



Australia's National  
Science Agency

# UAV–LiDAR and spaceborne remote sensing for habitat structure monitoring in the Beetaloo

Shaun R. Levick, Stephen B. Stewart, Stephanie M. Johnson,  
Pascal Castellazzi, and Chris Pavey

June 2025

## Citation

Levick SR, SB Stewart, SM Johnson, P Castellazzi, and C Pavey. 2025. UAV–LiDAR and spaceborne remote sensing habitat structure monitoring in the Beetaloo. CSIRO, Australia.

## Copyright

© Commonwealth Scientific and Industrial Research Organisation 2025. To the extent permitted by law, all rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of CSIRO.

## Important disclaimer

CSIRO advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, CSIRO (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

CSIRO is committed to providing web accessible content wherever possible. If you are having difficulties with accessing this document please contact [csiro.au/contact](https://www.csiro.au/contact).

# Contents

Acknowledgments .....	vii
Executive summary.....	viii
1 Background.....	1
1.1 Prior Research .....	1
2 Research project goals .....	3
2.1 Study objective.....	3
3 Research approach and methods.....	5
3.1 Study site description.....	5
3.2 UAV-LiDAR acquisition and processing .....	5
3.3 Spaceborne Synthetic Aperture Radar (SAR) processing .....	9
3.4 Spaceborne multi-spectral image processing .....	10
4 Statistical analyses and modelling.....	12
4.1 Structural metrics from LiDAR point-clouds.....	12
4.2 Explanatory variables and spatial sampling .....	15
4.3 Correlations between UAV-LiDAR structural metrics and satellite remote sensing 15	
4.4 Predictive modelling with quantile regression forests.....	15
5 Study results.....	17
5.1 Correlations between LiDAR structural metrics and SAR.....	17
5.2 Predictive modelling of UAV-LiDAR structural metrics .....	20
5.3 Extrapolating predictions of vegetation 3D structure over a larger extent .....	25
5.4 Quantifying temporal dynamics .....	28
6 Outlook for large area structural monitoring.....	32
6.1 Key findings from this study.....	32
6.2 Future directions .....	33
6.3 Conclusion .....	34
7 References.....	35
8 Appendix 1.....	38
8.1 Additional datasets and processing steps .....	38

9 Appendix 2.....40  
9.1 Supporting figures.....40

# Figures

Figure 1 – UAV-LiDAR flights were conducted out of Daly Waters in Central Beetaloo region (a). Surveys areas (green boxes) were concentrated along the southern side of the Carpentaria highway (b and c) and aimed to capture diverse vegetation patch structures. ....	4
Figure 2 - Acecore Noa hexacopter with Riegl VUX-120 LiDAR sensor (a). Typical launch site during Beetaloo field campaigns (b).....	6
Figure 3 - Typical site survey with eight flightlines (white) of 1 km each, spaced 120 m apart, with the resulting point cloud displayed underneath. Warmer colours in the point cloud represent higher elevation above sea level.....	7
Figure 4 – Digital terrain model (DTM) derived from UAV-LiDAR. Note detection of lower lying drainage patches and road verges.....	8
Figure 5 – Vegetation canopy height model (CHM) derived from UAV-LiDAR surveys, showing good delineation of roads and vegetation patches (a) and individual trees (b & c). note the detection of dead trees in (c). ....	8
Figure 6 – Examples of the point cloud data contained within 20 m ‘pixels’, showing the range of structures that are integrated into each SAR and SR pixel.....	12
Figure 7 – Derivation of structural metrics from UAV-LiDAR at a single site. (a) Canopy height from the normalised point-cloud. Examples of metrics computed to match the 20 m pixels of the spaceborne imagery: (b) 98 <sup>th</sup> percentile height, (c) GFP, (d) PAI, (e) std, and (f) density 5-10m.....	14
Figure 8 – Scatterplots of the relationship between VH backscatter and (a) Gap Fraction Probability (GFP), (b) Plant Area Index (PAI), (c) 75 <sup>th</sup> percentile height (p75) and (d) density above 1 m height (n = 2,923).....	17
Figure 9 - Pairwise correlations between SAR (y-axes) and LiDAR (x-axes) variables for (a) annual, (b) austral winter, (c) austral spring, (d) austral summer, (e) austral autumn, and (f) month of LiDAR acquisition median compositing periods. All comparisons were conducted on a randomised spatial sample of point locations (n = 2,923) using Spearman’s rank-based correlation coefficient (rho). Non-significant (p ≥ 0.05) correlations are not reported (grey cells). ....	18
Figure 10 - Absolute difference in pairwise correlations between SAR (y-axes) and LiDAR (x-axes) variables for (a) austral winter, (b) austral spring, (c) austral summer, (d) austral autumn, and (e) month of LiDAR acquisition in comparison to annual median compositing periods (i.e., blue values indicate better correlations with lidar variables than the annual composite). All comparisons were conducted on a randomised spatial sample of point locations (n = 2,923) using Spearman’s rank-based correlation coefficient (rho). Non-significant (p ≥ 0.05) correlations are not reported (grey cells).....	19
Figure 11 – Cross validated (a) Lin’s Concordance Correlation Coefficient and (b) root mean squared error (RMSE) / mean for predictive models (quantile regression forests) of LiDAR structural metrics (x-axes) using SAR variables (y-axes). Statistics were calculated using 10-fold cross-validation on 70% of samples selected for model training (n = 2,043). ....	20

Figure 12 – Observed versus predicted (a) Gap Fraction Probability (GFP), (b) Plant Area Index (PAI), (c) 75<sup>th</sup> percentile height (p75), and (c) density above 1 m height (dens\_>1m) at held-out point locations (n = 880) when modelled using SAR variables. The red dashed line shows the 1:1 relationship between observed and predicted values. Model names are specified above each plot. Lin’s Concordance Correlation Coefficient (CCC) is reported in the top left corner of each subplot.....22

Figure 13 – Observed versus predicted (a) Gap Fraction Probability (GFP), (b) Plant Area Index (PAI), (c) 75<sup>th</sup> percentile height (p75), and (c) density above 1 m height (dens\_>1m) at held-out point locations (n = 880) when modelled using both SAR and SR variables. The red dashed line shows the 1:1 relationship between observed and predicted values. Model names are specified above each plot. Lin’s Concordance Correlation Coefficient (CCC) is reported in the top left corner of each subplot. ....24

Figure 14 – Modelled (a) Gap Fraction Probability (GFP), (b) width of the GFP prediction interval (PI), (c) difference in modelled GFP when using surface reflectance (SR) in addition to SAR for prediction, and (d) difference in the prediction interval width when using both SR and SAR for prediction (blue indicates reduced uncertainty). Q10, Q50 and Q90 represent the 10<sup>th</sup>, 50<sup>th</sup> and 90<sup>th</sup> percentile of predictions, respectively. Model names and their order for the purposes of calculating differences are specified above each plot.....25

Figure 15 – Modelled (a) Plant Area Index (PAI), (b) width of the PAI prediction interval (PI), (c) difference in modelled PAI when using surface reflectance (SR) in addition to SAR for prediction, and (d) difference in the prediction interval width when using both SR and SAR for prediction (blue indicates reduced uncertainty). Q10, Q50 and Q90 represent the 10<sup>th</sup>, 50<sup>th</sup> and 90<sup>th</sup> percentile of predictions, respectively. Model names and their order for the purposes of calculating differences are specified above each plot.....26

Figure 16 - Modelled (a) return density above 1 m (dens\_>1m), (b) width of the prediction interval (PI) for density above 1 m, (c) difference in modelled density above 1 m when using surface reflectance (SR) in addition to SAR for prediction, and (d) difference in the prediction interval width when using both SR and SAR for prediction (blue indicates reduced uncertainty). Q10, Q50 and Q90 represent the 10<sup>th</sup>, 50<sup>th</sup> and 90<sup>th</sup> percentile of predictions, respectively. Model names and their order for the purposes of calculating differences are specified above each plot.....27

Figure 17 – Modelled predictions of (a) PAI and (b) density above 1 m height for year beginning 01/June/2024, the difference in (c) PAI and (d) density above 1 m height since the year beginning 01/June/2019, and the annual trend in mean (e) PAI and (f) density above 1 m height calculated across the whole study region. The decrease in PAI and density above 1 m height (c, d) is associated with development activity, including roadworks. Red shading (e, f) shows the prediction interval given by the 10<sup>th</sup> and 90<sup>th</sup> percentile. ....29

Figure 18 – Visualisation of change in density above 1 m (dens\_>1m) over time in a subset of the study region. Images show annual RGB composites for (a) the year beginning 01/June/2019 and (e) year beginning 01/June/2023, return density above 1 m height for (b) the year beginning 01/June/2019 and (f) year beginning 01/June/2019, and change in density above 1 m between (c) 01/June/2019 and 31/May/2021, and (d) 01/June/2022 and 31/May/2024. ....30

Figure 19 – Time-series of Plant Area Index (PAI) and density above 1 m height (dens\_>1m) at (a) three point locations where vegetation cover (b, c) decreases, (d, e) remains stable, and (f, g) increases. Red shading shows the prediction interval given by the 10th and 90th percentile. ....31

Figure 20 - Pairwise correlations between LiDAR structural metrics. All comparisons were conducted on a randomised spatial sample of point locations (n = 2,923) using Spearman’s rank-based correlation coefficient (rho). Non-significant (p ≥ 0.05) correlations are not reported (grey cells). ....40

Figure 21 – Scatterplots of spatial random samples (n = 15,000) of (a) Gap Fraction Probability (GFP) versus density above 1 m height (dens\_>1m), (b) Plant Area Index (PAI) versus density above 1 m height, (c) 75<sup>th</sup> percentile height (p75) versus density above 1 m height, and (d) PAI versus GFP. Point samples were extracted from extrapolated predictions from models fitted using SAR and SR as explanatory variables.....41

# Tables

Table 1 – Sentinel-2A and 2B multi-spectral bands used for modelling LiDAR structural metrics (Geoscience Australia, 2022a, 2022b). .....	11
Table 2 – LiDAR structural metrics evaluated for correlations with SAR and used as dependent variables for modelling vegetation structure. All metrics were calculated with a spatial resolution of 20 m. ....	13
Table 3 – Validation statistics for held-out point locations (n = 880) and optimised parameters (mtry, min_n) for predictive models of LiDAR structural metrics using SAR variables. Specific statistics reported are the coefficient of determination ( $R^2$ ), Lin’s Concordance Correlation Coefficient (CCC), root mean squared error (RMSE), bias (mean error) and the mean value of the observations (Mean). ....	21
Table 4 – Validation statistics for held-out point locations (n = 880) and optimised parameters (mtry, min_n) for predictive models of LiDAR structural metrics using both SAR and SR variables. Specific statistics reported are the coefficient of determination ( $R^2$ ), Lin’s Concordance Correlation Coefficient (CCC), root mean squared error (RMSE), bias (mean error) and the mean value of the observations (Mean). ....	23

# Acknowledgments

This research was funded through CSIRO's Gas Industry Social and Environmental Research Alliance (GISERA) with contributions from the Australian Government's Department of Industry, Science and Resources. GISERA is a collaboration between CSIRO, Commonwealth, state and territory governments and industry established to research the impacts of onshore gas exploration and development on the environment and the socioeconomic costs and benefits. For information about GISERA's governance structure, projects and research findings, visit <https://gisera.csiro.au>

We thank the members of the Technical Reference Group (TRG) for providing advice and guidance. The TRG included Tim Whiteside (Department of Climate Change, Energy, the Environment and Water) and Ian Leiper (Flora and Fauna Division, Northern Territory Government).

We thank Matt Paget, Adriana Parra Ruiz, and Zheng-Shu Zhou for their assistance with Sentinel-1 SAR data processing workflows. Rafael Bohn Reckziegel and Kahill Mitchell's assistance with UAV-LiDAR fieldwork was greatly appreciated.

