Estimating vegetation structure and aboveground carbon storage in Western Australia using GEDI LiDAR, Landsat and Sentinel data

Natasha Lutz, Pedro Rodriguez-Veiga, Imma Oliveras Menor

PAPER • OPEN ACCESS

Estimating vegetation structure and aboveground carbon storage in Western Australia using GEDI LiDAR, Landsat and Sentinel data

To cite this article: Natasha Lutz *et al* 2024 *Environ. Res.: Ecology* **3** 045004

ENVIRONMENTAL RESEARCH
ECOLOGY

OPEN ACCESS

RECEIVED
28 March 2024

REVISED
13 September 2024

ACCEPTED FOR PUBLICATION
25 September 2024

PUBLISHED
8 November 2024

PAPER

Estimating vegetation structure and aboveground carbon storage in Western Australia using GEDI LiDAR, Landsat and Sentinel data

Natasha Lutz[1,*], **Pedro Rodriguez-Veiga**[3,4] **and Imma Oliveras Menor**[1,2]

[1] Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, United Kingdom
[2] AMAP (Botanique et Modélisation de l'Architecture des Plantes et des Végétations), CIRAD, CNRS, INRA, IRD, Université de Montpellier, Montpellier, France
[3] Sylvera Ltd, London, United Kingdom
[4] School of Geography, Geology and the Environment, University of Leicester, Leicester, United Kingdom
[*] Author to whom any correspondence should be addressed.

E-mail: natasha.lutz@ouce.ox.ac.uk

Keywords: GEDI, canopy height, vegetation structure, aboveground biomass, Random forest

Supplementary material for this article is available online

Abstract

Worsening climate change impacts are amplifying the need for accurate estimates of vegetation structure and aboveground biomass density (AGBD) to assess changes in biodiversity and carbon storage. In Australia, increasing wildfire frequency and interest in the role of forests in the carbon cycle necessitates biomass mapping across large geographic extents to monitor forest change. The availability of spaceborne Light Detection and Ranging optimised for vegetation structure mapping through the Global Ecosystem Dynamics Investigation (GEDI) provides an opportunity for large-scale forest AGBD estimates of higher accuracy. This study assessed the use of the GEDI canopy height product to predict woody AGBD across five vegetation types in Western Australia: tall eucalypt forests, eucalypt open−woodlands, low-lying heathland, tropical eucalypt savannas, and tussock and hummock grasslands. Canopy height models were developed using random forest regressions trained on GEDI canopy height discrete point data. Predictor variables included spectral bands and vegetation indices derived from synthetic aperture radar Sentinel−1 data, and multispectral Landsat and Sentinel−2 data. AGBD was subsequently estimated using power-law models derived by relating the predicted canopy heights to field AGBD plots. Mapping was conducted for 2020 and 2021. The accuracy of canopy height predictions varied with height quantiles; models underestimated the height of taller trees and overestimated the height of smaller trees. A similar underestimation and overestimation trend was observed for the AGBD estimates. The mean carbon stock was estimated at 69.0 $\pm$ 12.0 MgCha^{-1} in the tall eucalypt forests of the Warren region; 33.8 $\pm$ 5.0 MgCha^{-1} for the open eucalypt woodlands in the South Jarrah region; 7.1 $\pm$ 1.4 MgCha^{-1} for the heathland and shrublands in the Geraldton Sandplains region; 43.9 $\pm$ 4.9 MgCha^{-1} for the Kimberley eucalypt savanna; and 3.9 $\pm$ 1.0 MgCha^{-1} for the Kimberley savanna grasslands. This approach provides a useful framework for the future development of this process for fire management, and habitat health monitoring.

1. Introduction

Forests, woodlands, and savannas provide a range of critical ecosystem and climate-regulating services. Globally, forests sequester approximately 2.6 Gt CO_2 per year through photosynthesis and tree growth (Tubiello *et al* 2021). Forests also buffer microclimate temperatures under the canopy, support biodiversity by providing habitat and food, and influence disturbance regimes (Gao *et al* 2014, De Frenne *et al* 2019). However, shifts in temperature and precipitation associated with anthropogenic climate change are expected to reduce taxonomic diversity and vegetation health and alter carbon dynamics through events such as

wildfires of increased frequency and severity (Chuvieco *et al* 2021, Balch *et al* 2022, Díaz and Malhi 2022). Amid these worsening climate change and biodiversity impacts, future forest management requires accurate quantification of existing carbon stocks (Filbee-Dexter and Wernberg 2020), estimations of vegetation degradation (Viana-Soto *et al* 2022), and mapped changes in the distribution of plant functional types (Guo *et al* 2017). Further, a growing Australian carbon market necessitates accurate baseline carbon estimates to ensure carbon credit integrity is maintained in national and international markets (Renwick *et al* 2014, Axelsson *et al* 2024). Carbon markets provide opportunities for Aboriginal and Torres Strait Islander stewardship and associated finance flows to climate frontline communities (Renwick *et al* 2014, Russell-Smith *et al* 2015). Inadequate monitoring and verification of carbon stocks undermines trust in a market that may afford significant opportunities for sustainable development and biodiversity co-benefits. Estimates of stored carbon on regional landscape scales currently remain marred by considerable uncertainty (Hill *et al* 2013, Girardin *et al* 2021, Friedlingstein *et al* 2023).

Biomass consists of aboveground biomass (AGB), below-ground biomass, deadwood, and soil. Vegetation biomass assessments typically focus on AGB and exclude the biomass of roots and other below-ground components due to the challenges of obtaining these measurements (Lu *et al* 2016). In forest and woodland ecosystems, AGB is an indicator of forest productivity and contributes significantly to the overall carbon storage in vegetation, and so forms a key part of carbon accounting for emissions removal projects (Calders *et al* 2015, Standish and Prober 2020). AGB can be estimated either in absolute terms (Mg or tonnes) or as above-ground biomass density (AGBD) expressed as megagrams per hectare (Mg ha^{-1}). Carbon stock in Australia is generally estimated to be approximately 50% of the AGB stock for a given area (Burrows *et al* 2002, Volkova *et al* 2015). In woody vegetation, canopy-top height is a useful predictor for AGB, carbon stock and habitat heterogeneity (Simard *et al* 2011, Nandy *et al* 2021, Lang *et al* 2023). AGBD and canopy height therefore form key variables of interest to environmental policymakers focusing on climate change mitigation and biodiversity protection.

AGB estimates have been traditionally obtained from *in situ* destructive or non-destructive field measurements, which are time and resource-costly, logistically challenging, and are restricted to measurements at the tree or plot level (McCaw *et al* 1996, Stovall *et al* 2018, Brede *et al* 2022). Destructive sampling provides accurate estimates but requires the felling of the trees involved (Catchpole and Wheeler 1992). Non-destructive field sampling techniques have conventionally involved the use of statistical allometric models; however, the site or species-specific nature of these models can limit their applicability over large regions (Calders *et al* 2015). More recently, non-destructive methods have also included indirect measurements of AGBD derived from proximal Light Detection and Ranging (LiDAR) (Morais *et al* 2021). High-density LiDAR point clouds from ground-collected terrestrial laser scanning, and airborne laser scanning systems such as airplanes or unmanned aerial vehicle laser scanners (UAV-LS) such as drones, have been increasingly deployed as methods to characterize three-dimensional (3D) forest structures such as diameter at breast height, canopy height and canopy coverage (Calders *et al* 2015, Wang *et al* 2021, Brede *et al* 2022, Demol *et al* 2022). However, while ground and airborne LiDAR sensors produce relatively accurate AGBD estimates, they have high operational costs, often rendering a regional coverage uneconomical (Erdody and Moskal 2010, Stojanova *et al* 2010).

Broad-scale forest structure characterization is typically achieved using remote sensing methods (Goetz and Dubayah 2011, Montesano *et al* 2013, Brigot *et al* 2019, Salum *et al* 2020), which provide a cost-effective method of assessing the biomass of forests and woodlands at the scales needed to monitor vegetation changes with climate. A range of sensors, including passive multi-spectral sensors (Li *et al* 2020a), active sensors such as synthetic aperture radar (SAR) (Mitchard *et al* 2009), and space-borne LiDAR (Narine *et al* 2020) have been used to map AGBD over a range of scales and spatial resolutions. Previous studies have shown that utilising a combination of these active and passive sensors improves the accuracy of vegetation structure and biomass characterisation, as each data source presents different strengths and limitations (Montesano *et al* 2013, Ediriweera *et al* 2014, Rodríguez-Veiga *et al* 2019, Silva *et al* 2021, Li *et al* 2022a, Shendryk 2022, Guo *et al* 2023). Passive optical sensors such as Sentinel−2 and Landsat detect biochemical vegetation attributes through surface reflectance (Chaves *et al* 2020). The spectral indices, or band ratios, from these multispectral sensors can be useful for detecting vegetation attributes such as green biomass, density and canopy architecture through photosynthetic absorbance (Tucker 1979, Ustin *et al* 2004). Sentinel−2 multispectral imagery contains additional red-edge bands that capture plant stress or growth through sensitivities to chlorophyll levels (Forkuor *et al* 2018, Li *et al* 2020b, Hua and Zhao 2021). Landsat multispectral data provides the most comprehensive time-series of earth observation data. However, both Landsat and Sentinel−2 multispectral sensors are sensitive to cloud cover (Li *et al* 2020b). In contrast, Sentinel−1 is an active SAR sensor capable of penetrating clouds and some tree canopies. SAR instruments such as Sentinel−1 determine vegetation structure by emitting microwave signals that elucidate the target surface using the scattering of the returning echo signals, or back-scatter (Ferro-Famil and Pottier 2016).

However, Landsat, Sentinel−2, and Sentinel−1 sensors all tend to underestimate biomass in dense forests due to signal saturation and low canopy penetration (Luckman *et al* 1998, Joshi *et al* 2017), and overestimate in sparse open forests due to an inability to distinguish between spectral reflectance from understorey and canopy vegetation strata (Li *et al* 2020b, Wang *et al* 2021).

Spaceborne LiDAR offers an opportunity to improve ecosystem mapping at regional and global scales (Liu and Popescu 2022, Mulverhill *et al* 2022). NASA's Global Ecosystem Dynamics Investigation (GEDI) is the first spaceborne LiDAR dedicated to monitoring vertical vegetation structure; capable of penetrating up to 99% canopy cover (Hancock *et al* 2019). This improved technical capability enables more accurate and widespread AGBD and vegetation structure mapping, particularly in dense forests where AGBD estimation is typically associated with considerable uncertainty (Duncanson *et al* 2022, Leite *et al* 2022). GEDI data is currently available from 2019 to 2023, with data collection recommencing in 2024 after a brief period of hibernation. The GEDI sensor is aboard the International Space Station (ISS) and collects full-waveform LiDAR (Dubayah *et al* 2020) across tropical, subtropical and temperate biogeographical zones between the latitudes of ∼51.6° north and south (Hancock *et al* 2019). As the GEDI sensor collects transects of discrete circular footprints (snapshots of ∼25 m diameter) separated by unsampled geographic extents, it is necessary to combine it with co-located Landsat, and Sentinel−1 and Sentinel−2 remote sensing imagery to achieve complete coverage (wall-to-wall) canopy height or AGBD maps (Silva *et al* 2021, Shendryk 2022).

A range of parametric modelling and non-parametric machine learning methods have been used in the estimation of canopy height and AGBD and to quantify the associated uncertainties (Fassnacht *et al* 2016). Parametric models such as multiple regression or Ordinary Least Squares (OLS) have been found to be effective, and afford easy output interpretation due to observable model equations (Clark *et al* 2011, Ferraz *et al* 2016, Duncanson *et al* 2022). Other studies have found non-parametric models outperform statistical regression models in their ability to handle non-normally distributed input data and complex non-linear relationships between variables (Hojo *et al* 2023, Wu *et al* 2023). Commonly used machine learning methods include k-nearest neighbour (Jiang *et al* 2022), support vector regression (Gleason and Im 2012), decision tree algorithms such as random forest (RF) (Baccini *et al* 2008, Powell *et al* 2010, Leite *et al* 2022) and more recently, deep learning methods such as neural networks (Mugabowindekwe *et al* 2023, Seely *et al* 2023). Comparable results from both parametric models and machine learning models have also been achieved (Tang *et al* 2021). The applicability of various modelling approaches varies depending on the scale of analysis, and the type and amount of data available. The RF algorithm is one of the most widely used modelling approaches for ecological remote sensing applications, in part due to its ability to model complex interactions between input variables and robustness to outliers (Wang *et al* 2016, Morais *et al* 2021).

Recent studies have utilised GEDI with a data fusion method to estimate canopy height globally (Lang *et al* 2022, 2023) and regionally (Rishmawi *et al* 2021, Dhargay *et al* 2022, Shendryk 2022), and to map canopy cover and height diversity (Schneider *et al* 2020). GEDI metrics have also been used to quantify AGBD (Saarela *et al* 2018). However, the irregular spatial and temporal coverage of the ISS's precessing orbit means it is unlikely that GEDI LiDAR footprints can be spatially matched opportunistically with existing field biomass inventory plots (Bullock *et al* 2023). To overcome this, previous studies have simulated GEDI waveforms from UAV-LiDAR spatially coincident with field plot AGBD data, before applying this relationship to on-orbit GEDI data across a region (Hancock *et al* 2019, Leite *et al* 2022). While effective, this approach is limited to the small regions where field data plots have also been overflown with airborne LiDAR. The GEDI mission team produces a global 1 km−spatial resolution AGBD product; however, this may be of limited applicability at the finer scales of carbon projects or local policy implementation (Duncanson *et al* 2022). Field AGBD inventory data or airborne LiDAR are lacking in many regions of the world due to location remoteness or the cost of such data collection. Therefore, it is necessary to develop methods of effectively utilising GEDI height metrics to estimate AGBD given existing field data availability or limited calibration plots.

The utilisation of GEDI to estimate AGBD presents an opportunity to improve the accuracy of vegetation mapping on regional scales in Australia. Current AGBD estimation methods in Australia for National Inventory Reporting under the UNFCCC use the Full Carbon Accounting Model (Brack and Richards 2002, Paul and Roxburgh 2020), for which spatial mapping uses passive optical Landsat satellite data and a RF algorithm trained on field AGBD data (Roxburgh *et al* 2019). While airborne LiDAR data and field campaigns have been used to map vegetation across Australia (Liao *et al* 2020), the use of currently available spaceborne LiDAR such as GEDI to map vegetation structure and AGBD is scarce (Potapov *et al* 2021, Duncanson *et al* 2022, Shendryk 2022).

