions

An extensive dataset of field measurements and airborne lidar surveys across Australia was analysed to test the calibration of lidar tree canopy density metrics against three field measurements of canopy density: canopy gap probability ($P_{gap}$), foliage projective cover (FPC) and crown projective cover (CPC). The best metric to use as a proxy for $1 - P_{gap}$ was the proportion of all returns greater than a canopy height threshold ($t_{canopy}$) of 1.5 m above ground ($d_{all}$), with a root mean square error (RMSE) of 6.7% (95% confidence intervals of 6.1–7.3%). The best metric for FPC was similar to $d_{all}$, but where each return was weighted as the fraction of the number of returns recorded from each pulse ($d_{weighted}$), and where $t_{canopy}$ was 1.7 m above the ground, with a RMSE of 6.0% (5.3–6.7%). The best metric for CPC was found to be the proportion of 0.5 m pixels greater than 0.8 m above the ground, for an interpolated canopy height model, with an RMSE of 7.0% (6.4–7.7%). No consistent relationship could be discerned between lidar instrument/survey parameters and the bias between lidar metric and field measurements. In some surveys the lidar metric overestimated or underestimated most of the field measurements, however, these biases were not consistent with changes in the instrument/survey parameters considered.

The lidar metrics can be applied to most lidar survey data acquired in Australia, that have instrument/survey parameters and vegetation types like those used in this research (Table 1). The errors reported need to be taken into consideration when using lidar derived vegetation metrics as inputs to other models, in change detection between repeat lidar surveys, or when comparing with models derived from satellite data.

The lidar tree canopy density metrics in this research were calculated from data extracted from point clouds coincident with 100 m diameter field plots. Preliminary work comparing gridded products at 10 m and 100 m resolution has shown that the metrics were not scale dependent and can be calculated at different grid sizes, which is useful for different applications, such as comparisons with different resolution satellite imagery. As the archive of airborne lidar data acquired in Australia covers more than 600,000 km$^2$, the gridded metric products will provide a valuable source of information on tree canopy density and facilitate improved environmental monitoring and natural resource management. For example, changes in TCDM between repeat lidar surveys will allow tree clearing and growth to be quantified at high spatial resolution, and provide validation data for satellite monitoring programs.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to

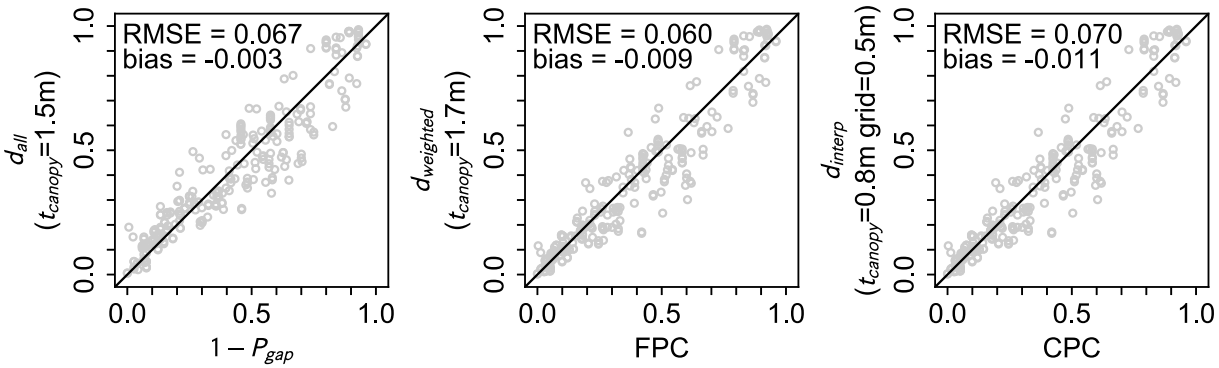



**Fig. 10.** The best matches (lowest root mean square error and bias) between lidar tree canopy density metrics ($d_{all}$, $d_{weighted}$ and $d_{interp}$) calculated according to Table 2 with optimum values of $t_{canopy}$, and field measurements of $P_{gap}$, FPC and CPC.

influence the work reported in this paper.

## Acknowledgements

Thanks to those who acquired and processed the lidar data, including the Department of Finance, Services and Innovation NSW (Spatial Services) and Airborne Research Australia, and to all who helped acquire the field measurements.