Sana Khan and Adriana Parra Ruiz provided excellent and timely feedback on drafts of this report. The project team is very appreciative of Jizelle Khoury for her timely and thorough support throughout this project. We thank Paul Cunningham and the GISERA communications team for excellent work on the communications component of the project.

# Executive summary

The development of an onshore gas industry in the Beetaloo region has the potential to impact on terrestrial biodiversity, ecosystem function and landscape integrity. Recent inquiries into hydraulic fracturing in the region have noted potential impact pathways on a broad range of environmental values, highlighting the need for mitigation measures and ongoing monitoring to assess impacts. The community has concerns about potential long-term impacts on the biodiversity of the region. They need reassurance that potential impacts are being mitigated and that any effects of resource development can be detected and managed.

Petroleum interest holders are required to describe the existing environment that may be affected by proposed activities in their Environmental Management Plans and must have a monitoring program that allows for auditing of the environmental performance of the activity. Field surveys are spatially and temporally limited, and the sensitivity of satellite remote sensing to subtle disturbances, and those that compound through time, is still poorly understood. Robust, repeatable, and easily deployable technologies that allow habitat condition to be reliably measured and monitored will be highly valuable in achieving these requirements and will provide reassurance to the broader community.

This project explored the development of a scalable and systematic approach for monitoring the structural condition of vegetation in the Beetaloo region. We aimed to: i) test the sensitivity of Synthetic Aperture Radar (SAR) imagery to spatio-temporal variation in habitat structural properties; ii) use machine learning to establish relationships between UAV-LiDAR structural metrics and spaceborne imagery using the high-quality UAV-LiDAR data for calibration and validation; and iii) use these models to predict vegetation structure, and structural change, across broader landscapes.

High-quality LiDAR point clouds were captured over diverse vegetation structures with survey grade UAV equipment in October 2023 and April 2024. Sentinel-1 SAR backscatter intensity images were processed as monthly, seasonal and annual composites from July 2019 to June 2024. UAV-LiDAR structural metrics were strongly correlated with Sentinel-1 backscatter intensity across all temporal aggregation periods, but the strongest correlations were found late in the dry season. Random Forest modelling showed that key structural metrics (such as Gap Fraction Probability, Plant Area Index, canopy density, height diversity, and height percentiles) could be predicted from combinations of Sentinel-1 VH and VV backscatter intensity across large landscape scales after training the models with high-quality point cloud data. The inclusion of Sentinel-2 multispectral imagery in the models further improved the predictive reliability, with Concordance Correlation Coefficients ranging from 0.76-0.89 for the metrics mentioned above.

Results from this study highlight the power of fusing open access and readily available SAR and multi-spectral imagery for regional scale habitat characterisation and monitoring. The collection of high-quality UAV-LiDAR data provided a strong basis for building predictive machine learning models, closing the gap between what data can be collected in the field, and what can be estimated from space. These findings lay the foundation for systematic on-going monitoring of vegetation structure over the broader region. This approach addresses the limitations of spatial

and temporal limitations of field surveys, and greatly improves the utility of satellite remote sensing. A systematic monitoring program using the methods described in this study may provide an objective dataset that reassures stakeholders, including the broader community, that potential impacts are detected and mitigated.

These models can undergo continuous improvement through the inclusion of additional LiDAR collections, and by building long-term temporal records of vegetation structure from LiDAR to better assess sensitivity to dynamics. Temporal trends in vegetation structure in these landscapes need to be interpreted in the context of broader patterns in landscape water availability. Leveraging additional information from spaceborne sensors, such as gravity measurements and SAR phase coherence, will assist in attributing causality to vegetation dynamics emerging through statistical or machine learning models. Understanding longer-term seasonal and inter-annual patterns of change is imperative for assessing potential impacts arising from development.



# 1 Background

The development of an onshore gas industry in the Beetaloo region has the potential to impact terrestrial biodiversity, ecosystem function and landscape integrity. Community concerns about these impacts were noted in *The Final Report of the Inquiry into Hydraulic Fracturing in the Northern Territory* (2018). The Geological and Bioregional Assessment (GBA) for the Beetaloo region (Geological and Bioregional Assessments Program, 2021) also noted a number of impact pathways with potential to affect a broad range of environmental values (see [GBA website](#)). Nevertheless, the GBA suggested that these potential impacts could be mitigated and emphasised the importance of ongoing monitoring to assess impacts.

Monitoring trends in habitat condition and attributing causality to specific drivers is extremely challenging in heterogenous ecosystems. Traditional monitoring approaches rely on field-based estimates of habitat, but these approaches are limited in their spatial and temporal representation of landscapes. Field-based measures of habitat focus on a small number of habitat metrics that are feasible to measure on the ground, but which may not holistically capture habitat status. Remote sensing technology offers solutions for habitat monitoring that are spatially continuous and systematically collected through time (Nagendra et al., 2013). Such measurement approaches are more feasible from a monitoring perspective as spatio-temporal trends in data can be interrogated to understand natural system dynamics and identify deviations from expected patterns. However, the trade-off for most satellite sensors that provide wall-to-wall mapping at regular time intervals (e.g. Landsat and Sentinel missions) is that they acquire information of limited dimensionality and at pixel resolutions coarser than individual plant canopies. To maximise the utility of these sensors for monitoring habitat condition and dynamics, it is critical to calibrate the reflectance and backscatter signals with high-resolution data at the scale of individual plants.

## 1.1 Prior Research

The Beetaloo Sub-region has been a focus of environmental research since 2018. The major projects undertaken in the period 2018-2023 that are relevant to the current project are summarised below.

The Geological and Bioregional Assessment (GBA) of the Beetaloo Sub-basin produced a series of reports covering the natural and social landscape of the region. It also provided a full impact assessment and developed causal pathways that linked activities from an onshore gas industry with impacts to threatened species and various aspects of biodiversity (refer to [Beetaloo GBA region synthesis](#)).

The GBA found that most potential impact pathways are related to surface disturbance such as civil construction, decommissioning, and rehabilitation. The pathways of concern connect these activities with the protected matters, protected fauna and terrestrial vegetation endpoints, reflecting the potential impact of surface disturbance. The assessment determined that potential impacts could be minimised or mitigated by existing management controls (Huddleston-Holmes et al., 2020).

The causal network for the Beetaloo identified the points along a pathway where monitoring would be most useful. Results from the impact assessment informed four broad monitoring objectives: (i) estimating baseline and trend, (ii) comparing areas of potential impact with areas where no changes occur (control sites), (iii) monitoring compliance with, and effectiveness of, mitigation strategies, and (iv) monitoring to validate and refine the causal network.

The Strategic Regional Environmental and Baseline Assessment (SREBA) for the Beetaloo Sub-basin was undertaken following the recommendations of the Scientific Inquiry into Hydraulic Fracturing in the Northern Territory (2018). The SREBA addressed several key topics across the Beetaloo including water quality and quantity, aquatic ecosystems, terrestrial ecosystems, greenhouse gases, human health and social, and cultural and economic studies. Regional vegetation mapping and systematic flora (1,818 taxa) and fauna (354 vertebrate species) surveys have greatly increased knowledge of the terrestrial ecosystems in the Beetaloo region. The main terrestrial biodiversity values within the study area were stated to be: (a) Sites of Conservation Significance (Mataranka Thermal Pools and Lake Woods/Longreach Waterhole), (b) high value vegetation types, and (c) important habitat for waterbirds and threatened species (Department of Environment Parks and Water Security, 2022).

The GISERA program in the Northern Territory has funded several important projects, at least two of which are relevant to this project. The GISERA *Putting land management knowledge into practice project* explored the use of a high resolution 'digital twin' created from photogrammetry to evaluate the design and placement of gas infrastructure, protect surface water and vegetation, and reduce erosion, soil damage and dust (Huth et al., 2019). The GISERA *Beetaloo biodiversity project* (Pavey et al. 2025) assessed the impacts to biodiversity of roads and other infrastructure, and of habitat fragmentation in the Beetaloo Sub-region, focussing on the area east of Daly Waters where most development is likely to occur. The study examined vegetation fragmentation and connectivity using high-resolution (10 m) surfaces modelled using Sentinel-2 composite multi-spectral images from the dry season of 2021 (July-September). The study focussed on vegetation with a woody cover fraction of > 20 % and a height of > 2 m. Pixel-based methods for assessing fragmentation and connectivity were applied to four regions each of 50 km by 50 km. The study showed that indicators of fragmentation (FAD; Foreground Area Density) and connectivity (MSPA; Morphological Spatial Pattern Analysis), although designed for forest ecosystems, can also be applied to savanna and associated vegetation in northern Australia including the Beetaloo Sub-region.

The study also assessed patch mosaics in areas of bullwaddy (*Macropteranthes kekwickii*) and lancewood (*Acacia shirleyi*) - vegetation types that potentially support the highest concentrations of species 'at risk' from linear corridors. Floristic composition and vegetation structure differed according to patch size, and results suggest that the species-rich small bullwaddy stands may be persisting as low flammability islands in the landscape matrix of highly flammable savanna. In contrast, lancewood is currently absent from portions of the landscape where fire is frequent because, compared to bullwaddy, it is fire-prone, and it is very sensitive to fire interval due to its long juvenile period and low resistance to fire in adult trees. The small patches of bullwaddy may therefore be a sign that fire is presently too frequent in this landscape and lancewood stands are being lost. Overall, the study suggested that small patches dominated by bullwaddy should be of management and monitoring priority.

## 2 Research project goals

Petroleum interest holders are required to describe the existing environment that may be affected by proposed activities in their Environmental Management Plans (see [EMP Guidelines](#)). They must also have a monitoring program that allows for auditing of the environmental performance of the activity. Robust, repeatable, and easily deployable technologies that allow habitat condition to be reliably measured and monitored will be highly valuable in achieving these requirements.

This project aimed to extend a feasibility study conducted as part of the GBA program to use remote sensing technology to monitor the condition of habitat in the Beetaloo region. The scientific value spans both applied and empirical research, as there is growing momentum internationally and nationally for Earth Observation science to inform the monitoring of Essential Biodiversity Variables (EBVs) and the United Nation's System of Environmental Economic Accounting (SEEA) (Pereira et al., 2013; Vihervaara et al., 2017). However fundamental research into new sensor technologies, their sensitivities to different magnitudes of change, and their scalability, needs to be undertaken to provide confidence in these assessments (Daniel Kissling et al., 2024). Overcoming these gaps in knowledge is key to advancing routine operational monitoring of habitat condition over large and remote landscapes, and to verify effectiveness of mitigation methods.

Three-dimensional structure is a fundamental physical element of habitat and has long been identified as a key determinant of biological diversity, particularly in woodland ecosystems (MacArthur & MacArthur, 1961). Light-detection-and-ranging (LiDAR) can supply detailed information about the three-dimensional structure of vegetation elements and their spatial variability, and thus its use in habitat assessment has gained rapid traction since its emergence in ecology (Lefsky et al., 2002). Most studies utilising LiDAR for assessing habitat focus on deriving a spatial model of one or more elements of vegetation vertical or horizontal structure and then apply these models for predicting habitat suitability for a specific organism based on known habitat requirements (Hill & Hinsley, 2015). Very few studies have tried to explore the upscaling of LiDAR to Synthetic Aperture Radar (SAR) in the context of biodiversity monitoring, but there is evidence from temperate forests that this approach has merit (see Bae et al., 2019).

### 2.1 Study objective

The primary objective of this project was to develop a scalable approach for monitoring the structural condition of vegetation in the Beetaloo region. We aimed to demonstrate the application of high-resolution UAV-LiDAR to study site characterisation and test the potential for upscaling to larger areas via satellite-based remote sensing platforms.