The overall aim of this study was to develop a method of estimating vegetation structure, AGBD and carbon stocks at the regional scale, utilising a fusion of the best freely available passive and active remote sensing data and spaceborne GEDI LiDAR, and given limitations on the availability of existing biomass calibration data. This method was applied to estimate canopy height and AGBD across four biogeographic

regions of considerable carbon, biodiversity, cultural and economic value in Western Australia, namely the Warren, South Jarrah, Geraldton Sandplains and Kimberley regions (Wardell-Johnson *et al* 2017, Jackson *et al* 2017, IBRA 2018). Across these regions, five separate models were implemented to account for different vegetation types; (i) tall eucalypt forests in the Warren region; (ii) mid−open eucalypt forests and woodlands in the South Jarrah region; (iii) low heathlands and shrublands in the Geraldton Sandplains; and within the Kimberley region, (iv) eucalypt savanna, and (v) tussock and hummock grasslands. This workflow overcomes a lack of extensive field AGBD data or co-located airborne LiDAR data by first creating RF canopy height models (CHMs) for each of the five vegetation types using GEDI height metrics and spatially continuous Landsat, Sentinel and Shuttle Radar Topography Mission (SRTM) topographical data. Power-law models were then derived for each vegetation type by relating the mean canopy height from the CHM with existing co-located field plot data. This approach was undertaken for two separate years, 2020 and 2021, to verify the consistency of the results given the irregular orbit track of GEDI. The results of this method were then assessed to determine whether they are sufficiently accurate to be useful in guiding climate mitigation policy and ecological monitoring.

2. Study areas

Vegetation structure and AGBD mapping were undertaken in four Interim Biogeographic Regionalisation (IBRA, Version 7) regions of Western Australia: the Warren region and South-Jarrah region in the south, Geraldton Sandplains region in the mid-west, and Kimberley region in the north (IBRA 2018) (figure 1). The tall eucalypt forests and open woodlands in the south-west of Western Australia present some of the tallest and most carbon-dense forests in Australia and are important hotspots for carbon and biodiversity conservation (Wardell-Johnson *et al* 2017). The open kwongan heathlands and shrublands and Wandoo woodlands of the Geraldton Sandplains region in the mid-west also occur within a region of high biodiversity and species endemism (Griffin 1994). The tropical savanna and grassland ecosystems in the north of Western Australia constitute some of the most structurally intact and fire-prone ecosystems in the state (Russell-Smith *et al* 2015). They are consequently the subject of carbon abatement projects, an emissions avoidance methodology, and numerous Indigenous cultural burning projects (Russell-Smith *et al* 2015, Yates *et al* 2023). They also currently form some of the most economically significant carbon stores in Western Australia (Russell-Smith *et al* 2015).

2.1. Warren region

The Warren IBRA region covers 844, 270 ha in the south-west region of Western Australia. Vegetation in the region is characterised by extensive tall Karri (*Eucalyptus diversicolor*) forests, often interspersed with Marri (*Corymbia calophylla*), open forests with Jarrah (*Eucalyptus marginata*), and pockets of tingle forest featuring Yellow Tingle (*Eucalyptus guilfoylei*), Red Tingle (*Eucalyptus jacksonii*) and coastal shrublands (figures 1 and 2). Canopy cover is generally between 30%–70% (Specht 1970), with recorded tree heights as tall as 83.3 m (Keith 2017). The Warren region experiences a Mediterranean climate and is defined by the high rainfall zone of southwest Australia, with annual rainfall around 1200 mm (BoM 2024). Approximately 68% of the region is legislatively protected as conservation reserve tenure managed by the state government's conservation agency (DBCA 2023a). Primary vegetation disturbance in the region is through prescribed burning and wildfires, and infestations of the soil-borne pathogen *Phytophthora cinnamomi* (Boer *et al* 2009, McDougall *et al* 2024).

2.2. South Jarrah region

The South Jarrah IBRA bioregion covers 2, 608, 550 ha in the south-west of Western Australia. Vegetation is dominated by eucalypt mid-open forest and woodlands, with *E. marginata, C. calophylla, and Eucalyptus aubangusta* (figures 1 and 2). Western regions of the Jarrah forests receive rainfall greater than 1200 m due to the orographic uplift along the Darling escarpment (McCaw *et al* 2011) (figure 1). Approximately 37% of the region is legislatively protected as conservation reserve tenure managed by the state government's conservation agency (DBCA 2023a). The flora and fauna in the region display a high degree of endemism (Rix *et al* 2015). Similar to the Warren region, the South-Jarrah forests and woodlands are subject to frequent wildfires and *P. cinnamomi* infestation (Boer *et al* 2009, McDougall *et al* 2024).

2.3. Geraldton Sandplains region

The Geraldton Sandplains region covers an area of 1, 171, 842 ha in the mid-west region of Western Australia. The region is characterized by kwongan heathlands, *Acacia* shrublands and scattered eucalypt woodlands with species including *Eucalyptus wandoo* (NVIS Technical Working Group 2017). Annual rainfall in the region is 498 mm (BoM 2024) (figures 1 and 2). Heathlands are dominated by low shrubs up

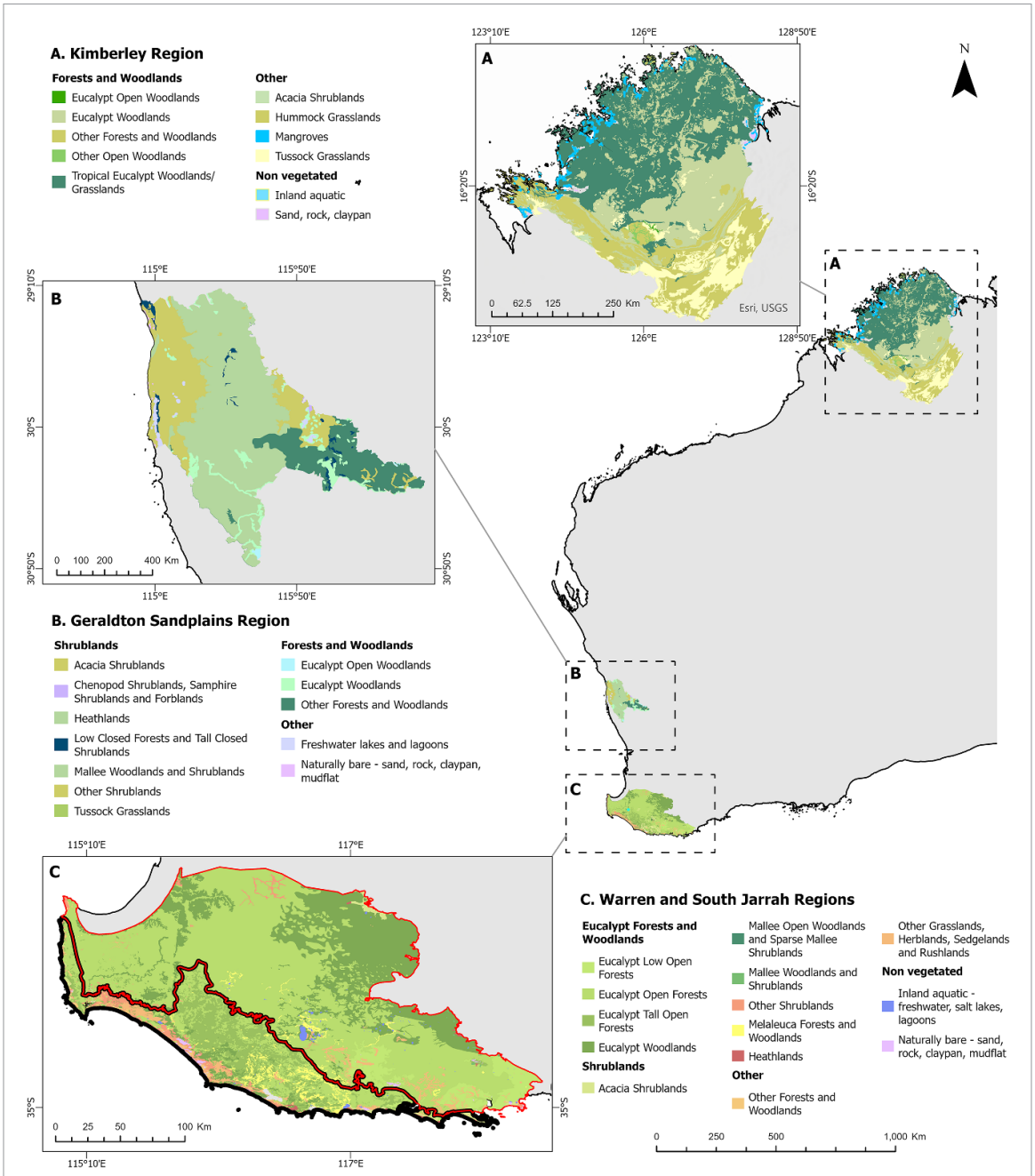

Figure 1. Different vegetation types found within the (A) Kimberley (B) Geraldton Sandplains and (C) Warren (within black boundary) and South Jarrah (within the red boundary) IBRA Regions in Western Australia, with the National Vegetation Information System, (NVIS) Major Vegetation Groups, Version 6 (NVIS Technical Working Group 2017).

to 2 m tall, with trees rarely exceeding 10 m tall (Keith 2017). Most biomass is contributed by *Banksia, Allocasuarina*, and *Xanthorrhoea* species (Keith 2017). *Acacia* shrublands include trees less than 4 m tall, dominated by *Banskia attenuata, Nuytsia floribunda*, and *Allocasuraina* low shrubland. Approximately 23% of the region is legislatively protected as conservation reserve tenure managed by the State Government's conservation agency (DBCA 2023a).

2.4. Kimberley region

The Kimberley bioregion covers 11, 079, 500 ha in the north of Western Australia, and is characterised by tropical savanna vegetation, with a discontinuous tree canopy layer and an understory of grasslands (NVIS Technical Working Group 2017, O'Grady *et al* 2000). Vegetation heights typically range from 2 to 25 m (Werner and Peacock 2019). The north Kimberley is characterised by the tropical savanna comprised of open woodlands of *Eucalyptus tectifica* and *Eucalyptus brevifolia, Acacia* Pindan shrublands, and *Sorghum* grasses, moving towards riparian closed forests and mangroves along the coast (figures 1 and 2) (Fox *et al* 2001). The central and southern regions of the Kimberley feature *Trioda* and *Sorgham* hummock grasslands (Keith

Figure 2. Example vegetation types observed in the study regions: (a) Karri forest (eucalypt tall forest vegetation type) found in Warren IBRA region, and (b) Jarrah and Marri forest and woodland (eucalypt open forest and woodlands) found in Warren and South Jarrah IBRA regions; (c) *Banksia* low shrubland (*Acacia* shrubland vegetation type), (d) Kwongan heath (heathlands) and (e) Wandoo woodland (eucalypt woodlands) found in Geraldton Sandplains IBRA region; and (f) *Acacia* Pindan woodland (*Acacia* shrublands vegetation type) found in the Kimberley region.

2017). Only 5.28% of the terrestrial Kimberley region is legislatively conserved (DBCA 2023a); however, 25% is covered by registered carbon projects (Clean Energy Regulator 2023).

3. Methodology

3.1. Overall approach

Vegetation canopy height was spatially estimated using a RF model trained with GEDI L2A canopy height data with multi-spectral and SAR data as predictors (figure 3). The model was cross-validated using a pixel-level k-fold approach. The canopy height prediction map was then related to *in-situ* field-derived AGBD data to spatially predict AGBD across the study regions. Vegetation data from the National Vegetation Information System database (NVIS Technical Working Group 2017) was used to ensure the GEDI training data and opportunistically sampled biomass data were representative of the vegetation types in the region. A stratified sampling approach was implemented for GEDI footprints to prevent over-representation of a particular vegetation type.

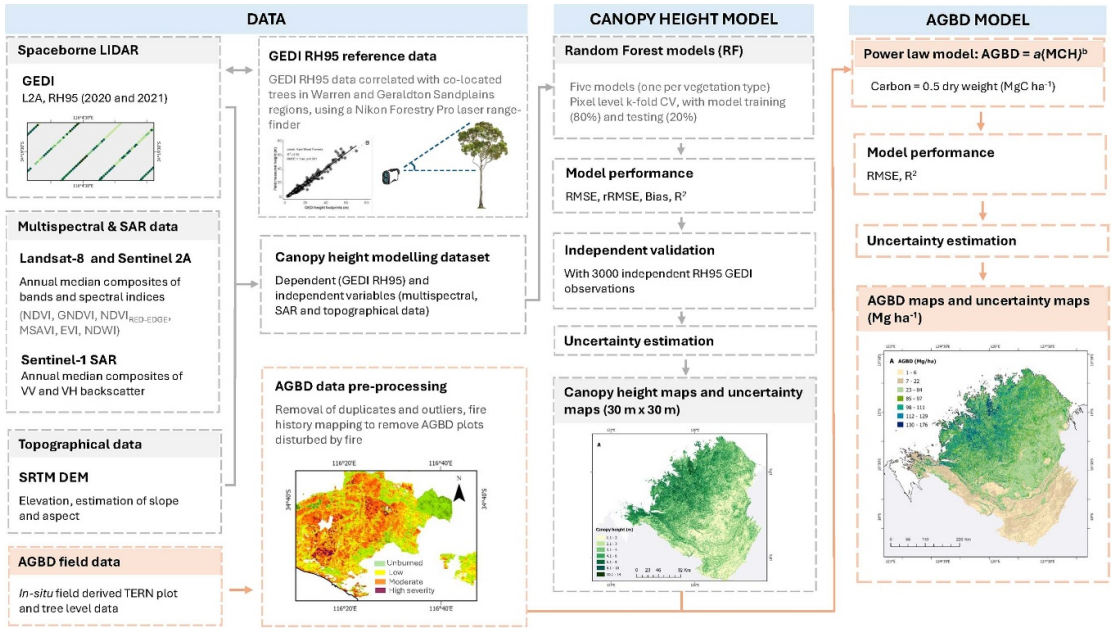

Figure 3. Overall canopy height modelling and AGBD estimation workflow.

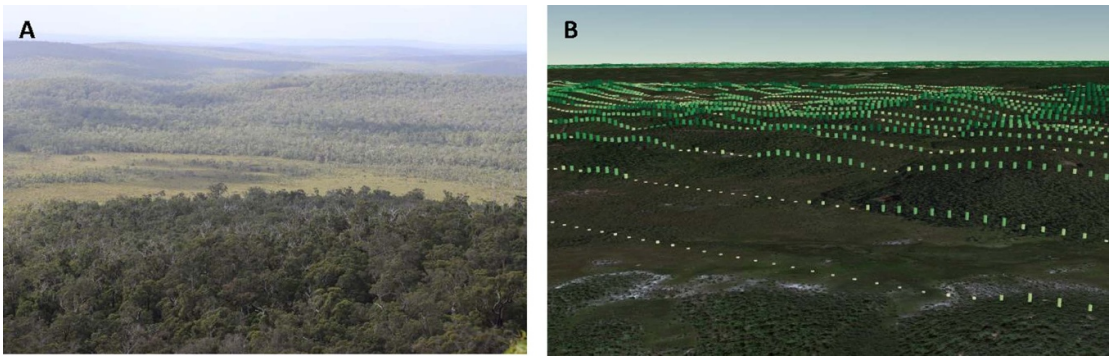

Figure 4. Quality processed transects of GEDI 2 A data across Mount Franklin National Park, Warren Region; (A) shows a photograph of the region mapped in (B).