## References

Adam, P., King, R.J., 2013. Introduction to royal national park. In: King, R.J. (Ed.), Field Guide to Royal National Park. Linnean Society of New South Wales, Sydney, pp. 1–8.

Armston, J.D., Denham, R.J., Danaher, T.J., Scarth, P.F., Moffiet, T.N., 2009. Prediction and validation of foliage projective cover from Landsat-5 TM and Landsat-7 ETM+ imagery. J. Appl. Remote Sens. 3, 033540. https://doi.org/10.1117/1.3216031.

Armston, J., Scarth, P., Disney, M., Phinn, S., Lucas, R., Bunting, P., Goodwin, N., 2012. Estimation of foliage projective cover, leaf area index and crown cover from airborne lidar across multiple bioregions in eastern Australia. In: Proceedings of the XXII Congress of the International Society for Photogrammetry and Remote Sensing, 25 August - 1 September 2012., Melbourne.

Armston, J., Disney, M., Lewis, P., Scarth, P., Phinn, S., Lucas, R., Bunting, P., Goodwin, N., 2013a. Direct retrieval of canopy gap probability using airborne waveform lidar. Remote Sens. Environ. 134, 24–38. https://doi.org/10.1016/j.rse.2013.02.021.

Armston, J., Newnham, G., Strahler, A., Schaaf, C., Danson, F.M., Gaulton, R., Zhang, Z., Disney, M., Sparrow, B., Phinn, S., Schaefer, M., Burt, A., Counter, S., Erb, A., Goodwin, N., Hancock, S., Howe, G., Johansen, K., Li, A., Lollback, G., Martel, J., Muir, J., Paynter, I., Saenz, E., Scarth, P., Tindall, D., Walker, L.A., Witte, C., Woodgate, W., Wu, D., 2013b. Intercomparison of terrestrial laser scanning instruments for assessing forested ecosystems: a Brisbane field experiment. In: *2013* Fall Meeting, AGU, 9-13 December, San Francisco, *Abstract B11G-0443*.

Armston, J., Bunting, P., Flood, N., Gillingham, S., 2015. PyLidar. http://pylidar.org.

Bater, C.W., Wulder, M.A., Coops, N.C., Nelson, R.F., Hilker, T., Næsset, E., 2011. Stability of sample-based scanning-lidar-derived vegetation metrics for forest monitoring. IEEE Trans. Geosci. Remote Sens. 49, 2385–2392. https://doi.org/10.1109/TGRS.2010.2099232.

Boggs, P.T., Spiegelman, C.H., Donaldson, J.R., Schnabel, R.B., 1988. A computational examination of orthoginal distance regression. J. Econom. 38, 169–201. https://doi.org/10.1016/0304-4076(88)90032-2.

Bunting, P., Armston, J., Gillingham, S., 2015. pynninterp. https://bitbucket.org/petebunting/pynninterp.

Chasmer, L., Hopkinson, C., Treitz, P., 2006. Investigating laser pulse penetration through a conifer canopy by integrating airborne and terrestrial lidar. Can. J. Remote Sens. 32, 116–125. https://doi.org/10.5589/m06-011.

Danaher, T., Scarth, P., Armston, J., Collett, L., Kitchen, J., Gillingham, S., 2010. Remote sensing of tree-grass systems: the eastern Australian woodlands. In: Hill, M.J., Hanan, N.P. (Eds.), Ecosystem Function in Savannas: Measurement and Modeling at Landscape to Global Scales. CRC Press, Boca Raton, pp. 175–194.

Dawson, S., Fisher, A., Lucas, R., Hutchinson, D., Berney, P., Keith, D., Catford, J., Kingsford, R., 2016. Remote sensing measures restoration successes, but canopy heights lag in restoring floodplain vegetation. Remote Sens. 8, 542. https://doi.org/10.3390/rs8070542.

Disney, M.I., Kalogirou, V., Lewis, P., Prieto-Blanco, A., Hancock, S., Pfeifer, M., 2010. Simulating the impact of discrete-return lidar system and survey characteristics over young conifer and broadleaf forests. Remote Sens. Environ. 114, 1546–1560. https://doi.org/10.1016/j.rse.2010.02.009.