Most remote sensing research in the field of habitat condition monitoring has focused on multi-spectral reflectance, however in dynamic tree-grass ecosystems and in landscapes with prominent background soil reflectance, spectral indicators of habitat structural state may saturate in the multi-spectral domain (Mutanga et al., 2023). Here, we aimed to test an active remote sensing to

this challenge by using Synthetic Aperture Radar (SAR) satellites to assess subtle variations in structural characteristics that are associated with habitat condition. Specifically, we aimed to:

1. Map the 3D structure of woody vegetation canopies with high-quality UAV-LiDAR surveying.
2. Test the sensitivity of Synthetic Aperture Radar (SAR) imagery to spatio-temporal variation in habitat structural properties, using the high-quality UAV-LiDAR data for calibration and validation.
3. Use the relationship between UAV-LiDAR structural metrics and spaceborne imagery to predict vegetation structure across broader unsampled areas.

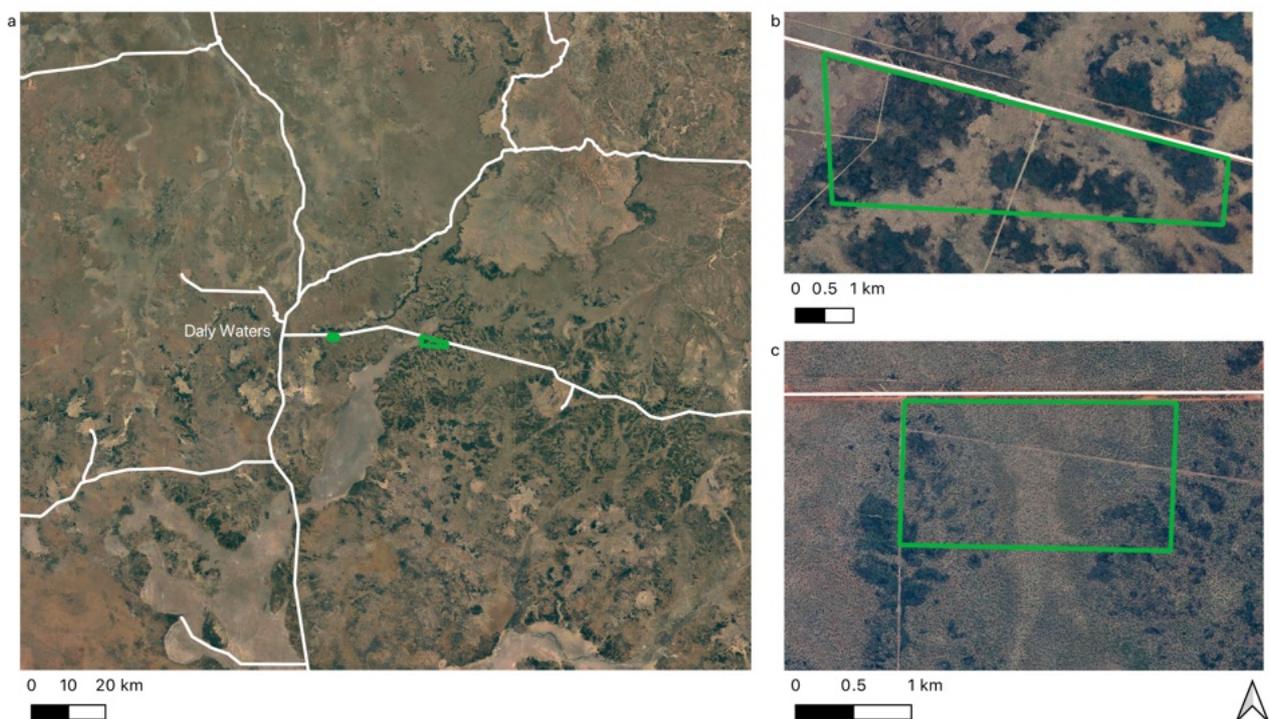


Figure 1 – UAV-LiDAR flights were conducted out of Daly Waters in Central Beetaloo region (a). Surveys areas (green boxes) were centred along the southern side of the Carpentaria highway (b and c) and aimed to capture diverse vegetation patch structures.

## 3 Research approach and methods

### 3.1 Study site description

The study region is part of the Sturt Plateau Bioregion, a predominantly flat erosional plain of shallow infertile soil and dominated by savanna woodlands of mixed eucalypts and bloodwoods with a perennial grass understorey, and open woodlands on clay soil floodplains (Huddleston-Holmes et al., 2020). In this part of the bioregion, forest, and woodland patches of lancewood and bullwaddy are an additional prominent landscape component. The lancewood canopy dominant, *Acacia shirleyi*, is a single-stemmed tree to 15 m that grows in shallow gravelly or skeletal sandy soils on sandstone or laterite, often forming dense stands, but also occurring in closed forests, low open forests, or mixed savanna woodlands. The bullwaddy dominant, *Macropteranthes kekwickii*, is a tree or shrub that can form dense impenetrable thickets and is also associated with lateritic soils. These canopy species often co-occur but their relative abundance is variable across their range. Woodland with dense bullwaddy cover supports high species richness relative to woodland/forest dominated by lancewood. Thick bullwaddy patches ('thickets') are further known to support a high richness of rainforest allied taxa, predominantly climbing vines, trees, and shrubs. Average annual rainfall is 675.9 mm. The majority of that rainfall (85% on average) falls during the four months from December to March.

### 3.2 UAV-LiDAR acquisition and processing

#### 3.2.1 Data acquisition

UAV-LiDAR data were acquired in the central region of the Beetaloo to capture a range of different vegetation structural classes (Figure 1). The UAV platform was an Acecore Noa hexacopter, fitted with a RIEGL survey-grade VUX-120 laser scanner (Figure 2), which has a ranging precision of 5 mm. The VUX-120 uses a forward-nadir-backward scan pattern, alternating between scanning at  $-10^\circ$  from the surface directly below the sensor (nadir), nadir and  $+10^\circ$  from the nadir, which increases the probability of receiving laser returns from vertical surfaces. The VUX-120 sensor was operated at a scanning rate of 1,200 kHz with a line speed of 291 lines per second.

Each UAV flight covered an area of 80-100 ha, flown at an altitude of 120 m above ground level using a north-south pattern (e.g., Figure 3). Surveys were conducted at a speed of  $8 \text{ ms}^{-1}$  and with a line spacing of 120 m, with total flight times ranging between 24 and 28 mins. A field base station, consisting of a Leica GS16 Global Navigation Satellite System (GNSS) receiver mounted on a survey tripod with a leveling tribrach, logged data over a 2-hour period per flight, spanning the flight duration.

The UAV-LiDAR datasets were collected in October 2023 (200 ha) and April 2024 (1000 ha).

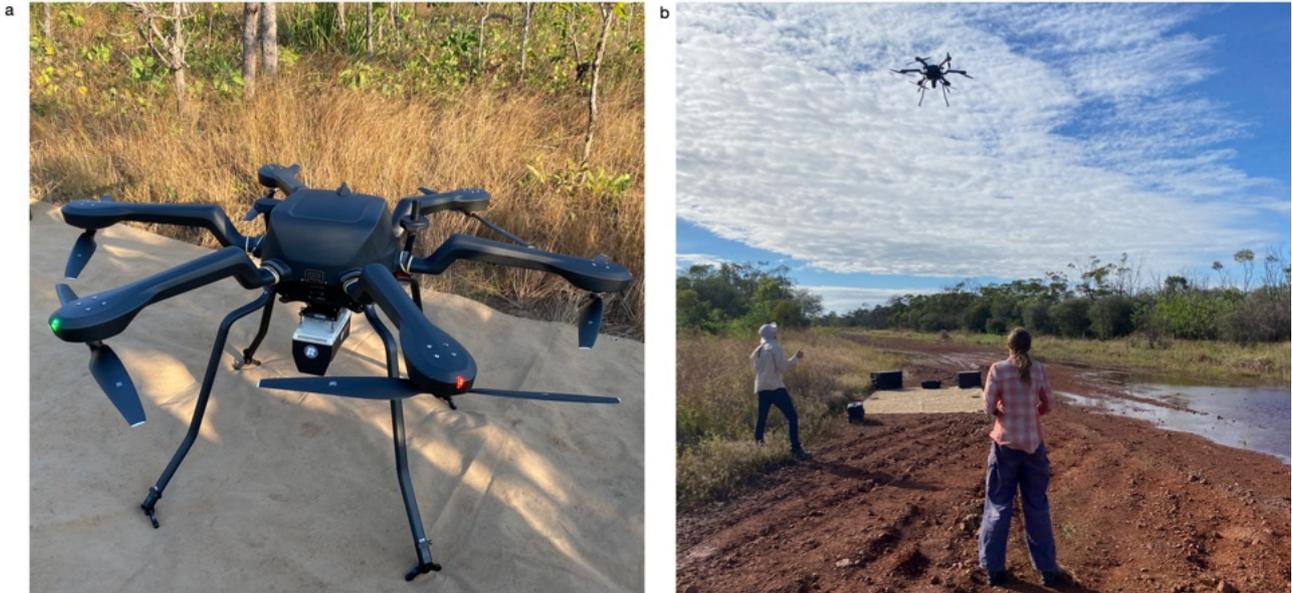


Figure 2 - Acecore Noa hexacopter with Riegl VUX-120 LiDAR sensor (a). Typical launch site during Beetaloo field campaigns (b).

### 3.2.2 UAV-LiDAR raw data processing

The GNSS logs collected by the base station were processed using Geoscience Australia's online GPS processing service, AUSPOS (version 2.4). AUSPOS uses up to 15 nearby reference stations from the International GNSS Service and the Asia-Pacific Reference Frame Network to perform double differencing. This method calculates the difference between signals received from two satellites at two different receivers, reducing the influence of clock and atmospheric errors and ensuring high coordinate precision.

The corrected geographic positions were then used to refine the raw flight path data collected by the APX-20 Inertial Measurement Unit (IMU) and GPS onboard the Acecore Noa. This refinement was performed using Trimble's Applanix POSPac UAV Suite 9.1 (dedicated manufacturer's software for processing and adjusting UAV trajectories).

The raw UAV-LiDAR flight lines were processed along with the corrected trajectory data using RIEGL's RiPROCESS software suite (v1.9.5). The flight lines were first clipped using the RXP cutter tool to remove lower quality data collected during flight turns or while transiting to the planned mission. Next, strip alignment was performed with RIEGL's RiPrecision tool to reduce systematic errors and biases in overlapping flight lines, which may arise from variations in GNSS signal quality and flight conditions. The refined flight lines and trajectory files were then exported in .laz and .txt format respectively, and their alignment was further fine-tuned using BayesStripAlign 2.24 (dedicated software using Bayesian inference to solve flightline misalignments).