3.2. Data

3.2.1. GEDI LiDAR data

Canopy height LiDAR data was obtained from NASA's GEDI instrument, which fires short pulses of light (14 nano-seconds long) towards the earth's surface with a laser wavelength of 1064 nm, resulting in 25 m diameter footprints separated by 60 m along the orbital track (Dubayah *et al* 2020) (figure 4). The GEDI instrument has three lasers; two full power and a third coverage laser split into two beams, resulting in four beams in total. These are optically dithered to produce eight ground tracks 600 m apart (Dubayah *et al* 2020). GEDI data was obtained as HDF5 files and converted to shapefiles clipped to the study areas using an adapted version of the *GEDI Subsetting Python script* (LP DAAC 2021). Several quality filters were implemented to select the most accurate data. Footprints collected from full-strength power beams (beams 0101, 0110, 1000, and 1011) were included as they are the most suitable for dense vegetation, and all coverage beam footprints were omitted. Data that passed threshold requirements for energy, sensitivity, amplitude, and surface tracking quality (indicated in the ancillary data by a 'quality flag' value of 1) were included and all others (with a value of 0) were omitted (Dubayah *et al* 2020). To reduce solar radiation background noise, only night-time LiDAR shots were included, indicated by footprints with a negative solar elevation angle metric (Dubayah *et al* 2020).

Canopy height measurements were obtained from the relative height (RH) data of the GEDI L2A Version 2 Elevation and Height Metrics product (Dubayah *et al* 2020). RH metrics provide a measurement of the canopy height relative to the elevation of the ground terrain by indicating the altitude at which a certain

Table 1. Vegetation spectral indices considered for prediction of canopy height and AGBD. NIR, red, blue, green and red edge refer to the near-infrared, red, blue, green and red edge bands respectively. For Sentinel−2 these refer to band 8 (NIR), band 4 (red), band 3 (green) and bands 5, 6, 7, 8A (red-edge bands (1–4)). For Landsat–8 OLI these refer to band 5 (NIR), band 4 (red) and band 2 (blue).

Radiometric Index		Formula	Platform	Reference
NDVI	Normalised differenced vegetation index	$\frac{(\mathrm{NIR}-\mathrm{Red})}{(\mathrm{NIR}+\mathrm{Red})}$	Sentinel 2	Pettorelli *et al* (2005)
GNDVI	Green normalised differenced vegetation index	$\frac{(\mathrm{NIR}-\mathrm{Green})}{(\mathrm{NIR}+\mathrm{Green})}$	Sentinel 2	Gitelson *et al* (1996)
NDVI-RE	NDVI red edge (bands 1–4)	$\frac{(\mathrm{NIR}-\mathrm{Red\ Edge})}{(\mathrm{NIR}+\mathrm{Red\ Edge})}$	Sentinel 2	Danson and Plummer (1995)
MSAVI	Modified soil adjusted vegetation index	$\frac{\left(2\mathrm{NIR}+1-\left[(2\mathrm{NIR}+1)^2-8(\mathrm{NIR}-\mathrm{Red})\right]^{\frac{1}{2}}\right)}{2}$	Sentinel 2	Qi *et al* (1994)
EVI	Enhanced vegetation index	$2.5\frac{(\mathrm{NIR}-\mathrm{Red})}{(\mathrm{NIR}+6\mathrm{Red}-7.5\mathrm{Blue})+1}$	Landsat 8	Huete *et al* (2002)
NDWI	Normalized difference water index	$\frac{(\mathrm{Green}-\mathrm{NIR})}{(\mathrm{Green}+\mathrm{NIR})}$	Sentinel 2	McFeeters (1996)

quantile of energy is returned (Dubayah *et al* 2020). The GEDI processing algorithm uses different settings to interpret the LiDAR waveform, resulting in one hundred RH metrics (RH1 to RH100). The RH metric RH95 (95th percentile of energy return) was used for this study as RH95 has been found to correlate closely with airborne ALS LiDAR data (Potapov *et al* 2021) and be the best estimate of top canopy height. The coverage of GEDI measurements across the study areas is not consistent spatially or temporally, as the GEDI instrument is constrained by the orbital tracking and inclination of its host platform, the ISS. Consequently, stratified sampling of the GEDI data was used to ensure the representation of vegetation types was proportional to the area of their distribution within the study region. Data was compiled for 2020 and 2021 to explore the difference of GEDI sampling orbital tracks between years.

3.2.2. SAR data

Sentinel−1 dual polarisation C−band SAR data with vertical-vertical (VV) and vertical horizontal (VH) backscatter polarisations from the ascending or descending orbit was used to predict canopy height and AGBD (Li *et al* 2020b, ESA 2022). Both ascending and descending data was available for the Warren and South Jarrah regions, while only descending data was available for the Geraldton and Kimberley regions. Median composites for 2020 and 2021 were compiled at 30 m pixel resolution. Sentinel−1 C− band data is provided calibrated and ortho-corrected by the Sentinel−1 Toolbox. SAR data was used to complement the optical data, particularly in regions with persistent cloud cover.

3.2.3. Multi-spectral data

Sentinel−2 Level 2A multi-spectral surface reflectance (10 m resolution) and Landsat-8 Operational Land Imager (OLI) data (30 m resolution) were used to compute vegetation spectral indices and obtain surface reflectance median image composites for 2020 and 2021 to coincide with the GEDI data.

3.2.4. Vegetation indices

Several spectral indices were explored as predictors of canopy height metrics and biomass (table 1). The normalised difference vegetation index (NDVI) is used as a proxy for canopy cover, as it indicates the absorbance of chlorophyll and is correlated to the fraction of photosynthetically active radiation (Pettorelli *et al* 2005). Red-edge NDVIs (NDVI-RE) are a derivation of NDVI that include the four red-edge bands in Sentinel-2 (the narrow wavelengths between the visible red and infra-red bands) that are more sensitive to changes in chlorophyll content and are useful for vegetation stress detection (Tillack *et al* 2014). The modified soil adjusted vegetation index (MSAVI) is a form of NDVI that includes an adjustment that

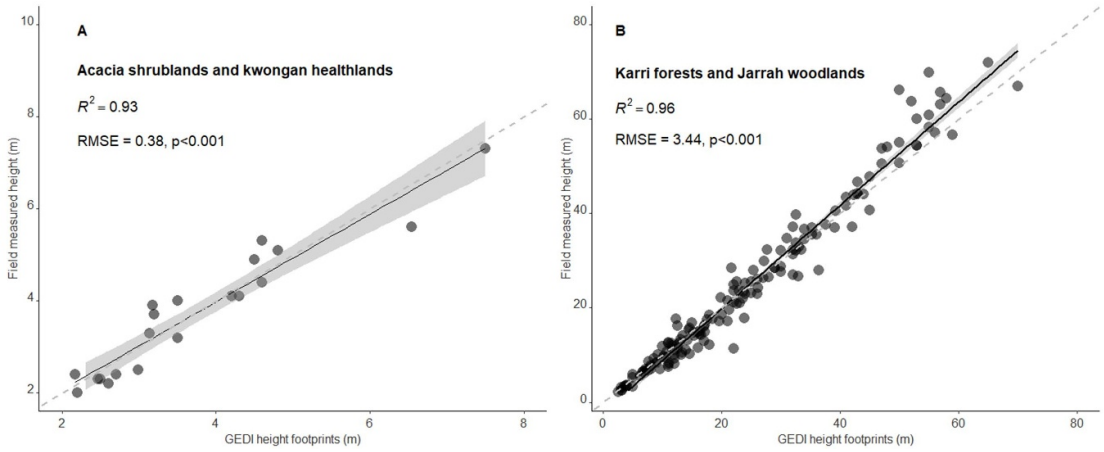

Figure 5. Comparison of GEDI RH95 and field collected canopy heights estimated using a laser rangefinder in (A) *Acacia* low shrubland and kwongan heathland vegetation in Leseur National Park in the Geraldton Sandplains region, and (B) Tall Karri eucalypt forests and Jarrah woodlands in Mount Franklin National Park in the Warren region; R^2 = coefficient of determination, RMSE = root mean square error. Dashed black line is 1:1 line.

minimises the influence of soil luminance and is beneficial in areas where vegetation cover is limited (Qi *et al* 1994). The enhanced vegetation index (EVI) is more responsive to canopy structural variations and performs better than NDVI in dense canopy where leaf area index is high (Huete *et al* 2002, Jiang *et al* 2008). The normalised difference water index was used to delineate inland water sources which other indices may confuse with vegetation (McFeeters 1996). Due to its ability to penetrate clouds, Sentinel-2 was used for most vegetation indices. The EVI composite data product from Landsat-8 was used for the EVI raster layer (USGS 2022). For the CHMs, vegetation index raster images were created at a spatial resolution of 30 m. To avoid erroneous non-zero biomass estimates in non-forested pixels, areas of naturally bare land such as sand dunes and open inland water sources were masked from analysis. Nine vegetation indices (including the four NDVI red-edge indices) and 21 bands were included, and between two and four SAR indices, depending on the availability of ascending or descending SAR back-scatter data (table 1).

3.2.5. Topographical data

Elevation, slope, and aspect were included as potential predictor variables, obtained from a digital elevation model derived from the 30 m NASA SRTM (Geoscience Australia 2020). At local and regional spatial scales, elevational gradients, slope angle, and orientation can influence vegetation structure and species distribution through microclimatic variations in temperature, precipitation and edaphic conditions (Colgan *et al* 2012, Jucker *et al* 2018, Rahman *et al* 2022, Zhang *et al* 2024). For example, trees in valleys tend to be taller due to greater competition for light resources, greater access to water and nutrients, protection from wind and deeper soils allowing increased anchorage for roots (Fortunel *et al* 2018). Slope aspect influences ecosystem characteristics due to variable exposure to solar radiation, air streams, cloud cover and precipitation (Bale *et al* 1998). These drivers of vegetation structure and biomass can be particularly relevant when modelling across heterogeneous or complex topography (Ediriweera *et al* 2014).

3.2.6. Field validation: canopy height field data

A Nikon Forestry Pro laser rangefinder was used to estimate the heights of 163 trees in Mount Franklin National Park in the Warren region and 30 trees in the Lesueur National Park in the Geraldton Sandplains region. These canopy height measurements were correlated with co-located or nearby GEDI RH95 canopy heights to determine the degree of agreement between RH95 measurements and the regions studied (figure 5). The RH95 metric GEDI canopy heights were not statistically different from the field measurements (t-test; $p > 0.05$), and as such the RH95 was deemed appropriate for use in the canopy height mapping. Greater variation was observed between the GEDI and field measurements of taller trees in denser forests.

3.2.7. AGB field data

AGBD maps created through remote sensing are modelled estimates of biomass based on spectral correlations, and such maps rely on quality field estimates for calibration and validation (Foody 2010, Olofsson *et al* 2014). AGB training data was obtained from the Terrestrial Ecosystem Research Network (TERN) Biomass Plot Library, a database of *in situ* stem inventory AGB data compiled from government, university, and research organisations (TERN 2021). Site-level AGB density per hectare data was utilised for

most sites. Where data was available at the tree level, allometric models relating tree height and diameter were used to estimate biomass, which was aggregated to create plot-level biomass estimates in megagrams (Mg) per hectare (Paul *et al* 2016, TERN 2021). The AGB data was not the result of a stratified sampling campaign, and in certain instances, had been specifically acquired from region with mature forests. Effort was therefore made to include a representative range of vegetation types in the final training data selection.

For the Warren region, site inventory data was obtained for trees and shrubs from field surveys between 2005 and 2015. This was supplemented by data obtained at the tree level at eight 100 m by 100 m plots, totalling 973 trees. These trees were used to obtain representative measurements of tree distributions in the heterogeneous vegetation for a range of species representative of the eucalypt forests in the region; Blackwood (*Acacia melanoxylon*), Willow myrtle, (*Agonis flexuosa*), Sheoak (*Allocasuarina decussata*), Marri (*C. calophylla*), Karri (*E. diversicolor*), Jarrah (*E. marginata*), Yellow Tingle (*E. guilfoylei*), Red Tingle, (*E. jacksonii*), and Karri Hazel (*Tricalcium odoratissimum*). For the remaining regions, tree-level data was unavailable, and site-level data was obtained at a 1-hectare resolution (TERN 2021). Plot-level data was available between 2005 and 2014 for the South-Jarrah region; the year 2018 for the Geraldton Sandplains region, and between 2010 and 2012 for the Kimberley region.

3.2.8. AGB data processing

The TERN AGB plot data was screened to remove duplicates, plots that were known to have experienced disturbance, or plots that were considered outliers. Two plots located in the Warren region were removed due to the presence of a high number of *E. jacksonii* individuals which as a species grow pronounced buttressing that results in inflated diameter measurements and elevated biomass estimates for a given height (TERN 2021).

Fire histories were obtained for regions where the AGB field samples were collected. Samples within the Warren, South Jarrah and Geraldton Sandplains regions subjected to fire in the time-since AGB data collection were omitted as fires in tall-open and temperate eucalypt forests and shrublands often have prolonged and significant impacts on biomass and ecosystem functioning (Enright *et al* 2014, McCaw and Middleton 2015). The timing of burns and the perimeters of fires that occurred after the AGB data collection were determined using the Fire History (DBCA-060) dataset (DBCA 2023b). Fire severity mapping was then undertaken to assess the region burnt by each fire. Landsat Thermal OLI imagery was used due to the long time series of data available. Pre-and post-fire cloud-masked imagery was obtained within three months before and following a fire, respectively, to optimise the burn signal while minimising changes due to phenology (Tran *et al* 2018). The differenced normalised burn ratio was used as a proxy for fire burn severity. Near infra-red (NIR) and short-wave infrared (SWIR) wavelengths were used to calculate a normalised burn ratio for pre- and post-fire images:

$$\mathrm{NBR} = \frac{(\mathrm{NIR} - \mathrm{SWIR})}{(\mathrm{NIR} + \mathrm{SWIR})} \tag{1}$$

Burnt regions were identified from the difference in pre- and post-burn NBR values. Thresholds of burn severity were adopted as per Key and Benson (2006). For field AGB samples from the fire-adapted Kimberley savanna region, fire histories of AGB sampling locations were checked using available fire history datasets (NAFI 2023) and samples were only omitted if a fire had occurred within three years prior to mapping, as a time interval of three or more years is considered long-unburnt for the region (Wysong *et al* 2021). Following this data cleaning process, 23 samples remained for the Warren region, 35 from the South-Jarrah region, 22 from the Geraldton Sandplains shrublands, 34 from the Kimberley eucalypt savanna, and 12 from the Kimberley grasslands region.