Ediriweera, S., Pathirana, S., Danaher, T., Nichols, D., Moffiet, T., 2013. Evaluation of different topographic corrections for Landsat TM data by prediction of foliage projective cover (FPC) in topographically complex landscapes. Remote Sens. 5, 6767–6789. https://doi.org/10.3390/rs5126767.

FAO, 2012. FRA 2015 Terms and Definitions. *Forest Resources and Assessment Working Paper 180*. Food and Agriculture Organization of the United Nations, Rome 36 pp. Available at: http://www.fao.org/docrep/017/ap862e/ap862e00.pdf, Accessed date: 18 July 2014.

Fisher, A., Day, M., Gill, T., Roff, A., Danaher, T., Flood, N., 2016. Large-area, high-resolution tree cover mapping with multi-temporal SPOT5 imagery, New South Wales, Australia. Remote Sens. 8, 515. https://doi.org/10.3390/rs8060515.

Fisher, A., Scarth, P., Armston, J., Danaher, T., 2018. Relating foliage and crown projective cover in Australian tree stands. Agric. For. Meteorol. 259, 39–47. https://doi.org/10.1016/j.agrformet.2018.04.016.

Friedman, J.H., Hastie, T., Tibshirani, R., 2010. Regularization path for generalized linear models by coordinate descent. J. Stat. Softw. 33. https://doi.org/10.18637/jss.v033.i01.

Geoscience Australia, 2018. Elvis - elevation and depth - foundation spatial data. Available at: http://elevation.fsdf.org.au/.

Geosystems, Leica, 2007. Leica ALS50-II Airborne Laser Scanner Product Specifications. Leica Geosystems AG, Heerbrugg, Switzerland.

Geosystems, Leica, 2008. Leica ALS60 Airborne Laser Scanner Product Specifications. Leica Geosystems AG, Heerbrugg, Switzerland.

Gill, T.K., Phinn, S.R., Armston, J.D., Pailthorpe, B.A., 2009. Estimating tree-cover change in Australia: challenges of using the MODIS vegetation index product. Int. J. Remote Sens. 30, 1547–1565. https://doi.org/10.1080/01431160802509066.

Gill, T., Johansen, K., Phinn, S., Trevithick, R., Scarth, P., Armston, J., 2017. A method for mapping Australian woody vegetation cover by linking continental-scale field data and long-term Landsat time series. Int. J. Remote Sens. 38, 679–705. https://doi.org/10.1080/01431161.2016.1266112.

Goodwin, N.R., Coops, N.C., Culvenor, D.S., 2006. Assessment of forest structure with airborne LiDAR and the effects of platform altitude. Remote Sens. Environ. 103, 140–152. https://doi.org/10.1016/j.rse.2006.03.003.

Hansen, E., Gobakken, T., Næsset, E., 2015. Effects of pulse density on digital terrain models and canopy metrics using airborne laser scanning in a tropical rainforest. Remote Sens. 7, 8453–8468. https://doi.org/10.3390/rs70708453.

Hnatiuk, R., Thackway, R., Walker, J., 2009. Vegetation. In: Australian Soil and Land Survey Field Handbook. National Committee on Soil and Terrain, CSIRO Publishing, Melbourne, pp. 73–125.

Hopkinson, C., Chasmer, L., 2009. Testing LiDAR models of fractional cover across multiple forest ecozones. Remote Sens. Environ. 113, 275–288. https://doi.org/10.1016/j.rse.2008.09.012.

Hopkinson, C., Lovell, J., Chasmer, L., Jupp, D., Kljun, N., van Gorsel, E., 2013. Integrating terrestrial and airborne lidar to calibrate a 3D canopy model of effective leaf area index. Remote Sens. Environ. 136, 301–314. https://doi.org/10.1016/j.rse.2013.05.012.

Isenburg, M., 2015. LAStools. https://rapidlasso.com/.

Keith, D., 2013. Vegetation of royal national park. In: King, R.J. (Ed.), Field Guide to Royal National Park. Linnean Society of New South Wales, Sydney, pp. 27–44.

Kim, S.-J., Koh, K., Lustig, M., Boyd, S., Gorinevsky, D., 2007. An interior-point method for large-scale L1-regularized least squares. IEEE Journal of Selected Topics in Signal Processing 1, 606–617. https://doi.org/10.1109/JSTSP.2007.910971.