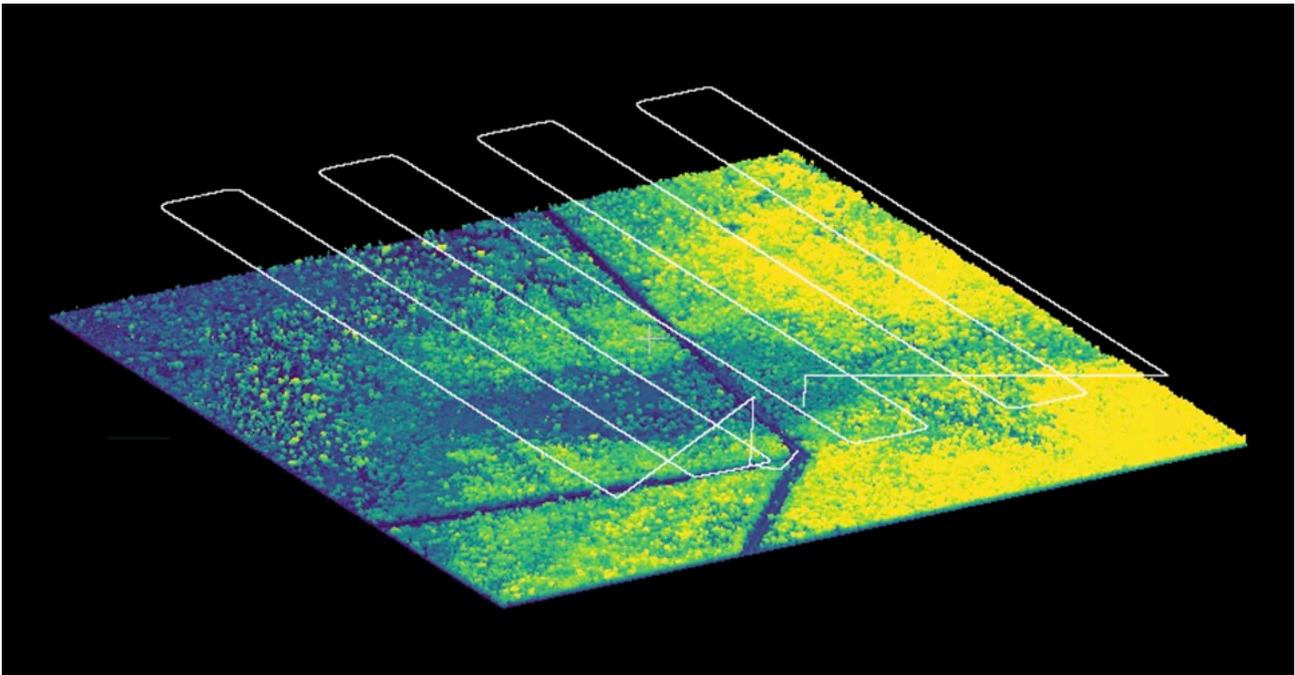


Figure 3 - Typical site survey with eight flightlines (white) of 1 km each, spaced 120 m apart, with the resulting point cloud displayed underneath. Warmer colours in the point cloud represent higher elevation above sea level.

### 3.2.3 UAV-LiDAR point cloud processing

The geolocated .laz flightlines were post-processed through batch scripts using LASTools (v2.0.3) routines to produce the analysis ready LiDAR datasets. The overlapping flightlines were cut into 500 m x 500 m tiles, with 10 m buffering to prevent edge artifacts, using *lastile*. For each UAV-LiDAR tile, ground returns were classified with *lasground*, and then normalised to height above ground level with the *lasheight* function. Any stray points that were either below ground level or higher than 40 m above ground level (erroneous data points that can arise from interactions with water on the ground surface, dust in the atmosphere, or birds in flight) were removed with *lasheight*. The cleaned and height normalised UAV-LiDAR .laz file tiles were used as the data inputs in the structural analyses described in Section 4.1. Examples of the high resolution digital terrain models (DTMs) and canopy height models (CHMs) derived from the UAV-LiDAR surveying are shown in Figure 4 and Figure 5 respectively.

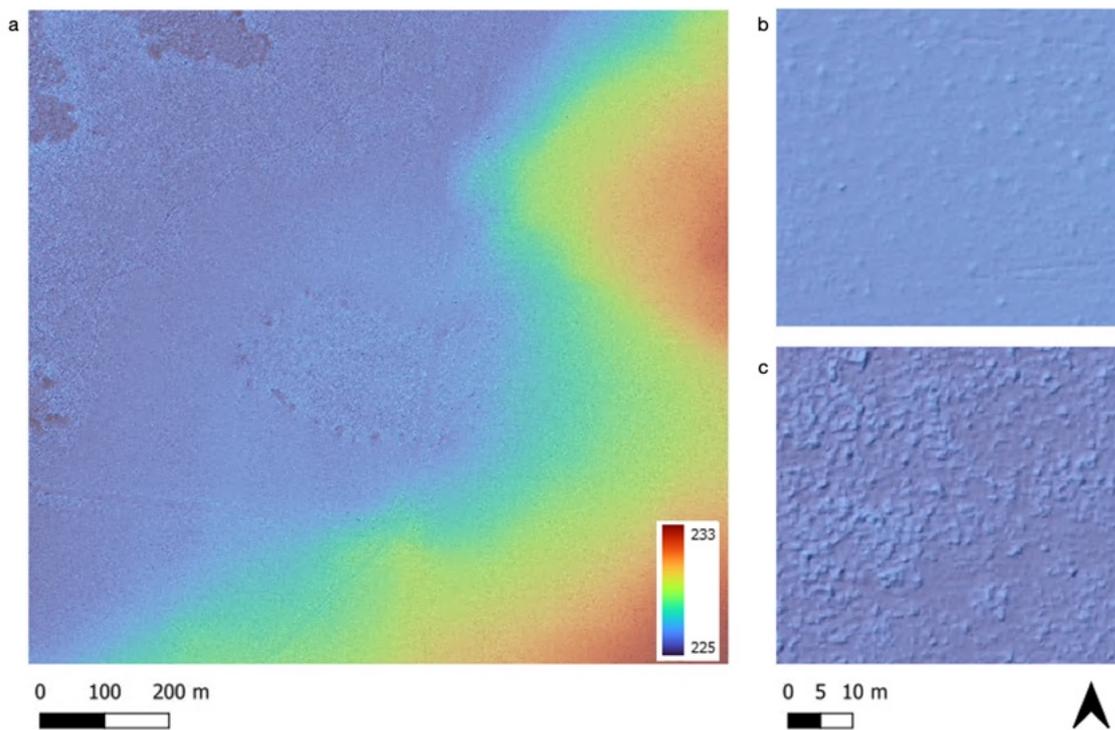


Figure 4 – Digital terrain model (DTM) derived from UAV-LiDAR. Note detection of lower lying drainage patches and road verges.

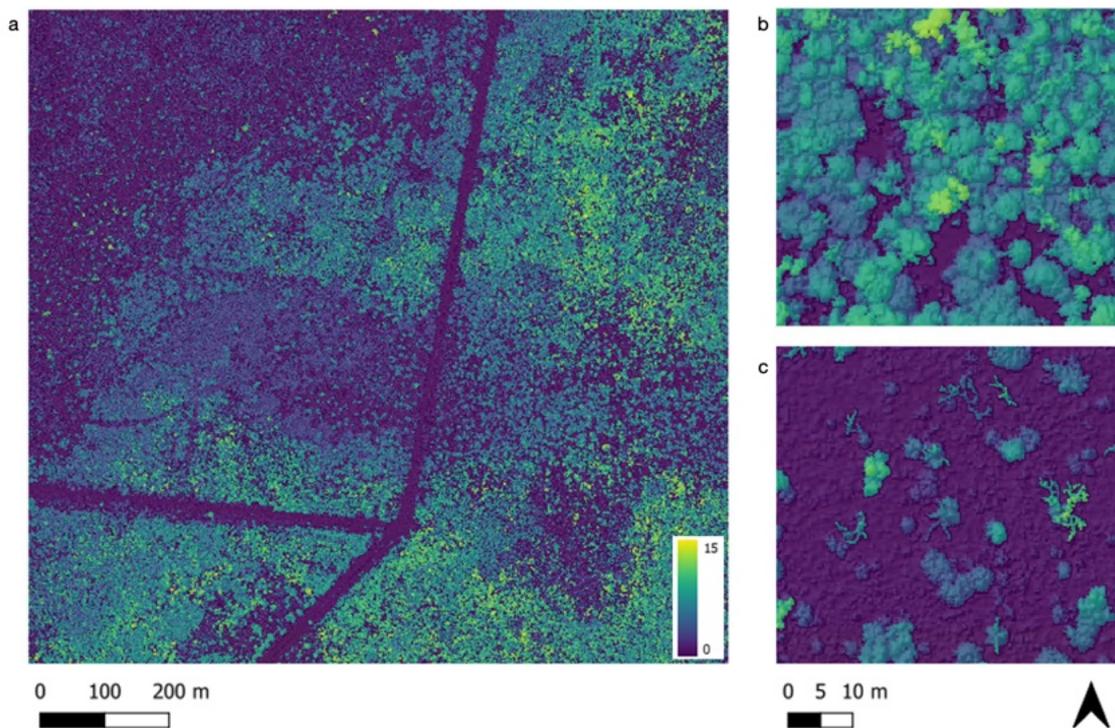


Figure 5 – Vegetation canopy height model (CHM) derived from UAV-LiDAR surveys, showing good delineation of roads and vegetation patches (a) and individual trees (b & c). note the detection of dead trees in (c).

## 3.3 Spaceborne Synthetic Aperture Radar (SAR) processing

### 3.3.1 Synthetic Aperture Radar (SAR)

Synthetic Aperture Radar (SAR) sensors are frequently used for Earth observation applications, either as a complement to optical sensors, or for some of their specific, unique capabilities. They are active sensors, emitting their own electromagnetic waves, with wavelengths in the centimetre scale, which is five orders of magnitude longer than the visible light (Jensen, 2014). Such active sensors have the advantage of not relying on another electromagnetic wave sources, enabling them to acquire data during the day and at night. The longer wavelength is also better at penetrating through clouds and smoke, so SAR images are less impacted by atmospheric conditions. Another advantage of SAR imagery is the retrieval of the radar phase information - recording the phase value along the sinusoidal signal when it hits the sensors after being reflected on Earth's surface. Comparing the phase information along a time series of radar images is referred to as Radar Interferometry, or InSAR (Massonnet & Feigl, 1998). This information allows detecting changes of elevation and structure of ground targets over which the signal bounces, with major applications in different science domains including hydrology, geology and ecology.

### 3.3.2 Sentinel-1 imagery

Sentinel-1 satellites, built and operated by the European Space Agency (ESA), are the first SAR systems designed for global monitoring of the Earth's surface (Torres et al., 2012). The Sentinel-1 SAR constellation comprises three satellites, Sentinel-1A, Sentinel-1B and Sentinel-1C (launched in 2014, 2016 and 2024, respectively). A malfunction in Sentinel-1B power systems in December 2021 led to its decommissioning shortly after, so current data is being collected by Sentinel-1A and Sentinel-1C. All Sentinel-1 data products are publicly available at no cost to the end-user, in accordance with open access data sharing policy from the European Commission. In Australia, distribution of Sentinel-1 data is managed through the Copernicus Australasia Regional Data Hub. The Hub is located and operated from the National Computational Infrastructure (NCI) in Canberra.

The Sentinel-1 constellation follows an orbital cycle of 12 days, meaning that repeat-pass images are theoretically acquired every 12 days over all land areas across the planet. Sentinel-1's nominal spatial resolution, in the Interferometric Wide (IW) swath mode typically used for land applications, is approximately 5 meters in the along-track direction (azimuth) and 20 meters in the across-track direction (range). The medium resolution, high temporal frequency and global coverage of Sentinel-1 data makes it particularly suited for large-scale applications. These include ground deformation monitoring for tectonics, volcanology (e.g., Albino et al., 2020) or groundwater (e.g., Parker et al., 2021). It also includes large scale monitoring of vegetation, soil moisture and erosion (e.g., Castellazzi et al., 2023). These applications typically use backscatter intensity images, representing the strength of the radar signal scattered back to the sensor by the features on the Earth's surface. Backscatter intensity is typically expressed in decibels (db). Some applications also leverage the phase information, which provides insight into on the changes of distance between the sensors and the ground features.

### **3.3.3 Radar backscatter processing**

Different types of backscatter intensity images can be produced from spaceborne radar sensors, according to the polarisation of the radar wave emitted and the polarisation of the receiving antenna. Being a dual-polarised sensor, two image bands are produced by Sentinel-1, one corresponding to a vertically emitted wave and a vertically received wave (VV), and one corresponding to a vertically emitted wave and a horizontally received wave (VH). VV signal is typically stronger than VH, as it only measures a portion of the original signal, which is spread angularly, typically by bouncing over rough or complex structures such as soils and vegetation canopy. The VH band, or an indication of its proportion in comparison to VV, are widely used in vegetation monitoring applications (e.g., Veloso et al., 2017).