3.3. RF CHM

RF machine learning algorithms were used to predict canopy height from remotely sensed earth observation data and GEDI data (Breiman 2001). RF is a decision tree-based machine learning modelling approach which can be used for non-parametric datasets exhibiting complex non-linear relationships (Belgiu and Drăguţ 2016). A separate RF model was developed for each major vegetation type within the IBRA regions studied to account for distinctions in tree density and structure. Vegetation was delineated according to the Major Vegetation Groups of the National Vegetation Information System (NVIS Technical Working Group 2017). The resulting five models were (i) tall eucalypt (Karri) forests in the Warren region, (ii) the mid-open eucalypt (Jarrah) forests and woodlands of the South Jarrah region; (iii) the heathlands and shrublands of the Geraldton Sandplains region; (iv) the eucalypt savannas of the Kimberley region; and (v) the tussock and hummock grasslands of the Kimberley region. Each model was built using only the GEDI, multi-spectral, and SAR information collected from that region or vegetation type. RF models were run with 500 decision trees, with the variables used for each decision-tree split being the square of the total number of variables

(Breiman 2001). To increase model parsimony and computational efficiency, highly correlated predictor variables were omitted (Wilkes *et al* 2015). The relative importance of the predictor variables in determining the canopy height were measured by the percentage increase in mean square error (%MSE).

Model performance and errors were evaluated with a random 10 k-fold cross-validation approach, in which observations are split into *k* sets. Model training was iteratively performed on *k-1* sets, with a different set withheld for validation each time. The RF model was performed with $k = 10$ (Rodríguez-Veiga *et al* 2020). The predictive performance of the k-fold cross-validated RF models was evaluated using the root mean square error (RMSE), relative RMSE (rRMSE), prediction bias, and R^2 (Luo *et al* 2019). The RMSE is the standard deviation of the residuals. A lower RMSE value indicates data is more concentrated around the line of best fit, and less error. The RMSE was calculated according to the following:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\widehat{y}_i - y_i)^2}{n}} \tag{2}$$

The relative RMSE was used to quantify uncertainty, compare performances across models and was calculated according to the following (Li *et al* 2022b):

$$\text{rRMSE}\ (\%) = \frac{\text{RMSE}}{\bar{y}} 100 \tag{3}$$

The bias, calculated by the mean error, was used to assess the average difference between the predicted and observed mean canopy height values. A positive bias means the predicted values are greater than the observed values on average. Bias was calculated according to the following (Lang *et al* 2022):

$$\text{bias} = \frac{1}{n} \sum_{i=1}^{n} (\widehat{y}_i - y_i) \tag{4}$$

where for each of [2], [3], and [4]: $\widehat{y}_i$ is the predicted canopy height (m), y_i is the observed canopy height, $\bar{y}$ is the mean canopy height, and n is the number of samples.

This accuracy assessment was calculated for contiguous stratified ranges of the reference GEDI canopy height as well as the overall model for each vegetation type. The selected ranges varied between regions and were determined based on the range of canopy heights and the number of reference data points within each range. The model height predictions were further validated by calculating the estimation bias of the model using a separate three thousand independent GEDI reference data points, stratified into the same contiguous ranges. The RF framework was implemented in the Google Earth Engine (GEE) javascript platform, with further statistical analysis undertaken in R (v 4.3) (R Core Team 2023), using the RF (Liaw and Wiener 2002), *caret* (Kuhn 2008) and *ggplot* (Wickham 2016) packages.

3.4. AGBD-height power model

The canopy height prediction map was used as a predictor for AGBD. This was achieved using power-law models derived by relating AGBD field data to co-located 30 m × 30 m canopy height estimates, and creating regressions of the general formula (Meyer *et al* 2018, Duncanson *et al* 2022):

$$\text{AGBD} = a(\text{MCH})^b \tag{5}$$

where AGBD is the abovegound biomass density (Mg ha^{-1}), MCH is the mean canopy height for the pixel, a is a scaling factor, and b is the power-law coefficient (Meyer *et al* 2018). Carbon stocks were then quantified from the AGBD estimates based on the assumption that carbon is 50% of dry weight, resulting in Megagrams of carbon per hectare (MgC ha^{-1}) (Burrows *et al* 2002, Volkova *et al* 2015).

3.5. Uncertainty estimation

Four sources of error were considered in the canopy height modelling: the error from the RF model; the vertical accuracy of the GEDI data (assumed to be 0.5 m) (Dubayah *et al* 2020); the within-pixel sampling error, which arises due to the difference in shape and size of the area of the GEDI footprint (25 m) and the pixel size chosen for the mapping (30 m); and the temporal error between the collection of the GEDI data and the predictor remote sensing imagery. As per Zhang *et al* (2014), the error associated with the difference in the spatial variability of GEDI footprints and remote sensing plots was considered negligible because the area of the 30 m pixel is approximately the same as the 25 m footprint, taking into account the 5 m geolocation error of the GEDI data (Zhang *et al* 2014). The temporal error was also assumed to be zero given the GEDI

and earth observation data (Landsat, Sentinel) were sampled over the same period. The overall uncertainty for each map pixel in the canopy height modelling was derived additively from these four sources of error:

$$\varepsilon_{\mathrm{MCH}} = \left(\varepsilon^2{}_{\mathrm{RF\ model}} + \varepsilon^2{}_{\mathrm{GEDI}} + \varepsilon^2{}_{\mathrm{within-pixel}} + \varepsilon^2{}_{\mathrm{time}}\right)^{\frac{1}{2}} \quad (6)$$

For the AGBD mapping, the uncertainty at the pixel level was assumed to be composed of six independent sources of error. Firstly, the error associated with the canopy height modelling earlier explained. Secondly, the uncertainty associated with temporal differences between the date of AGBD field plot collection and collection of the mean canopy heights used to predict AGBD, as trees may have grown or degraded in that time. The temporal error was estimated as per Weisbin *et al* (2014) and adopted to be relevant to the dates of the data used in each model. Thirdly, a measurement error from upscaling tree-level parameters to a plot scale (Chave *et al* 2004). Measurement errors originate due to human error in the measurement of tree diameter and height, which aggregate cumulatively at the plot level (Araza *et al* 2022). The measurement error was assumed to be 10% according to Mitchard *et al* (2011). Fourthly, an allometry error associated with the choice of empirical model used to estimate AGB. The error associated with estimating AGB using allometric equations was assumed to be 11%, as per Chave *et al* (2004). Fifthly, the error resulting from the empirical model used to convert canopy height into AGB. Finally, a within-pixel sampling error as a result of local variability in AGBD within pixels. The within-pixel sampling error is a function of field plot size (ha) and remote sensing pixel size and occurs where AGBD is locally heterogeneous, and where field plots are smaller or larger than the size of the map pixels. The within-pixel sampling error is larger as the sampling area decreases. The within-pixel sampling error for the AGBD model was estimated by adapting the within-pixel sampling error matrix by Réjou-Méchain *et al* (2014). Using the 30 × 30 m canopy map estimate and a 1 ha field AGBD plot area, the within-pixel sampling error for the AGBD model was assumed to be 4% (Réjou-Méchain *et al* 2014). The overall uncertainty for each map pixel, εAGBD, was then derived additively from these six sources of error, as per Rodríguez-Veiga *et al* (2016), as the following:

$$\varepsilon_{\mathrm{AGBD}} = \left(\varepsilon^2{}_{\mathrm{MCH}} + \varepsilon^2{}_{\mathrm{time}} + \varepsilon^2{}_{\mathrm{measurement}} + \varepsilon^2{}_{\mathrm{allometry}} + \varepsilon^2{}_{\mathrm{MCH\ to\ AGBD}} + \varepsilon^2{}_{\mathrm{within-pixel\ sampling}}\right)^{\frac{1}{2}} \quad (7)$$

4. Results

4.1. Canopy height modelling

4.1.1. Model accuracy

For the year 2021, the multispectral and SAR predictor variables explained more than 70% of the variation in the canopy height of the tall eucalypt and mid-open eucalypt forests, but as low as 39% of the variation in heights in eucalypt savanna vegetation (table 2). The accuracy of the canopy height prediction models between 2020 and 2021 were comparable (see supplementary table 1 for 2020 results). Although the overall bias approached zero for all regions, the models consistently overestimated the heights of low-lying vegetation and underpredicted the upper bounds of vegetation height. Negative bias was observed overall for the tall eucalypt forests in the Warren region and for the mid-open eucalypt forests and woodlands in the South Jarrah region in 2021, indicating these vegetation types were underpredicted on average. A positive bias in low-lying heathland, savanna and grassland vegetation indicates these vegetation types were overpredicted by the model on average.

Elevation and the SAR backscatter variables ($\mathrm{VV_{desc,}}$ $\mathrm{VV_{asc,}}$ $\mathrm{VH_{desc,}}$ or $\mathrm{VV_{asc}}$) were the most important features in predicting canopy height for all vegetation types except the heathlands and shrublands in the Geraldton Sandplains region. The GNDVI and the MSAVI, were important in the sparse heathlands and shrublands of the Geraldton Sandplains region, but less so in other regions and vegetation types (table 2).

4.1.2. CHM validation

The ability of the RF models to predict canopy height using GEDI RH95 data was tested in each of the five vegetation types using three thousand model output estimations against a co-located and independent subset of the GEDI data (figure 6). Again, all models showed non-uniform bias, with an underestimation (negative bias) of GEDI values for higher canopies and an overestimation (positive bias) of lower canopy heights. For all regions, a linear regression between the predicted height values and GEDI RH95 data resulted in a slope of between 1.01 and 1.1, with intercepts of between −1.5 for the Warren region (for the year 2021) and −0.12 for the eucalypt savanna in the Kimberley (for the year 2021), reflecting an approximate 1:1 relationship (figure S1).

4.1.3. Canopy height mapping

Predicted mean canopy heights estimated at the 30 m pixel level reached a maximum of 42 m in the south-western Warren region, 32 m in the South-Jarrah region, 11 m in the Geraldton Sandplains region and

Table 2. Performance of RF models predicting canopy height for 2021, stratified by reference canopy height ranges as well as the overall model (indicated in bold). The variable importance of predictor variables, and sample size for each model is also listed. Variables with * are derived from Landsat. RMSE is the Root Mean Square Error; R^2 is the coefficient of determination.

Study region	Reference CH range (m)	*n* training samples	RMSE (m)	Relative RMSE (%)	Bias (m)	R^2	Variable importance of overall model (top 10)
Warren	0–10	5116	5.86	70.60	3.40	**0.70**	Elevation, VV_{desc}, VV_{asc},
Tall eucalypt	10–20	3457	6.51	43.40	2.13		$NDVI_{RE2}$, NIR*, VH_{desc},
forests	20–30	3195	6.23	24.92	−0.56		SWIR1*, EVI, GNDVI,
	30–40	1548	8.88	25.37	−6.34		$NDVI_{RE3}$
	>40	1100	15.3	34.00	−13.85		
	Overall	**14 416**	**7.43**	**29.72**	**−0.02**		
South Jarrah	0–5	4927	3.16	89.12	2.37	**0.74**	Elevation, VV_{desc},
Mid- open	5–10	1087	5.09	64.25	2.81		$NDVI_{RE2}$, VV_{asc}, VH_{desc},
eucalypt forests	10–15	2707	3.27	30.62	0.61		$NDVI_{RE3}$, SWIR1*, VH_{asc},
and woodlands	15–20	3632	3.30	20.60	−1.36		$NDVI_{RE4}$, NIR*.
	>20	1635	9.49	31.75	−7.39		
	Overall	**13 988**	**3.81**	**37.96**	**− 0.18**		
Geraldton	0–2.0	5182	0.38	19.83	0.27	**0.41**	Elevation, GNDVI, EVI,
Sandplains	2.0–2.5	6136	0.41	19.34	0.20		VH_{desc}, MSAVI, $NDVI_{RE1}$,
Heathlands,	2.5–3	792	0.48	17.90	−0.11		$NDVI_{RE2}$, SWIR1*, VV_{desc},
shrublands	3–7	987	1.83	43.67	−1.36		NIR*.
	>7	204	4.35	54.88	−4.53		
	Overall	**13 301**	**0.99**	**40.31**	**0.013**		
Kimberley	0–2.5	2822	1.96	99.73	1.78	**0.39**	VH_{desc}, SWIR1*, Elevation,
Eucalypt savannas	2.5–5	4084	1.86	53.10	1.18		EVI, VV_{desc}, $NDVI_{RE3}$,
	5–10	4073	2.19	30.60	−1.15		$NDVI_{RE2}$, NIR*, Blue,
	>10	1044	6.75	48.36	−4.40		$NDVI_{RE4}$.
	Overall	**12 023**	**2.68**	**48.36**	**0.05**		
Tussock &	0–2.5	3107	1.16	54.08	0.84	**0.43**	VH_{desc}, VV_{desc}, $NDVI_{RE1}$,
hummock	2.5 −5	2861	1.37	38.58	0.81		SWIR2*, Blue, NIR*,
grasslands	5–8	1327	1.91	30.64	−1.56		$NDVI_{RE2}$, Elevation, EVI,
	>8	691	5.01	52.34	−3.21		$NDVI_{RE3}$.
	Overall	**7986**	**2.09**	**52.17**	**0.003**		

12 m in the Kimberley region (figure 7). The canopy height predictions between 2020 and 2021 were not significantly different (ANOVA; $p > 0.05$) (figure 7, see supplementary figure S2 for 2020 canopy height map). The Warren region exhibited the greatest spatial heterogeneity and variation in the canopy height estimations.

4.2. Above-ground biomass and carbon density mapping

4.2.1. Model performance

Overall, the canopy height and AGBD power-law models exhibited relatively good performance, with R^2 higher than 0.80, and RMSE less than 72.4 (table 3). The strongest correlation between the heights and AGBD was observed in the Geraldton Sandplains region, with the lowest observed in the Warren region.