Korhonen, L., Korpela, I., Heiskanen, J., Maltamo, M., 2011. Airborne discrete-return LIDAR data in the estimation of vertical canopy cover, angular canopy closure and leaf area index. Remote Sens. Environ. 115, 1065–1080. https://doi.org/10.1016/j.rse.2010.12.011.

Korhonen, L., Heiskanen, J., Korpela, I., 2013. Modelling lidar-derived boreal forest canopy cover with SPOT 4 HRVIR data. Int. J. Remote Sens. 34, 8172–8181. https://doi.org/10.1080/01431161.2013.833361.

Korpela, I., Hovi, A., Morsdorf, F., 2012. Understory trees in airborne LiDAR data — selective mapping due to transmission losses and echo-triggering mechanisms. Remote Sens. Environ. 119, 92–104. https://doi.org/10.1016/j.rse.2011.12.011.

Lim, K., Hopkinson, C., Treitz, P., 2008. Examining the effects of sampling point densities on laser canopy height and density metrics. For. Chron. 84, 876–885. https://doi.org/10.5558/tfc84876-6.

Lovell, J.L., Jupp, D.L.B., Culvenor, D.S., Coops, N.C., 2003. Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests. Can. J. Remote Sens. 29, 607–622. https://doi.org/10.5589/m03-026.

Lucas, R., Cronin, N., Moghaddam, M., Lee, A., Armston, J., Bunting, P., Witte, C., 2006.