Sentinel-1 Interferometric Wide (IW) images in the Ground Range Detected (GRD) format were used in this study. This format is a pre-processed, calibrated, ortho-corrected product. It is suitable for applications requiring both VV and VH intensity products. Postprocessing steps applied to the GRD imagery data are, in order: Calibration to Gamma Nought, Filtering using a Lee filter, Multi-looking, Resampling to a global 20-m resolution grid. The processing was performed using the Sentinel Application Platform (SNAP 11), developed by the ESA specifically for processing Sentinel-1 data. Temporal compositing of the VV and VH backscatter images by month and season was performed in CSIRO's EASI Hub platform. To explore other processing parameters, another processing chain based on the Sarscape software was tested (see section 8.1.1). Results from this processing chain are presented in the 'Future directions' (section 6.2).

## **3.4 Spaceborne multi-spectral image processing**

### **3.4.1 Background on Sentinel-2 multi-spectral data**

The Sentinel-2 satellite constellation, operated by the European Space Agency, has provided near-global (56°S to 83°N) publicly available multi-spectral imagery from mid-2015 to present (Drusch et al., 2012). These satellites acquire data for 13 spectral bands covering the visible, near-infrared and short-wave infrared wavelengths with a spatial resolution of between 10 and 60 m.

Historically, the constellation has been comprised of two satellites, Sentinel-2A and Sentinel-2B, that enable a revisit time of 5-days. As of January 2025, data updates from Sentinel-2A have been replaced by Sentinel-2C, the third Sentinel-2 satellite that was launched in September 2024. Sentinel-2D is scheduled to succeed Sentinel-2B in 2028, which will ensure a consistent and temporally contiguous record of multi-spectral imagery for earth observation applications (Toulemont et al., 2021).

### **3.4.2 Surface reflectance processing**

Storage and distribution of Sentinel-2 data in Australia is managed through the Copernicus Australasia Regional Data Hub. Our study utilised Sentinel-2A and Sentinel-2B surface reflectance (SR) images that were pre-processed by Geoscience Australia (2022a, 2022b). These data are considered 'analysis ready', as they have been corrected for atmospheric conditions, the

bidirectional reflectance distribution function (i.e., surface characteristics, solar radiation, and sensor view angle), and terrain illumination. All images intersecting the study region and available between June 2019 and May 2024 were downloaded from S3 cloud storage via the Spatio-Temporal Asset Catalog (STAC) maintained by Geoscience Australia.

Composite images (i.e., statistical summaries of surface reflectance over a fixed temporal window) were calculated using the geometric median at annual, seasonal and monthly frequencies for the purposes of modelling LiDAR structural metrics. While computationally expensive to calculate, the geometric median preserves the relationships between spectral bands and mitigates the effects of outliers and noise (see Roberts et al., 2017, 2018). The steps to calculate these composites included: (i) select spectral bands for modelling (see Table 1, the 60 m coastal aerosol, water vapour and cirrus bands were excluded), (ii) exclude images with greater than 80% cloud cover, (iii) resample all bands to a common resolution of 20 m, (iv) apply pixel-level cloud masking to the time-series, and (v) compute the geometric median at each pixel across the study region for the aggregation period of interest (i.e., annual, seasonal, monthly). Cloud masks were applied using the 's2cloudless' band provided with the Sentinel-2A and 2B imagery by calculating the maximum cloud probability within 100 m of each pixel and then masking all pixels where the cloud probability exceeded 0.5. All geometric median calculations were conducted in R, using 64 core / 512 GB RAM nodes on CSIRO's Petrichor High Performance Computing cluster.

Table 1 – Sentinel-2A and 2B multi-spectral bands used for modelling LiDAR structural metrics (Geoscience Australia, 2022a, 2022b).

Band	Description	Wavelength	Native / resampled resolution
nbart_blue	Blue	492.4 nm (2A), 492.1 nm (2B)	10 m / 20 m
nbart_green	Green	559.8 nm (2A), 559.0 nm (2B)	10 m / 20 m
nbart_red	Red	664.6 nm (2A), 665.0 nm (2B)	10 m / 20 m
nbart_red_edge_1	Red edge	704.1 nm (2A), 703.8 nm (2B)	20 m / 20 m
nbart_red_edge_2	Red edge	740.5 nm (2A), 739.1 nm (2B)	20 m / 20 m
nbart_red_edge_3	Red edge	782.8 nm (2A), 779.7 nm (2B)	20 m / 20 m
nbart_nir_1	Near infrared	832.8 nm (2A), 833.0 nm (2B)	10 m / 20 m
nbart_nir_2	Near infrared	864.7 nm (2A), 864.0 nm (2B)	20 m / 20 m
nbart_swir_2	Short-wave infrared	1613.7 nm (2A), 1610.4 nm (2B)	20 m / 20 m
nbart_swir_3	Short-wave infrared	2202.4 nm (2A), 2185.7 nm (2B)	20 m / 20 m

## 4 Statistical analyses and modelling

### 4.1 Structural metrics from LiDAR point-clouds

The capture of high-resolution and high-quality UAV-LiDAR point clouds provides the opportunity to summarise 3D vertical and horizontal structure at different spatial scales. Given the goal to test the upscaling of UAV-LiDAR to coarser and larger area satellite-based remote sensing, we generated a suite of structural metrics at 20 m resolution to match the SAR and SR data exactly on a per-pixel basis (Figure 6). For each pixel in the SAR and SR datasets, structural metrics were calculated in R using the lidR package (Roussel et al., 2020) from the height normalised UAV-LiDAR tiles. A total of 23 metrics were calculated (Table 2) with a spatial resolution of 20 m to align exactly on a per-pixel basis with the SAR and SR explanatory variables. A subset of these is shown for one site in Figure 7.

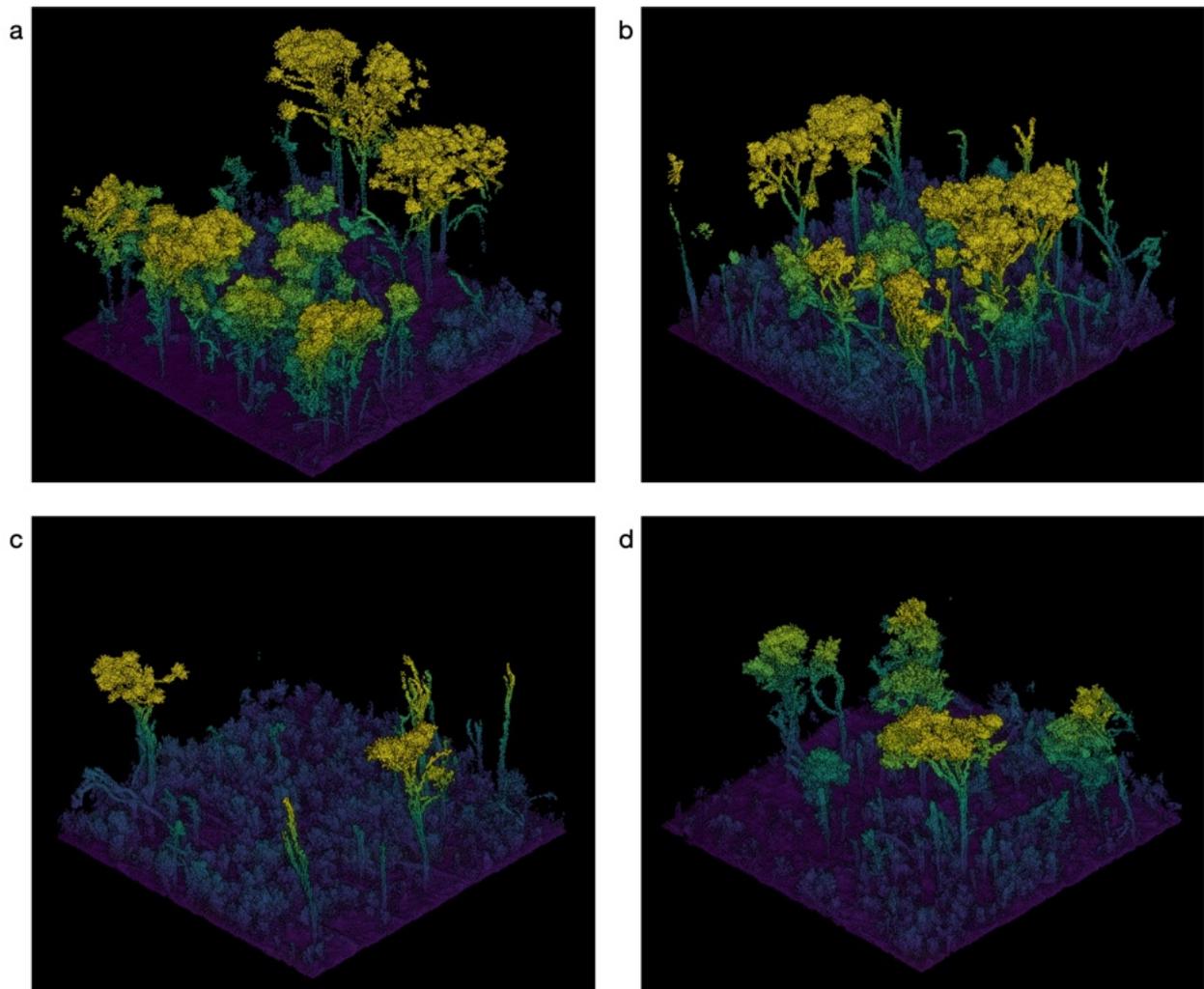


Figure 6 – Examples of the point cloud data contained within 20 m ‘pixels’, showing the range of structures that are integrated into each SAR and SR pixel.

Table 2 – LiDAR structural metrics evaluated for correlations with SAR and used as dependent variables for modelling vegetation structure. All metrics were calculated with a spatial resolution of 20 m.

Metric short name	Units	Description
GFP	Proportion	Gap Fraction Probability (GFP)
PAI	m <sup>2</sup> m <sup>-2</sup>	Plant Area Index (PAI)
stdev	m	Standard deviation of the height of all non-ground returns
p05	m	5 <sup>th</sup> percentile height of all non-ground returns
p10	m	10 <sup>th</sup> percentile height of all non-ground returns
p25	m	25 <sup>th</sup> percentile height of all non-ground returns
p50	m	50 <sup>th</sup> percentile height of all non-ground returns
p75	m	75 <sup>th</sup> percentile height of all non-ground returns
p90	m	90 <sup>th</sup> percentile height of all non-ground returns
p95	m	95 <sup>th</sup> percentile height of all non-ground returns
p98	m	98 <sup>th</sup> percentile height of all non-ground returns
dens_0_1m	Proportion	Number of non-ground returns from 0 to 1 m / total number of returns
dens_1_2m	Proportion	Number of non-ground returns from 1 to 2 m / total number of returns
dens_2_3m	Proportion	Number of non-ground returns from 2 to 3 m / total number of returns
dens_3_4m	Proportion	Number of non-ground returns from 3 to 4 m / total number of returns
dens_4_5m	Proportion	Number of non-ground returns from 4 to 5 m / total number of returns
dens_5_10m	Proportion	Number of non-ground returns from 5 to 10 m / total number of returns
dens_10_15m	Proportion	Number of non-ground returns from 10 to 15 m / total number of returns
dens_15_20m	Proportion	Number of non-ground returns from 15 to 20 m / total number of returns
dens_20_25m	Proportion	Number of non-ground returns from 20 to 25 m / total number of returns
dens_>1m	Proportion	Number of non-ground returns > 1 m / total number of returns
dens_>2m	Proportion	Number of non-ground returns > 2 m / total number of returns
dens_total	Proportion	Number of non-ground returns / total number of returns