4.2.2. AGBD characterisation across study regions

The spatial distribution of AGBD varied between vegetation types and regions (figure 8, see supplementary figure S3 for 2020 results). Regions with higher estimated AGBD were the tall eucalypt forests in the south-west of the Warren region, and the northern region of the Kimberley. The mean estimated values of AGBD in the Warren region were consistent between years, with 139.0 ± 16.7 Mg ha^{-1} to 138.0 ± 24.0 Mg ha^{-1} in 2020 and 2021 respectively. While the maximum AGBD values for the South Jarrah region remained relatively consistent between years (324 Mg ha^{-1}–336 Mg ha^{-1}), the mean AGBD reduced from 82.0 ± 10.0 Mg ha^{-1} in 2020 to 67.5 ± 10.0 Mg ha^{-1} 2021. In the Geraldton Sandplains region, the average AGBD decreased from 18.8 ± 2.9 Mg ha^{-1} to 14.2 ± 2.7 Mg ha^{-1} in 2021. In the northern eucalypt savanna of the Kimberley, the AGBD estimates were 83.0 ± 3.5 Mg ha^{-1} in 2020 and 87.8 ± 9.9 Mg ha^{-1} in 2021. In the Kimberley grasslands, the AGBD mean increased slightly from 7.5 ± 1.8 Mg ha^{-1} in 2020 to 7.7 ± 2.0 Mg ha^{-1} in 2021. For the year 2021, the mean carbon stock was estimated at 69.0 ± 12.0 MgC ha^{-1} in the Warren region; 33.8 ± 5.0 MgC ha^{-1} for South Jarrah; 7.1 ± 1.4 MgC ha^{-1} for Geraldton Sandplains

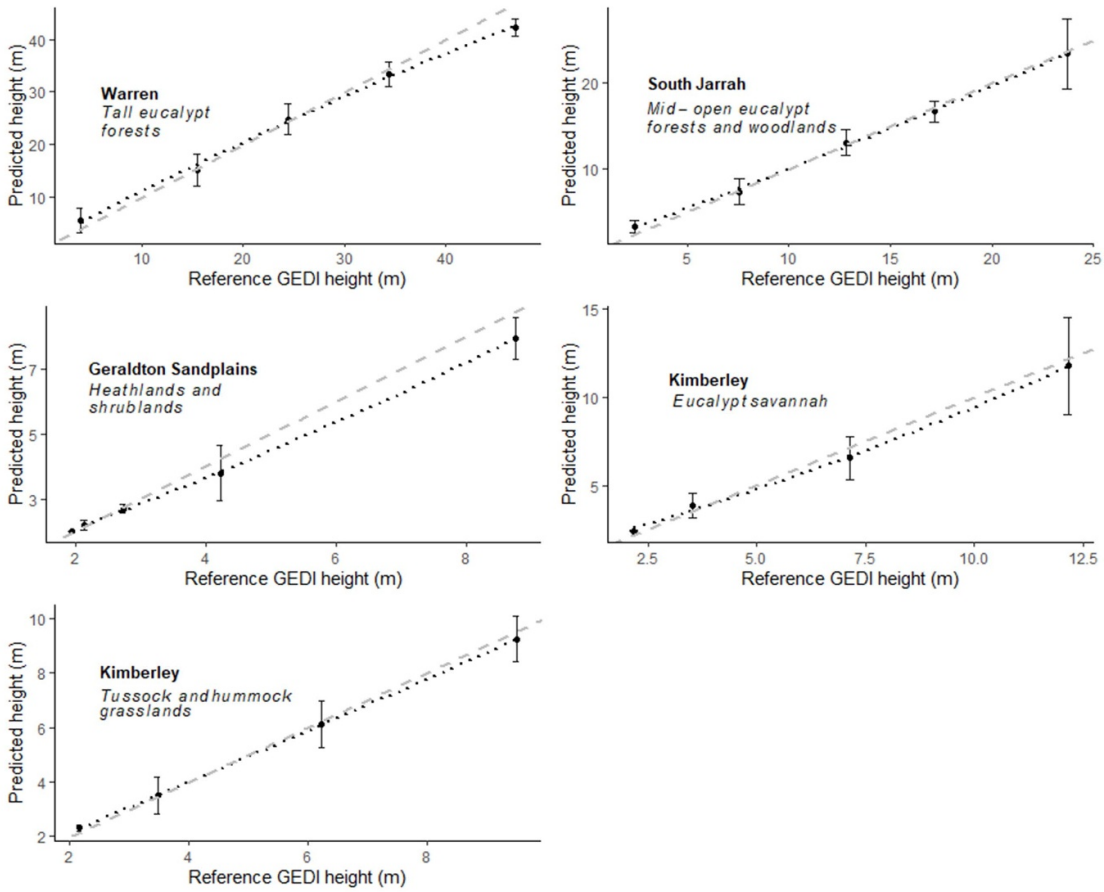

Figure 6. Average predicted canopy heights versus average reference GEDI RH95 heights at the pixel scale for each reference range, for the year 2021. Error bars indicate the standard deviation of the predicted canopy heights for each reference range. The dotted black line indicates a line-of-best fit; and the dashed grey line indicates a 1:1 line. Bias forms the main source of error for canopy height ranges where the error bars do not overlap the 1:1 line.

region; 43.9 ± 4.9 MgC ha^{-1} for the Kimberley eucalypt savanna; and 3.9 ± 1.0 MgC ha^{-1} for the Kimberley savanna grasslands.

5. Discussion

The availability of spaceborne GEDI LiDAR specifically designed to measure vertical vegetation structure provides new powerful opportunities for improved understanding of habitat heterogeneity and accuracy in AGBD mapping. This study demonstrated the potential for utilising a fusion of GEDI LiDAR, active SAR, passive optical multispectral, and topographical data, to create spatially continuous AGBD maps, at a resolution and accuracy appropriate for climate mitigation and biodiversity policy application. While this approach was implemented to evaluate differences over five vegetation types over three biomes of unique flora speciation within Western Australia, these vegetation types are structurally similar to vegetation both in other regions of Australia, and in Mediterranean and tropical regions globally, broadening the applicability of these methods and conclusions.

Published AGBD field estimates for Western Australian vegetation types are scarce at the plot level (Adams *et al* 2001, Raison *et al* 2003, TERN 2021) and negligible at the landscape or regional level, confirming the need for this study. The baseline plot estimates available in literature indicate that the AGBD predictions obtained here are credible for the vegetation types mapped. The AGBD estimates for the northern Kimberley tropical savanna were within the range obtained by Collins *et al* (2009), who reported AGB densities of between 10 and 174 Mg ha^{-1} in the tropical savanna ecosystems of the Northern Territory using SAR data and field measurements. The mean estimates obtained in this study for the Kimberley grasslands, Geraldton Sandplains heathlands and shrublands and South Jarrah mid-open forest and woodlands are in broad agreement with published field sampled estimates of those vegetation types: Adams *et al* (2001) estimated 4.5 Mg ha^{-1} and 6.7 Mg ha^{-1} for hummock *Triodia* grasslands; Grierson *et al* (2000) estimated 82 Mg ha^{-1} in Jarrah forests; and Low and Lamont (1990) estimated 13.9 Mg ha^{-1} in *Banksia*

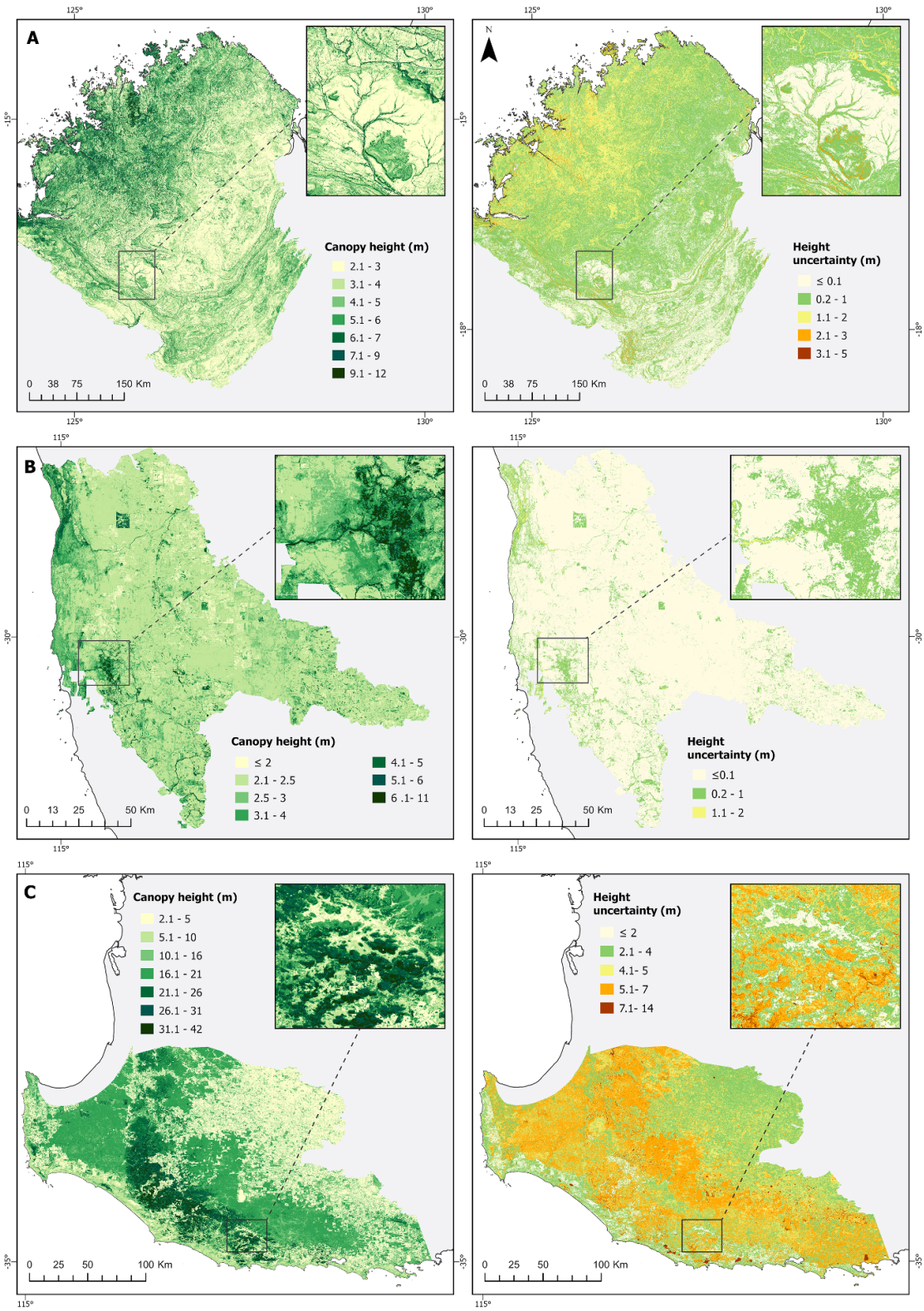

Figure 7. Canopy height (m) (left) and height uncertainty maps (right) for the (A) Kimberley, (B) Geraldton Sandplains, (C) Warren and South Jarrah regions for the year 2021. The inset map shows close-up views of canopy height gradients and vegetation structure heterogeneity.

heathlands and shrublands in the Geraldton region. This relative alignment between the AGBD estimates obtained through this study and field-based sampling, suggests that the methods used result in AGBD predictions that are reasonably accurate, enabling their use in decision making requiring a baseline assessment of AGBD and carbon stocks. The predicted AGBD estimates for the Warren region overlapped the lower bounds of range reported by Wood *et al* (2015), who reported plot–level estimates between

Table 3. Performance of AGBD models. RMSE = root mean square error; R^2 = coefficient of determination.

Study region	*n* training samples	RMSE	R^2
Warren	23	72.4	0.82
Tall eucalypt forests			
South Jarrah	35	67.6	0.87
Mid- open eucalypt forests and woodlands			
Geraldton Sandplains	22	21.1	0.94
Heathlands, shrublands			
Kimberley	34	34.7	0.84
Eucalypt savannas			
Tussock & hummock grasslands	12	5.7	0.87

310 Mg ha^{-1} and 540 Mg ha^{-1} obtained from non-destructive field sampling. These estimates obtained by Wood *et al* (2015) intentionally sampled tall, dense, old-growth forests, and as such their observations represent the upper tail of the distribution of biomass estimates and they are not wholly representative of the range of vegetation in the broader Warren region. The reduced maximum AGBD observed in the Warren region prediction maps, relative to these field sampled estimates may have also propagated from the non-uniform bias observed in the canopy height predictions (Réjou-Méchain *et al* 2019).

While the combination of GEDI, passive multispectral and active SAR data, and the use of RF models produced average canopy height estimates plausible for the vegetation types mapped in this study (Specht 1970, Rayner 1991, Vigilante and Bowman 2004) the method underestimated canopy heights in tall dense forests and overestimated the height of shorter sparse vegetation. This trend reflects a known bias observed in the use of RF models and data fusion methods to predict canopy height (Staben *et al* 2018, Li *et al* 2020b, Fassnacht *et al* 2021). This underestimation of taller heights in the RF canopy height prediction models is typically attributed to signal saturation of passive multi-spectral sensors and insufficient canopy penetration by SAR sensors (Duncanson *et al* 2010). The overestimation of the lower vegetation height values is likely associated with the limited ability of passive sensors to distinguish between the spectral reflectance of canopy and understory vegetation or soil, or the distorting influence of complex terrain on SAR backscatter (Li *et al* 2020b, Wang *et al* 2021). The reduced canopy height bias in the middle height quantiles is consistent with the statistical behaviour of regression machine learning models, which encourage the training data mean and prediction mean to be approximately equal (Araza *et al* 2022), and may explain how the mean AGBD estimates were approximately accurate relative to field obtained estimates.

For most regions, only slight variations in the height and AGBD estimates were observed between 2020 and 2021, reinforcing the robustness of the modelling process. The lower AGBD estimates in the South Jarrah region in 2021 relative to 2020, may reflect a loss of biomass from environmental disturbances such as fire; the region experienced an estimated 59 000 ha of woodland burns during 2020 (DBCA 2023b). It may also reflect variations in the raw GEDI height data distribution. The orbit of the ISS undertook a more complete coverage of the region in 2020 and reported a slightly higher mean height (11 m rather than 9 m in 2021). RF models predict within the bounds of the training data provided (Millard and Richardson 2015), and if the GEDI sensor onboard the ISS fails to sample the full distribution of tree heights in an area of high vegetation structural heterogeneity, the canopy map and resulting AGBD map are likely to underestimate or overestimate those values. While stratified sampling of the GEDI data was undertaken to avoid excess sampling of a given vegetation type in a given year, certain vegetation types, such as Jarrah (*E. marginata*) 'mid open forests', may still exhibit considerable height variation. This effect would likely be negligible in regions of vegetation structural homogeneity. The GEDI trained CHM is applicable beyond the temporal bounds explored in this study, given the same type of data is used (same processing level and version) from Sentinel−1, Sentinel−2 and Landsat. The power-law allometric relationship between canopy height and AGBD varies with vegetation type and climate (Chave *et al* 2014, Jucker *et al* 2017). Over a long period of time, the growth behaviour of vegetation may change due to changing climatic conditions, the vegetation may degrade from disturbances such as wildfire, or regrow through conservation efforts. In these cases, the allometric relationship between canopy height and AGBD for a given region may change and might benefit from recalibration with updated field data.