Integration of radar and Landsat-derived foliage projected cover for woody regrowth mapping, Queensland, Australia. Remote Sens. Environ. 100, 388–406. https://doi.org/10.1016/j.rse.2005.09.020.
Lucas, R., Bunting, P., Paterson, M., Chisholm, L., 2008. Classification of Australian forest communities using aerial photography, CASI and HyMap data. Remote Sens. Environ. 112, 2088–2103. https://doi.org/10.1016/j.rse.2007.10.011.
Melin, M., Korhonen, L., Kukkonen, M., Packalen, P., 2017. Assessing the performance of aerial image point cloud and spectral metrics in predicting boreal forest canopy cover. ISPRS J. Photogrammetry Remote Sens. 129, 77–85. https://doi.org/10.1016/j.isprsjprs.2017.04.018.
Montaghi, A., 2013. Effect of scanning angle on vegetation metrics derived from a nationwide Airborne Laser Scanning acquisition. Can. J. Remote Sens. 39, S152–S173. https://doi.org/10.5589/m13-052.
Morsdorf, F., Kötz, B., Meier, E., Itten, K.I., Allgöwer, B., 2006. Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. Remote Sens. Environ. 104, 50–61. https://doi.org/10.1016/j.rse.2006.04.019.
Morsdorf, F., Frey, O., Meier, E., Itten, K.I., Allgöwer, B., 2008. Assessment of the influence of flying altitude and scan angle on biophysical vegetation products derived from airborne laser scanning. Int. J. Remote Sens. 29, 1387–1406. https://doi.org/10.1080/01431160701736349.
Muir, J., Schmidt, M., Tindall, D., Trvithick, R., Scarth, P., Stewart, J., 2011. Field measurement of fractional ground cover: a technical handbook supporting ground cover monitoring for Australia: prepared by the Queensland department of environment and resource management for the Australian bureau of agricultural and resource economics and sciences, canberra. Available at: http://www.daff.gov.au/abares/aclump/pages/publications.aspx#landcover, Accessed date: 8 August 2014.
Ni-Meister, W., Jupp, D.L.B., Dubayah, R., 2001. Modeling lidar waveforms in heterogeneous and discrete canopies. IEEE Trans. Geosci. Remote Sens. 39, 1943–1958.
Næsset, E., 2009. Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. Remote Sens. Environ. 113, 148–159. https://doi.org/10.1016/j.rse.2008.09.001.
Pearse, G.D., Morgenroth, J., Watt, M.S., Dash, J.P., 2017. Optimising prediction of forest leaf area index from discrete airborne lidar. Remote Sens. Environ. 200, 220–239. https://doi.org/10.1016/j.rse.2017.08.002.
Riaño, D., Valladares, F., Condés, S., Chuvieco, E., 2004. Estimation of leaf area index and covered ground from airborne laser scanner (Lidar) in two contrasting forests. Agric. For. Meteorol. 124, 269–275. https://doi.org/10.1016/j.agrformet.2004.02.005.
Roussel, J.-R., Caspersen, J., Béland, M., Thomas, S., Achim, A., 2017. Removing bias from LiDAR-based estimates of canopy height: accounting for the effects of pulse density and footprint size. Remote Sens. Environ. 198, 1–16. https://doi.org/10.1016/j.rse.2017.05.032.
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 Sens. 11, 147. https://doi.org/10.3390/rs11020147.
Schliemann, S.A., Bockheim, J.G., 2011. Methods for studying treefall gaps: a review. For. Ecol. Manag. 261, 1143–1151. https://doi.org/10.1016/j.foreco.2011.01.011.
Sibson, R., 1981. A brief description of natural neighbor interpolation. In: Barnett, V. (Ed.), Interpolating Multivariate Data. John Wiley, Chichester, pp. 21–36.
Solberg, S., Brunner, A., Hanssen, K.H., Lange, H., Næsset, E., Rautiainen, M., Stenberg, P., 2009. Mapping LAI in a Norway spruce forest using airborne laser scanning. Remote Sens. Environ. 113, 2317–2327. https://doi.org/10.1016/j.rse.2009.06.010.
Specht, R.L., 1983. Foliage projective covers of overstory and understory strata of mature vegetation in Australia. Austral Ecol. 8, 433–439.
Specht, R.L., Specht, A., 1999. Australian Plant Communities: Dynamics of Structure , Growth and Biodiversity. Oxford University Press, Melbourne, Australia.
TERN AusCover, 2018. Vegetation height and structure - derived from ALOS-1 PALSAR, Landsat and ICESat/GLAS, Australia coverage. The AusCover facility of the terrestrial ecosystem research Network. .
Tickle, P.K., Lee, A., Lucas, R.M., Austin, J., Witte, C., 2006. Quantifying Australian forest floristics and structure using small footprint LiDAR and large scale aerial photography. For. Ecol. Manag. 223, 379–394. https://doi.org/10.1016/j.foreco.2005.11.021.
Trevithick, R., 2017. SLATS Star Transect Data. AusCover Remote Sesing Data Facility. Terrestrial Ecosystem Research Network Available at: http://data.auscover.org.au/Portal2/.
Trevithick, R., Muir, J., Denham, R., 2012. The effect of observer experience levels on the variability of fractional ground cover reference data. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. ISPRS Congress, Melbourne, Australia.
Vepakomma, U., St-Onge, B., Kneeshaw, D., 2008. Spatially explicit characterization of boreal forest gap dynamics using multi-temporal lidar data. Remote Sens. Environ. 112, 2326–2340. https://doi.org/10.1016/j.rse.2007.10.001.
Wang, Y., Weinacker, H., Koch, B., 2008. A lidar point cloud based procedure for vertical canopy structure analysis and 3D single tree modelling in forest. Sensors 8, 3938–3951. https://doi.org/10.3390/s8063938.
Weller, D., Denham, R., Witte, C., Mackie, C., Smith, D., 2003. Assessment and monitoring of foliage projected cover and canopy height across native vegetation in Queensland, Australia, using laser profiler data. Can. J. Remote Sens. 29, 578–591. https://doi.org/10.5589/m03-028.
Wilkes, P., Jones, S.D., Suarez, L., Haywood, A., Woodgate, W., Soto-Berelov, M., Mellor, A., Skidmore, A.K., 2015. Understanding the effects of ALS pulse density for metric retrieval across diverse forest types. Photogramm. Eng. Remote Sens. 81, 625–635. https://doi.org/10.14358/pers.81.8.625.
Wulder, M.A., Bater, C.W., Coops, N.C., Hilker, T., White, J.C., 2008. The role of LiDAR in sustainable forest management. For. Chron. 84, 807–826. https://doi.org/10.5558/tfc84807-6.
Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W., Gobakken, T., 2012. Lidar sampling for large-area forest characterization: a review. Remote Sens. Environ. 121, 196–209. https://doi.org/10.1016/j.rse.2012.02.001.
Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B 67, 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x.