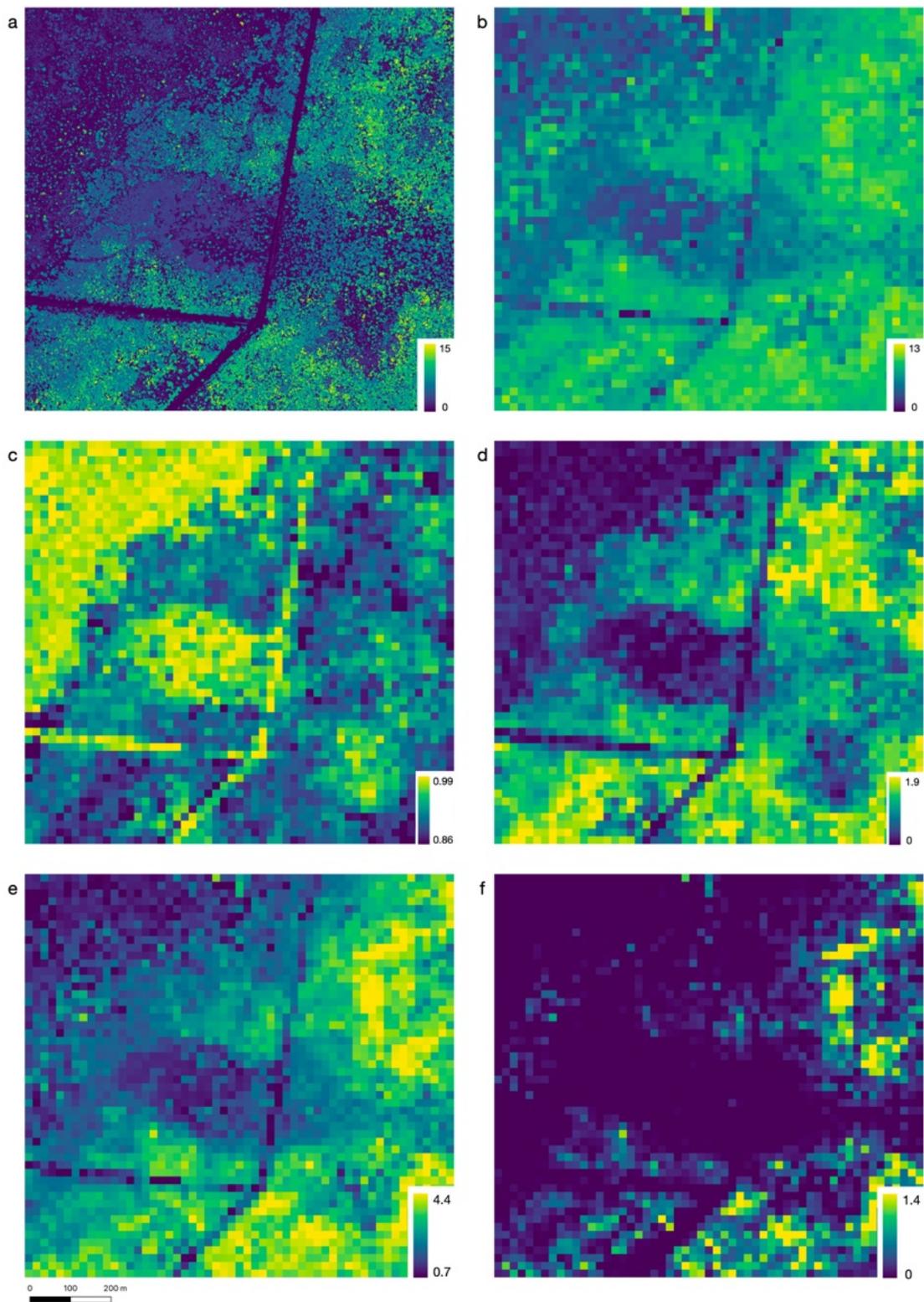


Figure 7 – Derivation of structural metrics from UAV-LiDAR at a single site. (a) Canopy height from the normalised point-cloud. Examples of metrics computed to match the 20 m pixels of the spaceborne imagery: (b) 98<sup>th</sup> percentile height, (c) GFP, (d) PAI, (e) std, and (f) density 5-10m.

## 4.2 Explanatory variables and spatial sampling

The SAR and SR datasets described in Section 3.3 and 3.4 were used as explanatory variables in our analysis. A regular spatial sample (i.e., where point samples are evenly distributed in a grid pattern) was taken by sampling every third pixel across the 20 m UAV-LiDAR structural metric grids, followed by a random jitter of up to 20 m in each horizontal axis. This sampling strategy provided near-complete spatial coverage of the UAV-LiDAR structural metric extents, while incorporating randomisation that would not lead to overlapping samples. These processes resulted in a total of 2,923 samples of the UAV-LiDAR structural metric values.

A series of explanatory variables representing annual ('Ann'), seasonal ('Seas'), late dry season (September to November, 'DrySeas') and conditions during the month of LiDAR acquisition ('MonthAqn'), in addition to inter-seasonal variability ('Var'), were calculated using the EASI SAR and Sentinel-2A and 2B SR imagery described in Sections 3.3 and 3.4. Seasons were defined as Austral winter from 1<sup>st</sup> June to 31<sup>st</sup> August (dry season), Austral spring from 1<sup>st</sup> September to 30<sup>th</sup> November (dry season), Austral summer from 1<sup>st</sup> December to 28<sup>th</sup>/29<sup>th</sup> February (wet season), and Austral autumn from 1<sup>st</sup> March to 31<sup>st</sup> May (wet season). Composite images for each aggregation period were summarised using the median of VV and VH backscatter intensity, and their ratio (VV:VH, calculated as  $VV - VH$  for dB which is on a log scale) for SAR, and geometric median of the 10 spectral bands described in Table 1 for SR. Inter-seasonal variability was calculated as the standard deviation of seasonal values for each variable. Each of the explanatory variables were extracted at all the sample locations for subsequent analyses.

## 4.3 Correlations between UAV-LiDAR structural metrics and satellite remote sensing

Correlations between SAR variables and UAV-LiDAR structural metrics were calculated for different compositing periods, including one year, each season, and the primary month of UAV-LiDAR acquisition using Spearman's rank-based correlation coefficient, which is robust to non-linear monotonic (i.e., as one variable increases or decreases, the other variable consistently moves in the same direction) relationships. The resulting correlations were assessed to explore the potential explanatory power of SAR for predicting structural metrics at different times of the year, when seasonal phenology or environmental conditions (e.g., soil moisture) could potentially impact upon the accuracy of predictive models. Pairwise correlations between UAV-LiDAR metrics were also calculated to quantify how each of the structural metrics relate to one another (see Appendix 2). All correlation analyses were conducted using the samples described in Section 4.2.

## 4.4 Predictive modelling with quantile regression forests

Predictive models of UAV-LiDAR structural attributes were fitted using quantile regression forests (Meinshausen, 2006), which use a series of regression trees (Breiman et al., 2017) to predict the conditional quantiles of the response variable. This enables quantification of the expected

response in addition to prediction intervals that can be used to quantify uncertainty. Furthermore, this algorithm is robust to overfitting (i.e., by fitting trees to random subsets of data and variables), can represent non-linear responses and variable interactions (i.e., by using an ensemble and allowing variable responses in different parts of the covariate domain), and has demonstrated strong predictive performance across many applications.

Quantile regression forests were fitted in R using the *ranger* package (Wright & Ziegler, 2017). The model training set was developed by selecting 70% of the sample data (see Section 4.2), and the remaining 30% of sample data were held out for model validation. The training set was further split into 10 cross-validation folds for optimising model parameters (i.e., *mtry*: the number of variables per split in the tree, and *min\_n*: the minimum number of samples for tree splitting to occur). These parameters were selected by running a grid-search for the optimal values, which involved iteratively running the 10-fold cross-validation for each model with 20 different random combinations of *mtry* (from 1 to the number of explanatory variables) and *min\_n* (from 1 to 30). Each model was fitted with 250 trees. The data used for training, validation, and cross-validation was partitioned such the original distribution of UAV-LiDAR density greater than 1 m ('dens\_>1m') was maintained. This ensured that the quantiles across sub-samples remained similar, thus mitigating the potential for this step to introduce bias.

UAV-LiDAR structural metrics were first modelled using the SAR explanatory variables. A total of 26,400 quantile regression forests were fitted to enable selection of the best performing approach (22 UAV-LiDAR structural metrics x 6 alternative sets of explanatory variables (i.e., 'Ann', 'AnnVar', 'Seas', 'SeasVar', 'DrySeas', 'MonthAqn'; see Section 4.2) x 10 cross-validation folds x 20 parameter combinations). The best performing model was selected by minimising the cross-validated root mean squared error (RMSE) for each UAV-LiDAR structural metric. The final models were fitted with all training data, and evaluated using the validation set (i.e., 30% of the sample data). The additional contribution provided by SR variables was then tested by repeating this procedure using the same set of explanatory variables identified as optimal by the SAR modelling results (e.g., annual, seasonal or monthly composites, with or without variability). An additional 4,400 models were fitted when incorporating SR (22 UAV-LiDAR structural metrics x 10 cross-validation folds x 20 parameter combinations).

# 5 Study results

## 5.1 Correlations between LiDAR structural metrics and SAR

The correlation between UAV-LiDAR structural metrics and SAR explanatory variables for annual, seasonal, and the primary month of acquisition temporal aggregation periods are reported in Figure 9. When statistically significant relationships were detected, all structural metrics, except for the Gap Fraction Probability ('GFP') and density below 1 m ('dens\_<1m'), were positively correlated with VV and VH backscatter intensity and negatively correlated with their ratio (VV:VH). This trend was consistent across all aggregation periods. UAV-LiDAR structural metrics were also typically more strongly correlated with VH backscatter intensity, which is known to be sensitive to volumetric scattering in the plant canopy, than with VV.

From a temporal perspective, the strongest correlations between UAV-LiDAR structural metrics and SAR were found late in the dry season (Figure 9c, Figure 9b), when there is less water present in the landscape that would otherwise introduce noise by increasing backscatter intensity.