The proportion of variance in GEDI canopy heights explained by the predictor variables of the canopy height RF models were within the range achieved at similar spatial resolutions from airborne LiDAR in regional studies in temperate and tropical biomes (García *et al* 2018, Torres de Almeida *et al* 2022), and other

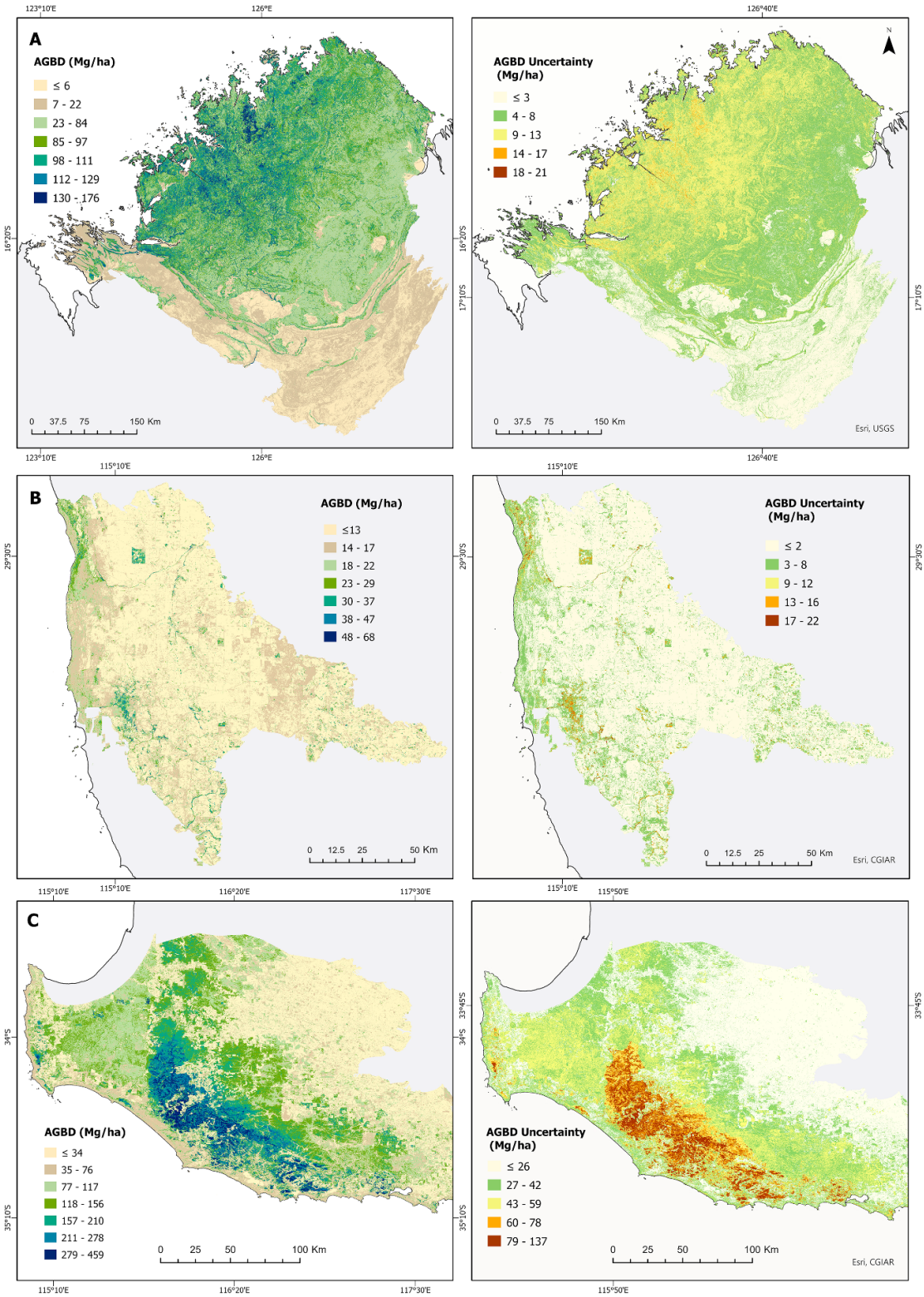

Figure 8. AGBD maps (left) and AGBD uncertainty maps (right) for the (A) Kimberley, (B) Geraldton, (C) Warren and South Jarrah regions for the year 2021.

space-borne LiDAR sensors such as full-waveform ICESat (Lefsky *et al* 2005, Iqbal *et al* 2013). The proportion of variation in AGBD explained by the GEDI CHM was similar or higher than in regional-level airborne LiDAR studies in temperate and tropical biomes utilizing power-law or OLS modelling approaches (Zhao *et al* 2009, Ferraz *et al* 2016, 2018, Naidoo *et al* 2016, Xu *et al* 2017, Labriere *et al* 2018). These results indicate the potential for this method to inform climate and biodiversity policy as accurate broadscale AGBD and carbon stock mapping is possible on a landscape scale, without the costs associated with airborne LiDAR.

The RF GEDI CHMs exhibited greater accuracy when applied to woody vegetation types than for heathland with low lying vegetation and sparse canopy. This is in agreement with Leite *et al* (2022) who found GEDI waveform metrics explained 88% of variation in woody fuels, and only 46% of herbaceous shrub fuels in a Cerrado savanna ecosystem in Brazil. This suggests that the utilisation of GEDI height metrics in this method is most appropriate for woody biomass regions, which does correlate to the vegetation types for which the GEDI instrument sensor has been designed (Dubayah *et al* 2020). However, while the savanna and grassland ecosystems in the Kimberley and sparse shrubland ecosystems in the Geraldton Sandplains exhibited the lowest proportion of explained variation in the RF models (indicated by lower coefficients of determination), the use of these canopy height prediction maps to generate AGBD estimations which are in relative agreement with published field-derived estimates suggests that this method remains useful for policy applications.

Uncertainty may also be introduced into the CHM through factors which impact GEDI's ability to detect the ground elevation. Relative to the Warren tall eucalypt forest model, the South Jarrah open forest model exhibited higher accuracy, potentially reflecting the impact of canopy cover on the strength of the GEDI return signal (Hofton and Blair 2019). Topographic relief and slope can be a source of uncertainty in CHMs derived from LiDAR as complex terrain may cause GEDI full-waveforms to broaden and ground and vegetation signals to overlap, degrading the accuracy of the ground and height estimation (Hyde *et al* 2005). Although GEDI's small 25 m footprint is designed to reduce mixing of ground and vegetation signals (Dubayah *et al* 2020), several studies have shown rugged terrain and steeper slopes may result in GEDI canopy height overestimation or underestimation due to inaccurate ground detection (Huettermann *et al* 2022, Wang *et al* 2022). Conversely, Olivera *et al* (2023) explored the effect in tropical Brazilian forests and found GEDI terrain slope only significantly impacted the quality of daytime GEDI shots. The data pre-processing undertaken in this study aimed to reduce the potential for slope and elevation impacts by selecting only the night-time and highest quality shots; however, there may be some error introduced from these factors that is transferred downstream to the biomass estimates that rely on the canopy height prediction.

The accuracy of AGBD estimations is dependent on the quality and quantity of the biomass field calibration data; although field AGBD estimates are commonly assumed as accurate ground-reference data, they can be a source of error (Réjou-Méchain *et al* 2019). While effort was made to include a representative range of vegetation types in the data selection, the AGBD data used in this study was not the result of a designed or stratified sampling campaign, and as such certain vegetation types were less represented (TERN 2021). The uncertainty of the AGBD modelling presented here could be improved by the inclusion of a larger training dataset through additional field AGBD plots, particularly in regions which exhibit high spatial heterogeneity in the vegetation structure, such as the Warren region. The AGBD estimation would also benefit from more recent field biomass estimates. While the temporal lapse was accounted for in the estimation of pixel-level error propagation, and fire area mapping, any remote sensing AGBD model validation process benefits from field data that is coincident spatially and temporally (Réjou-Méchain *et al* 2019).

Exploration of deep learning methods such as convolutional neural networks are emergent in AGBD estimation literature (Lang *et al* 2022, 2023) and present an alternative method which may help reduce the estimation bias observed with RF models. The GEDI instrument aboard the ISS begun data collection again in 2024 after a brief period of hibernation in 2023. A longer time series of GEDI data would support quantification of AGBD changes over timescales useful for applications such as measuring forest regeneration over five or ten years. While the inconsistent orbital tracking of the ISS and subsequent variations in coverage of GEDI between time periods presents some challenges, further research, over different spatial and time scales in regions with known disturbance histories would help to better quantify the significance of this effect on inter-temporal biomass estimations.

6. Conclusion

This is the first known attempt using this specific combination of GEDI, multispectral and SAR sensors, and modelling approaches to estimate AGBD at a regional scale for Western Australia. The accuracy of this method was higher or similar to previous spaceborne sensors such as ICESat, and regional airborne LiDAR methods, indicating that this is an exciting opportunity to undertake accurate, low cost, broadscale AGBD and carbon stock mapping. Although uncertainties remain, there is promising capability for this method to guide future biodiversity and climate mitigation policy through assessing fuel load dynamics and predicting wildfire occurence, validating baseline carbon estimates, and improving understanding of key ecosystem functioning attributes such as habitat structure and heterogeneity.

Data availability statement

The data cannot be made publicly available upon publication because the cost of preparing, depositing and hosting the data would be prohibitive within the terms of this research project. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgment

The biomass data for this study was sourced from the Terrestrial Ecosystem Research Network (TERN) Biomass Plot Library, which is enabled by the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS). This study would not have been possible without this freely available tree and plot inventory data, and we thank all the researchers who braved the ticks and bushflies in remote locations and generously shared their data through TERN for the purpose of validating satellite mapping across Australia. We also thank the reviewers for their helpful suggestions.

The authors acknowledge the Traditional Owners and Custodians of Australia, particularly those of the regions mapped in this study. We recognise their knowledge of Country, and continued stewardship of the land, water, biodiversity and culture. We pay our respects to Elders past and present.

Conflict of interest

The authors have no competing interests to declare.

ORCID iD

Natasha Lutz https://orcid.org/0000-0001-9765-1181

Reference

Adams M A *et al* 2001 *Biomass and Carbon in Vegetation of the Pilbara Region, Western Australia* (Ecosystems Research Group)

Araza A *et al* 2022 A comprehensive framework for assessing the accuracy and uncertainty of global above-ground biomass maps *Remote Sens. Environ.* **272** 112917

Axelsson K *et al* 2024 *Oxford Principles for Net Zero Aligned Carbon Offsetting (Revised 2024)* (Oxford University) (available at: www.smithschool.ox.ac.uk/sites/default/files/2024-02/Oxford-Principles-for-Net-Zero-Aligned-Carbon-Offsetting-revised-2024.pdf)

Baccini A, Laporte N, Goetz S J, Sun M and Dong H 2008 A first map of tropical Africa's above-ground biomass derived from satellite imagery *Environ. Res. Lett.* **3** 045011

Balch J K, Abatzoglou J T, Joseph M B, Koontz M J, Mahood A L, McGlinchy J, Cattau M E and Williams A P 2022 Warming weakens the night-time barrier to global fire *Nature* **602** 442–8

Bale C L, Williams J B and Charley J L 1998 The impact of aspect on forest structure and floristics in some Eastern Australian sites *For. Ecol. Manage.* **110** 363–77

Belgiu M and Drăguţ L 2016 Random forest in remote sensing: a review of applications and future directions *ISPRS J. Photogramm. Remote Sens.* **114** 24–31

Boer M, Sadler R J, Wittkuhn R S, McCaw L and Grierson P F 2009 Long-term impacts of prescribed burning on regional extent and incidence of wildfires—Evidence from 50 years of active fire management in SW Australian forests *For. Ecol. Manage.* **259** 132–42

BoM 2024 *Australia's Official Weather Forecasts & Weather Radar—Bureau of Meteorology* (available at: www.bom.gov.au/) (Accessed 21 February 2024)

Brack C L and Richards G P 2002 Carbon accounting model for forests in Australia *Environ. Pollut.* **116** S187–94

Brede B *et al* 2022 Non-destructive estimation of individual tree biomass: allometric models, terrestrial and UAV laser scanning *Remote Sens. Environ.* **280** 113180

Breiman L 2001 Random Forests *Mach. Learn.* **45** 5–32

Brigot G, Simard M, Colin-Koeniguer E and Boulch A 2019 Retrieval of forest vertical structure from PolInSAR data by machine learning using LIDAR-derived features *Remote Sens.* **11** 381

Bullock E L *et al* 2023 Estimating aboveground biomass density using hybrid statistical inference with GEDI lidar data and Paraguay's national forest inventory *Environ. Res. Lett.* **18** 085001

Burrows W H, Henry B K, Back P V, Hoffmann M B, Tait L J, Anderson E R, Menke N, Danaher T, Carter J O and. Mckeon G M 2002 Growth and carbon stock change in eucalypt woodlands in northeast Australia: ecological and greenhouse sink implications *Glob. Change Biol.* **8** 769–84

Calders K *et al* 2015 Nondestructive estimates of above-ground biomass using terrestrial laser scanning *Methods Ecol. Evol.* **6** 198–208

Catchpole W R and Wheeler C J 1992 Estimating plant biomass: a review of techniques *Aust. J. Ecol.* **17** 121–31

Chave J *et al* 2014 Improved allometric models to estimate the aboveground biomass of tropical trees *Glob. Change Biol.* **20** 3177–90

Chave J, Condit R, Aguilar S, Hernandez A, Lao S and Perez R 2004 Error propagation and scaling for tropical forest biomass estimates *Phil. Trans. R. Soc.* B **359** 409–20

Chuvieco E, Pettinari M L, Koutsias N, Forkel M, Hantson S and Turco M 2021 Human and climate drivers of global biomass burning variability *Sci. Total Environ.* **779** 146361

Clark M L, Roberts D A, Ewel J J and Clark D B 2011 Estimation of tropical rain forest aboveground biomass with small-footprint lidar and hyperspectral sensors *Remote Sens. Environ.* **115** 2931–42