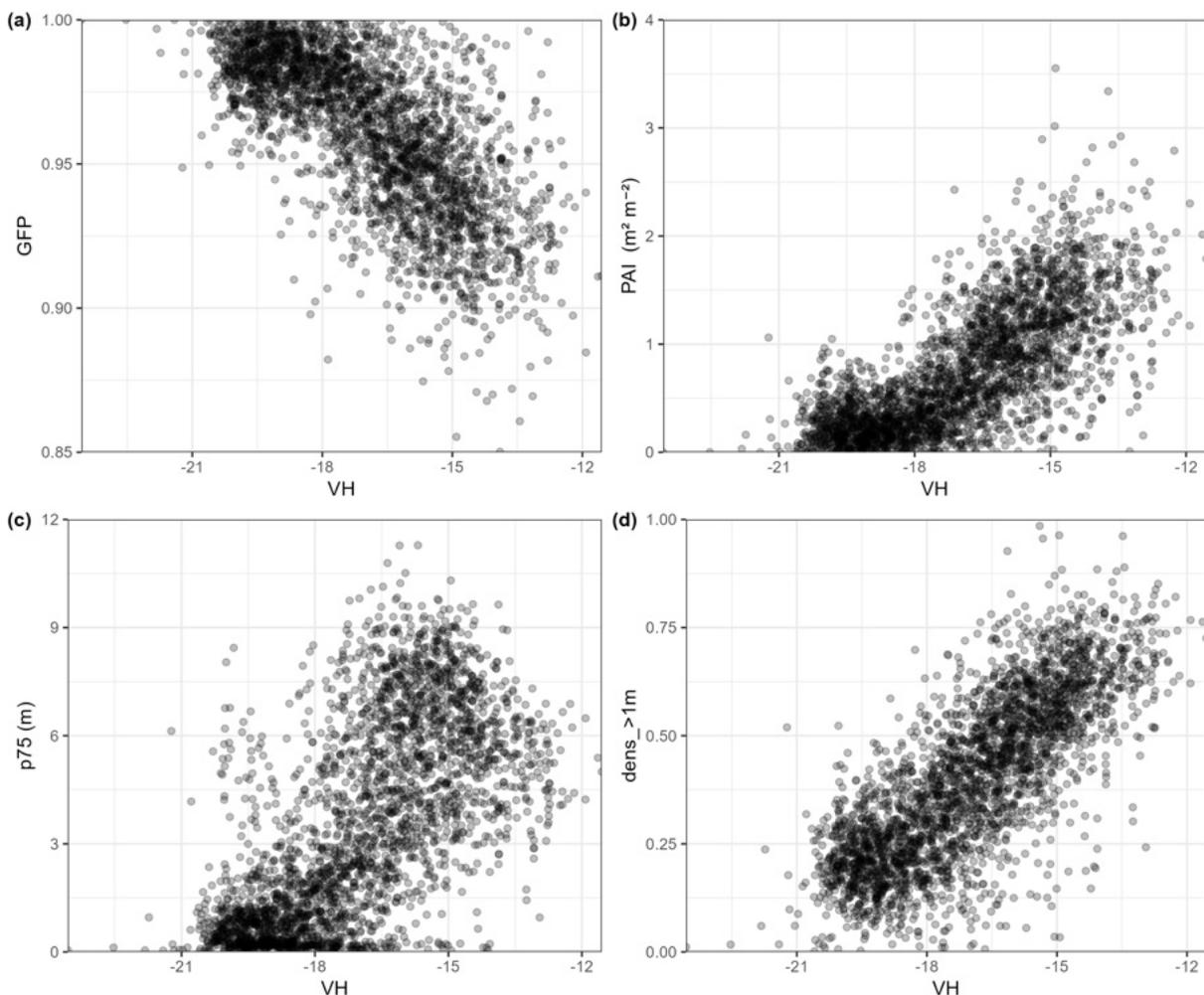


Figure 8 – Scatterplots of the relationship between VH backscatter and (a) Gap Fraction Probability (GFP), (b) Plant Area Index (PAI), (c) 75<sup>th</sup> percentile height (p75) and (d) density above 1 m height (n = 2,923).



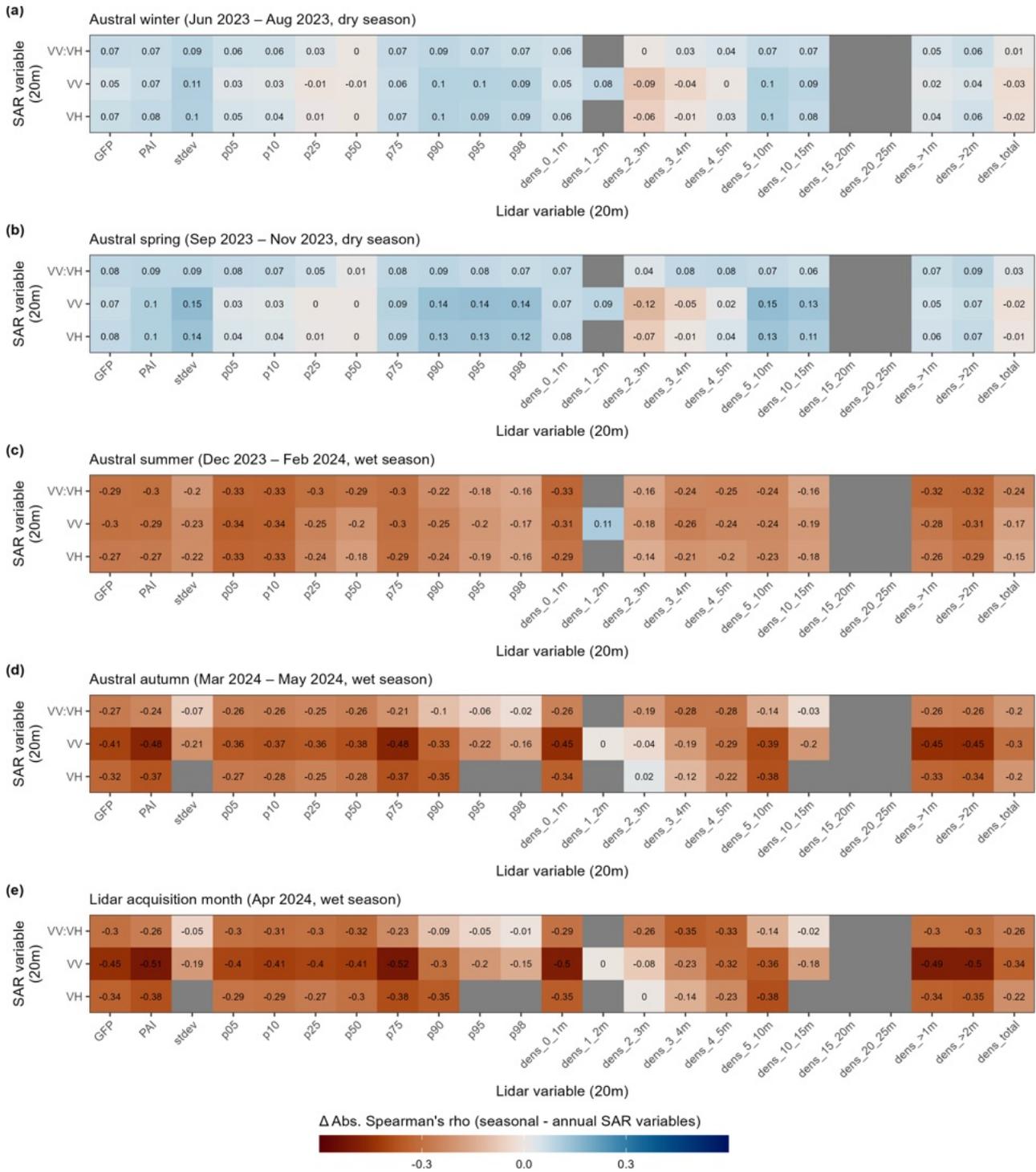


Figure 10 - Absolute difference in pairwise correlations between SAR (y-axes) and LiDAR (x-axes) variables for (a) austral winter, (b) austral spring, (c) austral summer, (d) austral autumn, and (e) month of LiDAR acquisition in comparison to annual median compositing periods (i.e., blue values indicate better correlations with lidar variables than the annual composite). All comparisons were conducted on a randomised spatial sample of point locations ( $n = 2,923$ ) using Spearman's rank-based correlation coefficient ( $\rho$ ). Non-significant ( $p \geq 0.05$ ) correlations are not reported (grey cells).

## 5.2 Predictive modelling of UAV-LiDAR structural metrics

### 5.2.1 Predictive modelling with SAR variables

Cross-validated statistical performance for quantile regression forest models of UAV-LiDAR structural metrics fitted with SAR explanatory variables (see Section 4.2) are presented in Figure 11. Note that high values of Lin's Concordance Correlation Coefficient, and low values of RMSE / standard deviation indicate stronger performance. The Plant Area Index (PAI), Gap Fraction Probability (GFP), 75<sup>th</sup> percentile height (p75), and density above 1 or 2 m height (dens\_>1m, dens\_>2m) were the best performing structural metrics. Explanatory variables that quantified how conditions vary throughout the year provided the best predictions of UAV-LiDAR structural metrics. These included 'SeasVar' (i.e., median backscatter for all seasons and standard deviation across seasons), 'Seas' (i.e., median backscatter for all seasons), and 'AnnVar' (i.e., median backscatter for a whole year, and standard deviation across seasons). Despite dry season imagery typically performing better than the median backscatter intensity over a full year, the inclusion of variability and data for additional seasons further improved statistical performance.

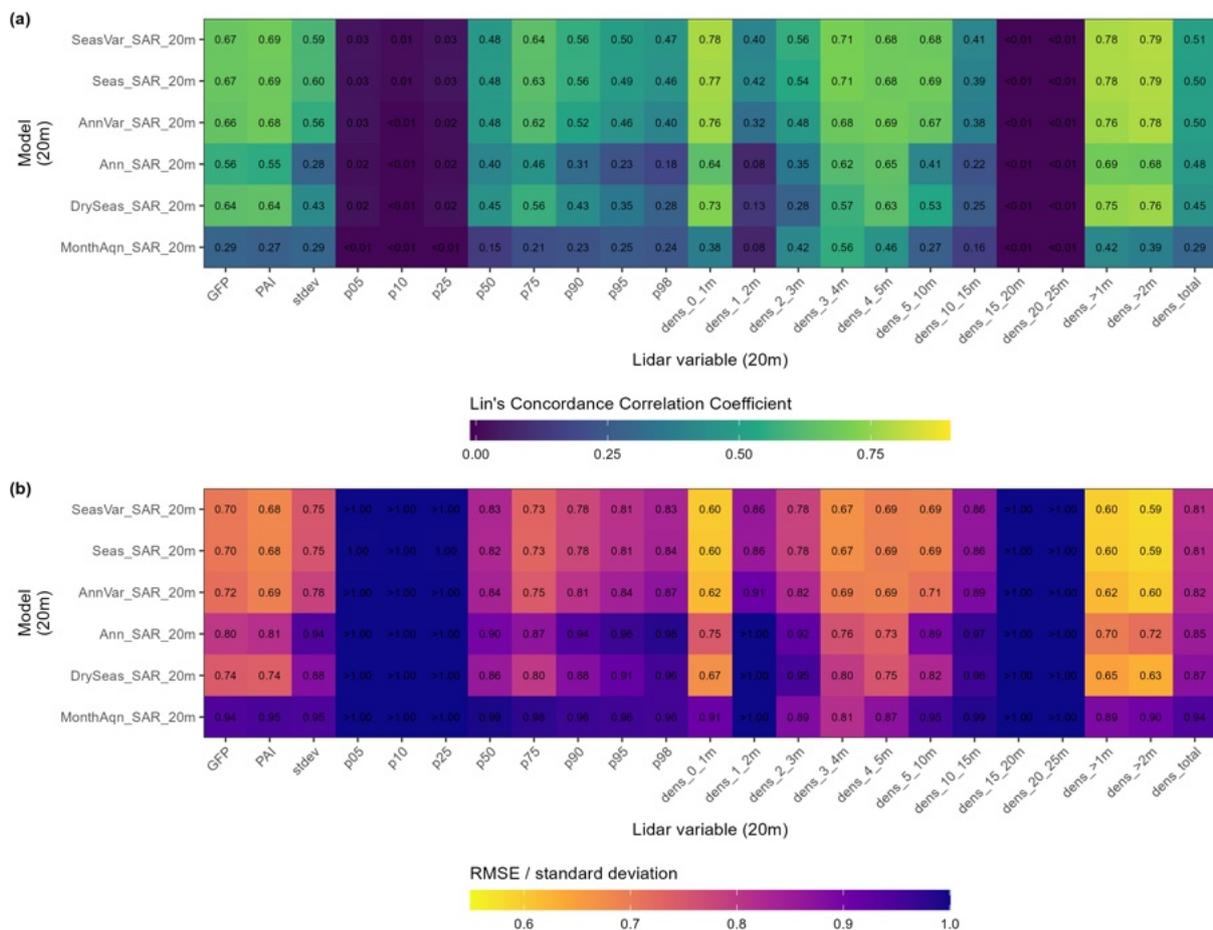


Figure 11 – Cross validated (a) Lin's Concordance Correlation Coefficient and (b) root mean squared error (RMSE) / mean for predictive models (quantile regression forests) of LiDAR structural metrics (x-axes) using SAR variables (y-axes). Statistics were calculated using 10-fold cross-validation on 70% of samples selected for model training (n = 2,043).