Clean Energy Regulator 2023 *Emissions Reduction Fund Project Register* (available at: www.cleanenergyregulator.gov.au/ERF/project-and-contracts-registers/project-register) (Accessed 21 February 2024)
Colgan M S, Asner G P, Levick S R, Martin R E and Chadwick O A 2012 Topo-edaphic controls over woody plant biomass in South African savannas *Biogeosciences* **9** 1809–21
Collins J, Hutley L B, Williams R J, Boggs G, Bell D and Bartolo R 2009 Full article: estimating landscape-scale vegetation carbon stocks using airborne multi-frequency polarimetric synthetic aperture radar (SAR) in the savannahs of north Australia *Int. J. Remote Sens.* **30** 1141–59
Danson F M and Plummer S E 1995 Red-edge response to forest leaf area index *Int. J. Remote Sens.* **16** 183–8
DBCA 2023a *DBCA—Legislated Lands and Waters (DBCA-011)—Datasets—data.wa.gov.au* (available at: https://catalogue.data.wa.gov.au/dataset/dbca-legislated-lands-and-waters) (Accessed 21 February 2024)
DBCA 2023b DBCA fire history (DBCA-060) (available at: https://catalogue.data.wa.gov.au/dataset/dbca-fire-history) (Accessed 21 February 2024)
De Frenne P, Zellweger F, Rodríguez-Sánchez F, Scheffers B R, Hylander K, Luoto M, Vellend M, Verheyen K and Lenoir J 2019 Global buffering of temperatures under forest canopies *Nat. Ecol. Evol.* **3** 744–9
Demol M *et al* 2022 Estimating forest above-ground biomass with terrestrial laser scanning: current status and future directions *Methods Ecol. Evol.* **13** 1628–39
Dhargay S, Lyell C S, Brown T P, Inbar A, Sheridan G J and Lane P N J 2022 Performance of GEDI space-borne LiDAR for quantifying structural variation in the temperate forests of South-Eastern Australia *Remote Sens.* **14** 3615
Díaz S and Malhi Y 2022 Biodiversity: concepts, patterns, trends, and perspectives *Annu. Rev. Environ. Resour.* **47** 31–63
Dubayah R *et al* 2020 The global ecosystem dynamics investigation: high-resolution laser ranging of the Earth's forests and topography *Sci. Remote Sens.* **1** 100002
Duncanson L I, Niemann K O and Wulder M A 2010 Estimating forest canopy height and terrain relief from GLAS waveform metrics *Remote Sens. Environ.* **114** 138–54
Duncanson L *et al* 2022 Aboveground biomass density models for NASA's global ecosystem dynamics investigation (GEDI) lidar mission *Remote Sens. Environ.* **270** 112845
ED Chaves M, CA Picoli M and D Sanches I 2020 Recent applications of landsat 8/OLI and sentinel-2/MSI for land use and land cover mapping: a systematic review *J. Remote Sens.* **12** 3062
Ediriweera S, Pathirana S, Danaher T and Nichols D 2014 Estimating above-ground biomass by fusion of LiDAR and multispectral data in subtropical woody plant communities in topographically complex terrain in North-eastern Australia *J. For. Res.* **25** 761–71
Enright N J, Fontaine J B, Lamont B B, Miller B P and Westcott V C 2014 Resistance and resilience to changing climate and fire regime depend on plant functional traits *J. Ecol.* **102** 1572–81
Erdody T L and Moskal L M 2010 Fusion of LiDAR and imagery for estimating forest canopy fuels *Remote Sens. Environ.* **114** 725–37
ESA 2022 *Sentinel-1* (available at: www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-1) (Accessed 21 February 2024)
Fassnacht F E, Latifi H, Stereńczak K, Modzelewska A, Lefsky M, Waser L T, Straub C and Ghosh A 2016 Review of studies on tree species classification from remotely sensed data *Remote Sens. Environ.* **186** 64–87
Fassnacht F E, Poblete-Olivares J, Rivero L, Lopatin J, Ceballos-Comisso A and Galleguillos M 2021 Using Sentinel-2 and canopy height models to derive a landscape-level biomass map covering multiple vegetation types *Int. J. Appl. Earth Observ. Geoinf.* **94** 102236
Ferraz A *et al* 2018 Carbon storage potential in degraded forests of Kalimantan, Indonesia *Environ. Res. Lett.* **13** 095001
Ferraz A, Saatchi S, Mallet C, Jacquemoud S, Gonçalves G, Silva C, Soares P, Tomé M and Pereira L 2016 Airborne lidar estimation of aboveground forest biomass in the absence of field inventory *Remote Sens.* **8** 653
Ferro-Famil L and Pottier E 2016 Synthetic aperture radar imaging *Microwave Remote Sensing of Land Surface* ed N Baghdadi and M Zribi (Elsevier) ch 1, pp 1–65
Filbee-Dexter K and Wernberg T 2020 Substantial blue carbon in overlooked Australian kelp forests *Sci. Rep.* **10** 12341
Foody G M 2010 Assessing the accuracy of land cover change with imperfect ground reference data *Remote Sens. Environ.* **114** 2271–85
Forkuor G, Dimobe K, Serme I and Tondoh J E 2018 Landsat-8 vs. Sentinel-2: examining the added value of sentinel-2's red-edge bands to land-use and land-cover mapping in Burkina Faso *GIScience Remote Sens.* **55** 331–54
Fortunel C, Lasky J R, Uriarte M, Valencia R, Wright S J, Garwood N C and Kraft N J B 2018 Topography and neighborhood crowding can interact to shape species growth and distribution in a diverse Amazonian forest *Ecology* **99** 2272–83
Fox I D, Neldner G W, Wilson G W and Bannink P J 2001 *The vegetation of the Australian tropical savannas* Environmental Protection Agency, Brisbane
Friedlingstein P *et al* 2023 Global carbon budget 2023 *Earth Syst. Sci. Data* **15** 5301–69
Gao T, Hedblom M, Emilsson T and Nielsen A B 2014 The role of forest stand structure as biodiversity indicator *For. Ecol. Manage.* **330** 82–93
García M, Saatchi S, Ustin S and Balzter H 2018 Modelling forest canopy height by integrating airborne LiDAR samples with satellite Radar and multispectral imagery *Int. J. Appl. Earth Observ. Geoinf.* **66** 159–73
Geoscience Australia 2020 *Digital Elevation Data, Geoscience Australia* (available at: www.ga.gov.au/scientific-topics/national-location-information/digital-elevation-data) (Accessed 22 March 2024)
Girardin C A J, Jenkins S, Seddon N, Allen M, Lewis S L, Wheeler C E, Griscom B W and Malhi Y 2021 Nature-based solutions can help cool the planet—if we act now *Nature* **593** 191–4
Gitelson A A, Kaufman Y J and Merzlyak M N 1996 Use of a green channel in remote sensing of global vegetation from EOS-MODIS *Remote Sens. Environ.* **58** 289–98
Gleason C J and Im J 2012 Forest biomass estimation from airborne LiDAR data using machine learning approaches *Remote Sens. Environ.* **125** 80–91
Goetz S and Dubayah R 2011 Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change *Carbon Manage.* **2** 231–44
Grierson P, Williams K and Adams M 2000 National Accounting System Technical Report *Review of Unpublished Biomass - Related Information: Western Australia, South Australia, New South Wales and Queensland* University of Western Australia
Griffin E A 1994 *Floristic Survey of Northern Sandplains between Perth and Geraldton* (Department of Primary Industries and Regional Development)
Guo Q, Du S, Jiang J, Guo W, Zhao H, Yan X, Zhao Y and Xiao W 2023 Combining GEDI and sentinel data to estimate forest canopy mean height and aboveground biomass *Ecol. Inf.* **78** 102348
Guo X, Coops N C, Tompalski P, Nielsen S E, Bater C W and John Stadt J 2017 Regional mapping of vegetation structure for biodiversity monitoring using airborne lidar data *Ecol. Inf.* **38** 50–61

Hancock S, Armston J, Hofton M, Sun X, Tang H, Duncanson L I, Kellner J R and Dubayah R 2019 The GEDI simulator: a large-footprint waveform lidar simulator for calibration and validation of spaceborne missions *Earth Space Sci.* **6** 294–310
Hill T C, Williams M, Bloom A A, Mitchard E T A and Ryan C M 2013 Are inventory based and remotely sensed above-ground biomass estimates consistent? *PLoS One* **8** e74170
Hofton M and Blair J B 2019 *Algorithm Theoretical Basis Document* (ATBD)
Hojo A, Avtar R, Nakaji T, Tadono T and Takagi K 2023 Modeling forest above-ground biomass using freely available satellite and multisource datasets *Ecol. Inf.* **74** 101973
Hua Y and Zhao X 2021 Multi-model estimation of forest canopy closure by using red edge bands based on Sentinel-2 images *Forests* **12** 1768
Huete A, Didan K, Miura T, Rodriguez E P, Gao X and Ferreira L G 2002 Overview of the radiometric and biophysical performance of the MODIS vegetation indices *Remote Sens. Environ.* **83** 195–213
Huettermann S, Jones S, Soto-Berelov M and Hislop S 2022 Intercomparison of real and simulated GEDI observations across sclerophyll forests *Remote Sens.* **14** 2096
Hyde P, Dubayah R, Peterson B, Blair J, Hofton M, Hunsaker C, Knox R and Walker W 2005 Mapping forest structure for wildlife habitat analysis using waveform lidar: validation of montane ecosystems *Remote Sens. Environ.* **96** 427–37
IBRA 2018 *Interim Biogeographic Regionalisation for Australia (IBRA), Version 7 (Subregions)—SEED* (available at: https://datasets.seed.nsw.gov.au/dataset/interim-biogeographic-regionalisation-for-australia-ibra-version-7-subregions) (Accessed 21 February 2024)
Iqbal I A, Dash J, Ullah S and Ahmad G 2013 A novel approach to estimate canopy height using ICESat/GLAS data: a case study in the new forest national park, UK *Int. J. Appl. Earth Observ. Geoinf.* **23** 109–18
Jackson S, Palmer L, McDonald F and Bumpus A 2017 Cultures of Carbon and the Logic of Care: The Possibilities for Carbon Enrichment and Its Cultural Signature *Ann. Assoc. Am. Geogr.* **107** 867–82
Jiang F, Sun H, Ma K, Fu L and Tang J 2022 Improving aboveground biomass estimation of natural forests on the Tibetan Plateau using spaceborne LiDAR and machine learning algorithms *Ecol. Indic.* **143** 109365
Jiang Z, Huete A, Didan K and Miura T 2008 Development of a two-band enhanced vegetation index without a blue band *Remote Sens. Environ.* **112** 3833–45
Joshi N, Mitchard E T A, Brolly M, Schumacher J, Fernández-Landa A, Johannsen V K, Marchamalo M and Fensholt R 2017 Understanding "saturation" of radar signals over forests *Sci. Rep.* **7** 3505
Jucker T *et al* 2017 Allometric equations for integrating remote sensing imagery into forest monitoring programmes *Glob. Change Biol.* **23** 177–90
Jucker T, Bongalov B, Burslem D F R P, Nilus R, Dalponte M, Lewis S L, Phillips O L, Qie L and Coomes D A 2018 Topography shapes the structure, composition and function of tropical forest landscapes *Ecol. Lett.* **21** 989–1000
Keith D A 2017 *Australian Vegetation* (Cambridge University Press) (available at: https://assets.cambridge.org/97811071/18430/frontmatter/9781107118430_frontmatter.pdf)
Key C H and Benson N C 2006 *Landscape Assessment: Ground Measure of Severity, the Composite Burn Index; and Remote Sensing of Severity, the Normalized Burn Ratio* (USDA Forest Service, Rocky Mountain Research Station) (available at: https://pubs.usgs.gov/publication/2002085) (Accessed 27 March 2024)
Kuhn M 2008 Building predictive models in R using the caret package *J. Stat. Softw.* **28** 1–26
Labriere N *et al* 2018 *In situ* reference datasets from the TropiSAR and AfriSAR campaigns in support of upcoming spaceborne biomass missions *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **11** 3617–27
Lang N, Jetz W, Schindler K and Wegner J D 2023 A high-resolution canopy height model of the Earth *Nat. Ecol. Evol.* **7** 1778–89
Lang N, Kalischek N, Armston J, Schindler K, Dubayah R and Wegner J D 2022 Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles *Remote Sens. Environ.* **268** 112760
Lefsky M A, Harding D J, Keller M, Cohen W B, Carabajal C C, Del Bom Espirito-Santo F, Hunter M O and de Oliveira R 2005 Estimates of forest canopy height and aboveground biomass using ICESat *Geophys. Res. Lett.* **32** 2005GL023971
Leite R V *et al* 2022 Large scale multi-layer fuel load characterization in tropical savanna using GEDI spaceborne lidar data *Remote Sens. Environ.* **268** 112764
Li C, Li M, Li Y and Qian P 2020a Estimating aboveground forest carbon density using Landsat 8 and field-based data: a comparison of modelling approaches *Int. J. Remote Sens.* **41** 4269–92
Li J, Hong D, Gao L, Yao J, Zheng K, Zhang B and Chanussot J 2022a Deep learning in multimodal remote sensing data fusion: a comprehensive review *Int. J. Appl. Earth Observ. Geoinf.* **112** 102926
Li W, Niu Z, Shang R, Qin Y, Wang L and Chen H 2020b High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 data *Int. J. Appl. Earth Observ. Geoinf.* **92** 102163
Li Z, Bi S, Hao S and Cui Y 2022b Aboveground biomass estimation in forests with random forest and Monte Carlo-based uncertainty analysis *Ecol. Indic.* **142** 109246
Liao Z, Van Dijk A I J M, He B, Larraondo P R and Scarth P F 2020 Woody vegetation cover, height and biomass at 25-m resolution across Australia derived from multiple site, airborne and satellite observations *Int. J. Appl. Earth Observ. Geoinf.* **93** 102209
Liaw A and Wiener M 2002 Classification and Regression by randomForest *RNews* **2** 18–22
Liu M and Popescu S 2022 Estimation of biomass burning emissions by integrating ICESat-2, Landsat 8, and Sentinel-1 data *Remote Sens. Environ.* **280** 113172
Low A and Lamont B 1990 Aerial and belowground phytomass of Banksia scrub-heath at eneabba, South-Western Australia *Aust. J. Bot.* **38** 351–9
LP DAAC 2021 *GEDI Subsetter Python Script* (available at: Https://Git.Earthdata.Nasa.Gov/Projects/LPDUR/Repos/Gedi-Subsetter/Browse)
Lu D, Chen Q, Wang G, Liu L, Li G and Moran E 2016 A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems *Int. J. Digit. Earth* **9** 63–105
Luckman A, Baker J, Honzák M and Lucas R 1998 Tropical forest biomass density estimation using JERS-1 SAR: seasonal variation, confidence limits, and application to image mosaics *Remote Sens. Environ.* **63** 126–39
Luo S *et al* 2019 Estimating forest aboveground biomass using small-footprint full-waveform airborne LiDAR data *Int. J. Appl. Earth Observ. Geoinf.* **83** 101922
McCaw L and Middleton T 2015 Recovery of tall open eucalypt forest in South-Western Australia following complete crown scorch *Fire Ecol.* **11** 95–107
McCaw W L, Neal J E and Smith R H 1996 Fuel accumulation following prescribed burning in young even-aged stands of karri (Eucalyptus diversicolor) *Aust. For.* **59** 171–7

McCaw W L, Robinson R and Williams M 2011 *Integrated Biodiversity Monitoring for the Jarrah (Eucalyptus Marginata) Forest in South-West Western Australia: the FORESTCHECK Project* (https://doi.org/10.1080/00049158.2011.10676369?needAccess=true) (Accessed 22 February 2024)
McDougall K L, Barrett S, Velzeboer R, Cahill D M, Rudman T and Reiter N 2024 Evaluating the risk to Australia's flora from Phytophthora cinnamomi *Aust. J. Bot.* **72**
McFeeters S K 1996 The use of the normalized difference water index (NDWI) in the delineation of open water features *Int. J. Remote Sens.* **17** 1425–32
Meyer V, Saatchi S, Clark D B, Keller M, Vincent G, Ferraz A, Espírito-Santo F, d'Oliveira M V N, Kaki D and Chave J 2018 Canopy area of large trees explains aboveground biomass variations across neotropical forest landscapes *Biogeosciences* **15** 3377–90
Millard K and Richardson M 2015 On the importance of training data sample selection in random forest image classification: a case study in peatland ecosystem mapping *Remote Sens.* **7** 8489–515
Mitchard E T A, Saatchi S S, Lewis S L, Feldpausch T R, Woodhouse I H, Sonké B, Rowland C and Meir P 2011 Measuring biomass changes due to woody encroachment and deforestation/degradation in a forest–savanna boundary region of central Africa using multi-temporal L-band radar backscatter *Remote Sens. Environ.* **115** 2861–73
Mitchard E T A, Saatchi S S, Woodhouse I H, Nangendo G, Ribeiro N S, Williams M, Ryan C M, Lewis S L, Feldpausch T R and Meir P 2009 Using satellite radar backscatter to predict above-ground woody biomass: a consistent relationship across four different African landscapes *Geophys. Res. Lett.* **36** L23401
Montesano P M, Cook B D, Sun G, Simard M, Nelson R F, Ranson K J, Zhang Z and Luthcke S 2013 Achieving accuracy requirements for forest biomass mapping: a spaceborne data fusion method for estimating forest biomass and LiDAR sampling error *Remote Sens. Environ.* **130** 153–70
Morais T G, Teixeira R F M, Figueiredo M and Domingos T 2021 The use of machine learning methods to estimate aboveground biomass of grasslands: a review *Ecol. Indic.* **130** 108081
Mugabowindekwe M *et al* 2023 Nation-wide mapping of tree-level aboveground carbon stocks in Rwanda *Nat. Clim. Change* **13** 91–97
Mulverhill C, Coops N C, Hermosilla T, White J C and Wulder M A 2022 Evaluating ICESat-2 for monitoring, modeling, and update of large area forest canopy height products *Remote Sens. Environ.* **271** 112919
NAFI 2023 *Northern Australian Fire Information* (available at: https://firenorth.org.au/nafi3/) (Accessed 22 March 2024)
Naidoo L, Mathieu R, Main R, Wessels K and Asner G P 2016 L-band synthetic aperture radar imagery performs better than optical datasets at retrieving woody fractional cover in deciduous, dry savannahs *Int. J. Appl. Earth Observ. Geoinf.* **52** 54–64
Nandy S, Srinet R and Padalia H 2021 Mapping forest height and aboveground biomass by integrating ICESat-2, Sentinel-1 and Sentinel-2 data using random forest algorithm in Northwest Himalayan Foothills of India *Geophys. Res. Lett.* **48** e2021GL093799
Narine L L, Popescu S C and Malambo L 2020 Using ICESat-2 to estimate and map forest aboveground biomass: a first example *Remote Sens.* **12** 1824
NVIS Technical Working Group 2017 *Australian Vegetation Attribute Manual: National Vegetation Information System, Version 7.0* (Department of the Environment and Energy)
O'Grady A P, Chen X, Eamus D and Hutley L B 2000 Composition, leaf area index and standing biomass of eucalypt open forests near Darwin in the Northern Territory, Australia *Aust. J. Bot.* **48** 629–38
Oliveira P, Zhang X, Peterson B and Ometto J P 2023 Using simulated GEDI waveforms to evaluate the effects of beam sensitivity and terrain slope on GEDI L2A relative height metrics over the Brazilian Amazon Forest *Remote Sens. Environ.* **7** 100083
Olofsson P, Foody G M, Herold M, Stehman S V, Woodcock C E and Wulder M A 2014 Good practices for estimating area and assessing accuracy of land change *Remote Sens. Environ.* **148** 42–57
Paul K I *et al* 2016 Testing the generality of above-ground biomass allometry across plant functional types at the continent scale *Glob. Change Biol.* **22** 2106–24
Paul K I and Roxburgh S H 2020 Predicting carbon sequestration of woody biomass following land restoration *For. Ecol. Manage.* **460** 117838
Pettorelli N, Vik J O, Mysterud A, Gaillard J-M, Tucker C J and Stenseth N C 2005 Using the satellite-derived NDVI to assess ecological responses to environmental change *Trends Ecol. Evol.* **20** 503–10
Potapov P *et al* 2021 Mapping global forest canopy height through integration of GEDI and Landsat data *Remote Sens. Environ.* **253** 112165
Powell S L, Cohen W B, Healey S P, Kennedy R E, Moisen G G, Pierce K B and Ohmann J L 2010 Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: a comparison of empirical modeling approaches *Remote Sens. Environ.* **114** 1053–68
Qi J, Chehbouni A, Huete A R, Kerr Y H and Sorooshian S 1994 A modified soil adjusted vegetation index *Remote Sens. Environ.* **48** 119–26
R Core Team 2023 *R Foundation for Statistical Computing* (available at: www.R-project.org/)
Rahman M F, Onoda Y and Kitajima K 2022 Forest canopy height variation in relation to topography and forest types in central Japan with LiDAR *For. Ecol. Manage.* **503** 119792
Raison J *et al* 2003 Spatial estimates of biomass in "Mature" native vegetation
Rayner M E 1991 Site index and dominant height growth curves for regrowth karri (*Eucalyptus diversicolor* F. Muell.) in south-western Australia *For. Ecol. Manage.* **44** 261–83
Réjou-Méchain M *et al* 2014 Local spatial structure of forest biomass and its consequences for remote sensing of carbon stocks *Biogeosciences* **11** 6827–40
Réjou-Méchain M *et al* 2019 Upscaling forest biomass from field to satellite measurements: sources of errors and ways to reduce them *Surv. Geophys.* **40** 881–911
Renwick A R, Robinson C J, Martin T G, May T, Polglase P, Possingham H P and Carwardine J 2014 Biodiverse planting for carbon and biodiversity on indigenous land *PLoS One* **9** e91281
Rishmawi K, Huang C and Zhan X 2021 Monitoring key forest structure attributes across the conterminous United States by integrating GEDI LiDAR measurements and VIIRS data *Remote Sens.* **13** 442
Rix M G, Edwards D L, Byrne M, Harvey M S, Joseph L and Roberts J D 2015 Biogeography and speciation of terrestrial fauna in the south-western Australian biodiversity hotspot *Biol. Rev.* **90** 762–93
Rodríguez-Veiga P *et al* 2019 Forest biomass retrieval approaches from earth observation in different biomes *Int. J. Appl. Earth Observ. Geoinf.* **77** 53–68

Rodríguez-Veiga P, Carreiras J, Smallman T, Exbrayat J-F, Ndambiri J, Mutwiri F, Nyasaka D, Quegan S, Williams M and Balzter H 2020 Carbon stocks and fluxes in kenyan forests and wooded grasslands derived from earth observation and model-data fusion *Remote Sens.* **12** 2380

Rodríguez-Veiga P, Saatchi S, Tansey K and Balzter H 2016 Magnitude, spatial distribution and uncertainty of forest biomass stocks in Mexico *Remote Sens. Environ.* **183** 265–81

Roxburgh S H, Karunaratne S B, Paul K I, Lucas R M, Armston J D and Sun J 2019 A revised above-ground maximum biomass layer for the Australian continent *For. Ecol. Manag.* **432** 264–75

Russell-Smith J, Yates C P, Edwards A C, Whitehead P J, Murphy B P and Lawes M J 2015 Deriving multiple benefits from carbon market-based savanna fire management: an Australian example *PLoS One* **10** e0143426

Saarela S, Holm S, Healey S, Andersen H-E, Petersson H, Prentius W, Patterson P, Næsset E, Gregoire T and Ståhl G 2018 Generalized hierarchical model-based estimation for aboveground biomass assessment using GEDI and landsat data *Remote Sens.* **10** 1832

Salum R B, Souza-Filho P W M, Simard M, Silva C A, Fernandes M E B, Cougo M F, Do Nascimento W and Rogers K 2020 Improving mangrove above-ground biomass estimates using LiDAR *Estuar. Coast. Shelf Sci.* **236** 106585

Schneider F D *et al* 2020 Towards mapping the diversity of canopy structure from space with GEDI *Environ. Res. Lett.* **15** 115006

Seely H, Coops N C, White J C, Montwé D, Winiwarter L and Ragab A 2023 Modelling tree biomass using direct and additive methods with point cloud deep learning in a temperate mixed forest *Sci. Remote Sens.* **8** 100110

Shendryk Y 2022 Fusing GEDI with earth observation data for large area aboveground biomass mapping *Int. J. Appl. Earth Observ. Geoinf.* **115** 103108

Silva C A *et al* 2021 Fusing simulated GEDI, ICESat-2 and NISAR data for regional aboveground biomass mapping *Remote Sens. Environ.* **253** 112234

Simard M, Pinto N, Fisher J B and Baccini A 2011 Mapping forest canopy height globally with spaceborne lidar *J. Geophys. Res.* **116** G04021

Specht R 1970 *Vegetation. In The Australian Environment* (University Press)

Staben G, Lucieer A and Scarth P 2018 Modelling LiDAR derived tree canopy height from Landsat TM, ETM+ and OLI satellite imagery—A machine learning approach *Int. J. Appl. Earth Observ. Geoinf.* **73** 666–81

Standish R J and Prober S M 2020 Potential benefits of biodiversity to Australian vegetation projects registered with the emissions reduction fund—is there a carbon-biodiversity trade-off? *Ecol. Manage. Restor.* **21** 165–72

Stojanova D, Panov P, Gjorgjioski V, Kobler A and Džeroski S 2010 Estimating vegetation height and canopy cover from remotely sensed data with machine learning *Ecol. Inf.* **5** 256–66

Stovall A E L, Anderson-Teixeira K J and Shugart H H 2018 Assessing terrestrial laser scanning for developing non-destructive biomass allometry *For. Ecol. Manage.* **427** 217–29

Tang H, Ma L, Lister A, O'Neill-Dunne J, Lu J, Lamb R L, Dubayah R and Hurtt G 2021 High-resolution forest carbon mapping for climate mitigation baselines over the RGGI region, USA *Environ. Res. Lett.* **16** 035011

TERN (Terrestrial Ecosystem Research Network) 2021 Biomass Plot Library - National collation of stem inventory data and biomass estimation, Australian field sites. Version 1.0. Terrestrial Ecosystem Research Network *Joint Remote Sensing Research Program* (https://doi.org/10.4227/05/552486484985D)

Tillack A, Clasen A, Kleinschmit B and Förster M 2014 Estimation of the seasonal leaf area index in an alluvial forest using high-resolution satellite-based vegetation indices *Remote Sens. Environ.* **141** 52–63

Torres de Almeida C, Gerente J, Rodrigo Dos Prazeres Campos J, Caruso Gomes Junior F, Providelo L A, Marchiori G and Chen X 2022 Canopy height mapping by Sentinel 1 and 2 satellite images, airborne LiDAR data, and machine learning *Remote Sens.* **14** 4112

Tran B N, Tanase M A, Bennett L T and Aponte C 2018 Evaluation of spectral indices for assessing fire severity in Australian temperate forests *Remote Sens.* **10** 1680

Tubiello F N, Conchedda G, Wanner N, Federici S, Rossi S and Grassi G 2021 Carbon emissions and removals from forests: new estimates, 1990–2020′ *Earth Syst. Sci. Data* **13** 1681–91

Tucker C J 1979 Red and photographic infrared linear combinations for monitoring vegetation *Remote Sens. Environ.* **8** 127–50

USGS 2022 Landsat 8 collection 1 Tier 1 8-day EVI composite | earth engine data catalog (available at: https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C01_T1_8DAY_EVI) (Accessed 28 March 2024)

Ustin S L, ROBERTS D A, Gamon J A, ASNER G P and GREEN R O 2004 Using imaging spectroscopy to study ecosystem processes and properties *BioScience* **54** 523–34

Viana-Soto A, Okujeni A, Pflugmacher D, García M, Aguado I and Hostert P 2022 Quantifying post-fire shifts in woody-vegetation cover composition in Mediterranean pine forests using Landsat time series and regression-based unmixing *Remote Sens. Environ.* **281** 113239

Vigilante T and Bowman D M J S 2004 Effects of fire history on the structure and floristic composition of woody vegetation around Kalumburu, North Kimberley, Australia: a landscape-scale natural experiment *Aust. J. Bot.* **52** 381–404

Volkova L, Bi H, Murphy S and Weston C 2015 Empirical estimates of aboveground carbon in open eucalyptus forests of South-Eastern Australia and its potential implication for national carbon accounting *Forests* **6** 3395–411

Wang C, Elmore A J, Numata I, Cochrane M A, Shaogang L, Huang J, Zhao Y and Li Y 2022 Factors affecting relative height and ground elevation estimations of GEDI among forest types across the conterminous USA *GIScience Remote Sens.* **59** 975–99

Wang H, Seaborn T, Wang Z, Caudill C C and Link T E 2021 Modeling tree canopy height using machine learning over mixed vegetation landscapes *Int. J. Appl. Earth Observ. Geoinf.* **101** 102353

Wang L, Zhou X, Zhu X, Dong Z and Guo W 2016 Estimation of biomass in wheat using random forest regression algorithm and remote sensing data *Crop J.* **4** 212–9

Wardell-Johnson G, Crellin L, Napier C, Meigs G, Stevenson A and Wong S I 2017 Has canopy height and biomass recovered 78 years after an intense fire in south-western Australia's red tingle (Eucalyptus jacksonii) forests? *Int. J. Wildland Fire* **26** 148–55

Weisbin C R, Lincoln W and Saatchi S 2014 A systems engineering approach to estimating uncertainty in above-ground biomass (AGB) derived from remote-sensing data *Syst. Eng.* **17** 361–73

Werner P A and Peacock S J 2019 Savanna canopy trees under fire: long-term persistence and transient dynamics from a stage-based matrix population model *Ecosphere* **10** e02706

Wickham H 2016 *ggplot2: Elegant Graphics for Data Analysis* (Springer) (available at: https://ggplot2.tidyverse.org)

Wilkes P, Jones S, Suarez L, Mellor A, Woodgate W, Soto-Berelov M, Haywood A and Skidmore A 2015 Mapping forest canopy height across large areas by upscaling ALS estimates with freely available satellite data *Remote Sens.* **7** 12563–87

Wood S W, Prior L D, Stephens H C and Bowman D M J S 2015 Macroecology of Australian tall eucalypt forests: baseline data from a continental-scale permanent plot network *PLoS One* **10** e0137811

Wu N, Crusiol L G T, Liu G, Wuyun D and Han G 2023 Comparing the performance of machine learning algorithms for estimating aboveground biomass in typical steppe of northern China using Sentinel imageries *Ecol. Indic.* **154** 110723

Wysong M *et al* 2021 The sum of small parts: changing landscape fire regimes across multiple small landholdings in north-western Australia with collaborative fire management *Int. J. Wildland Fire* **31** 97–111

Xu L *et al* 2017 Spatial distribution of carbon stored in forests of the Democratic Republic of Congo *Sci. Rep.* **7** 15030

Yates C, Evans J, Vernooij R, Eames T, Muir E, Holmes J, Edwards A and Russell-Smith J 2023 Incentivizing sustainable fire management in Australia's northern arid spinifex grasslands *J. Environ. Manage.* **344** 118384

Zhang G, Ganguly S, Nemani R R, White M A, Milesi C, Hashimoto H, Wang W, Saatchi S, Yu Y and Myneni R B 2014 Estimation of forest aboveground biomass in California using canopy height and leaf area index estimated from satellite data *Remote Sens. Environ.* **151** 44–56

Zhang X *et al* 2024 Improved random forest algorithms for increasing the accuracy of forest aboveground biomass estimation using Sentinel-2 imagery *Ecol. Indic.* **159** 111752

Zhao K, Popescu S and Nelson R 2009 Lidar remote sensing of forest biomass: a scale-invariant estimation approach using airborne lasers *Remote Sens. Environ.* **113** 182–96