The statistical performance of the validation set was typically consistent with the cross-validation analyses (Table 3), indicating that the models were not overfitted and predictions remained stable when applied to new data. The ‘Seas’ and ‘SeasVar’ variables remained most performant in nearly all cases, except for the 10<sup>th</sup> percentile height (p10) which performed poorly and was unlikely to produce reliable predictions. Height models performed best at the 75<sup>th</sup> percentile, with diminishing performance when predicting higher or lower parts of the distribution. This result is potentially due to: (i) the canopy penetration capacity of the C-band, and (ii) the strong correlation between density and height that can be observed both in the UAV-LiDAR structural metrics and correlation tests on our fitted models (Figure 20, Figure 21c). Similarly, density metrics were best predicted when modelling an aggregated representation of the vertical profile (‘dens\_>1m’, ‘dens\_>2m’). Scatterplots of observed and validation predictions are illustrated for select structural metrics in Figure 12.

Table 3 – Validation statistics for held-out point locations (n = 880) and optimised parameters (mtry, min\_n) for predictive models of LiDAR structural metrics using SAR variables. Specific statistics reported are the coefficient of determination (R<sup>2</sup>), Lin’s Concordance Correlation Coefficient (CCC), root mean squared error (RMSE), bias (mean error) and the mean value of the observations (Mean).

Lidar variable (units)	Model type	Variables per split (mtry)	Minimum node size (min_n)	R <sup>2</sup>	CCC	RMSE	Bias	Mean
GFP	SeasVar_SAR_20m	2	3	0.52	0.68	0.02	0.00	0.96
PAI (m <sup>2</sup> m <sup>-2</sup> )	SeasVar_SAR_20m	1	11	0.65	0.78	0.35	0.00	0.74
stdev (m)	Seas_SAR_20m	2	9	0.53	0.68	0.73	0.01	2.64
p05 (m)	Seas_SAR_20m	1	6	0.22	0.46	0.07	0.01	-0.09
p10 (m)	Ann_SAR_20m	1	6	0.06	0.22	0.09	0.02	-0.06
p25 (m)	SeasVar_SAR_20m	1	5	0.09	0.24	0.35	0.03	0.06
p50 (m)	SeasVar_SAR_20m	1	7	0.54	0.70	1.05	0.06	1.08
p75 (m)	Seas_SAR_20m	2	13	0.63	0.77	1.70	0.02	3.59
p90 (m)	Seas_SAR_20m	2	8	0.52	0.69	2.06	0.02	6.09
p95 (m)	Seas_SAR_20m	1	8	0.48	0.63	2.08	0.07	7.34
p98 (m)	SeasVar_SAR_20m	3	14	0.45	0.61	2.04	0.10	8.45
dens_0_1m	SeasVar_SAR_20m	2	15	0.63	0.77	0.08	0.00	0.33
dens_1_2m	Seas_SAR_20m	4	11	0.22	0.36	0.04	0.00	0.06
dens_2_3m	SeasVar_SAR_20m	6	3	0.42	0.57	0.04	0.00	0.05
dens_3_4m	Seas_SAR_20m	4	15	0.53	0.69	0.03	0.00	0.05
dens_4_5m	SeasVar_SAR_20m	2	11	0.52	0.69	0.03	0.00	0.04
dens_5_10m	Seas_SAR_20m	3	3	0.60	0.73	0.08	0.00	0.17
dens_10_15m	Seas_SAR_20m	1	8	0.26	0.40	0.04	0.00	0.03
dens_15_20m	SeasVar_SAR_20m	1	13	0.00	0.00	0.00	0.00	0.00
dens_20_25m	SeasVar_SAR_20m	1	15	0.00	0.00	0.00	0.00	0.00
dens_>1m	SeasVar_SAR_20m	2	9	0.66	0.79	0.11	0.00	0.40
dens_>2m	Seas_SAR_20m	2	14	0.68	0.81	0.11	0.00	0.34
dens_total	Seas_SAR_20m	1	14	0.37	0.53	0.07	0.00	0.73

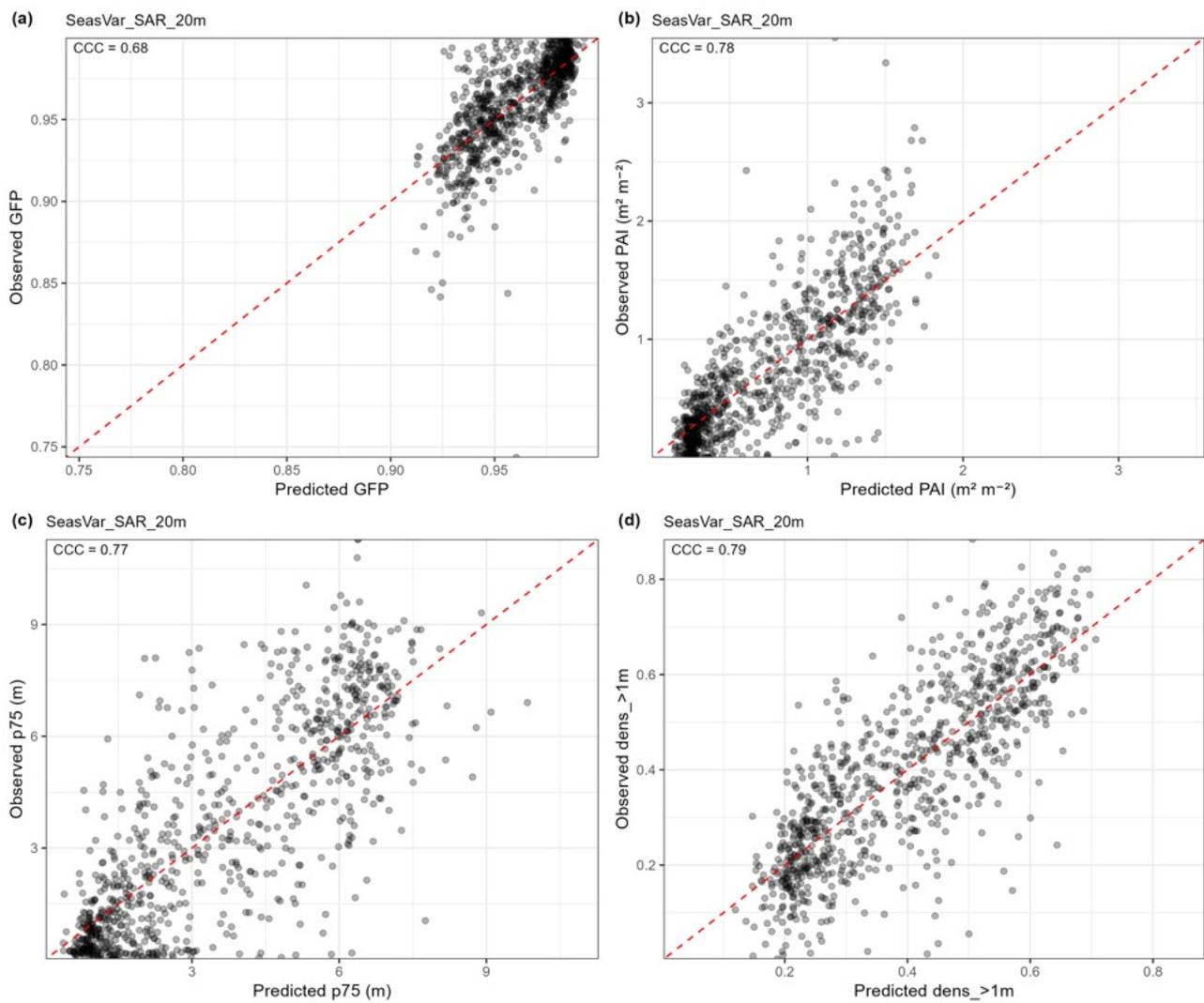


Figure 12 – Observed versus predicted (a) Gap Fraction Probability (GFP), (b) Plant Area Index (PAI), (c) 75<sup>th</sup> percentile height (p75), and (d) density above 1 m height (dens\_>1m) at held-out point locations ( $n = 880$ ) when modelled using SAR variables. The red dashed line shows the 1:1 relationship between observed and predicted values. Model names are specified above each plot. Lin's Concordance Correlation Coefficient (CCC) is reported in the top left corner of each subplot.

## 5.2.2 Predictive modelling with both SAR and SR variables

Statistical performance of the LiDAR structural metrics fitted with both SAR and surface reflectance (SR) explanatory variables are presented in Table 4. The relative performance of these models was consistent with the SAR only approach for each of the structural metrics. While these models included many more explanatory variables (see Table 1) with the addition of SR, they increased predictive performance by up to 0.2 for Lin's CCC relative to the SAR only models. This improved accuracy is evident in the scatterplots of observed and validation predictions for select structural metrics illustrated in Figure 13.

Table 4 – Validation statistics for held-out point locations (n = 880) and optimised parameters (mtry, min\_n) for predictive models of LiDAR structural metrics using both SAR and SR variables. Specific statistics reported are the coefficient of determination (R<sup>2</sup>), Lin's Concordance Correlation Coefficient (CCC), root mean squared error (RMSE), bias (mean error) and the mean value of the observations (Mean).

Lidar variable (units)	Model type	Variables per split (mtry)	Minimum node size (min_n)	R <sup>2</sup>	CCC	RMSE	Bias	Mean
GFP	Seas_SARSR_20m	22	8	0.62	0.76	0.02	0.00	0.96
PAI (m <sup>2</sup> m <sup>-2</sup> )	SeasVar_SARSR_20m	11	11	0.75	0.86	0.29	0.01	0.74
stdev (m)	SeasVar_SARSR_20m	15	6	0.69	0.81	0.59	0.03	2.64
p05 (m)	Seas_SARSR_20m	1	10	0.26	0.49	0.07	0.02	-0.09
p10 (m)	SeasVar_SARSR_20m	1	10	0.13	0.26	0.11	0.04	-0.06
p25 (m)	Seas_SARSR_20m	1	6	0.15	0.33	0.34	0.04	0.06
p50 (m)	SeasVar_SARSR_20m	3	13	0.67	0.81	0.89	0.10	1.08
p75 (m)	SeasVar_SARSR_20m	9	8	0.77	0.87	1.36	0.09	3.59
p90 (m)	SeasVar_SARSR_20m	8	8	0.67	0.81	1.70	0.07	6.09
p95 (m)	SeasVar_SARSR_20m	5	5	0.62	0.76	1.76	0.13	7.34
p98 (m)	Seas_SARSR_20m	7	10	0.57	0.72	1.79	0.16	8.45
dens_0_1m	SeasVar_SARSR_20m	18	8	0.74	0.85	0.07	0.00	0.33
dens_1_2m	Seas_SARSR_20m	18	4	0.43	0.56	0.03	0.00	0.06
dens_2_3m	SeasVar_SARSR_20m	24	5	0.62	0.72	0.03	0.00	0.05
dens_3_4m	SeasVar_SARSR_20m	13	9	0.65	0.77	0.03	0.00	0.05
dens_4_5m	SeasVar_SARSR_20m	7	4	0.63	0.76	0.02	0.00	0.04
dens_5_10m	Seas_SARSR_20m	11	5	0.71	0.82	0.07	0.00	0.17
dens_10_15m	SeasVar_SARSR_20m	11	2	0.42	0.59	0.03	0.00	0.03
dens_15_20m	SeasVar_SARSR_20m	1	7	0.00	-0.01	0.00	0.00	0.00
dens_20_25m	SeasVar_SARSR_20m	1	10	0.00	0.00	0.00	0.00	0.00
dens_>1m	Seas_SARSR_20m	19	12	0.80	0.89	0.08	0.00	0.40
dens_>2m	Seas_SARSR_20m	18	8	0.81	0.89	0.09	0.00	0.34
dens_total	Seas_SARSR_20m	12	14	0.52	0.68	0.06	0.00	0.73

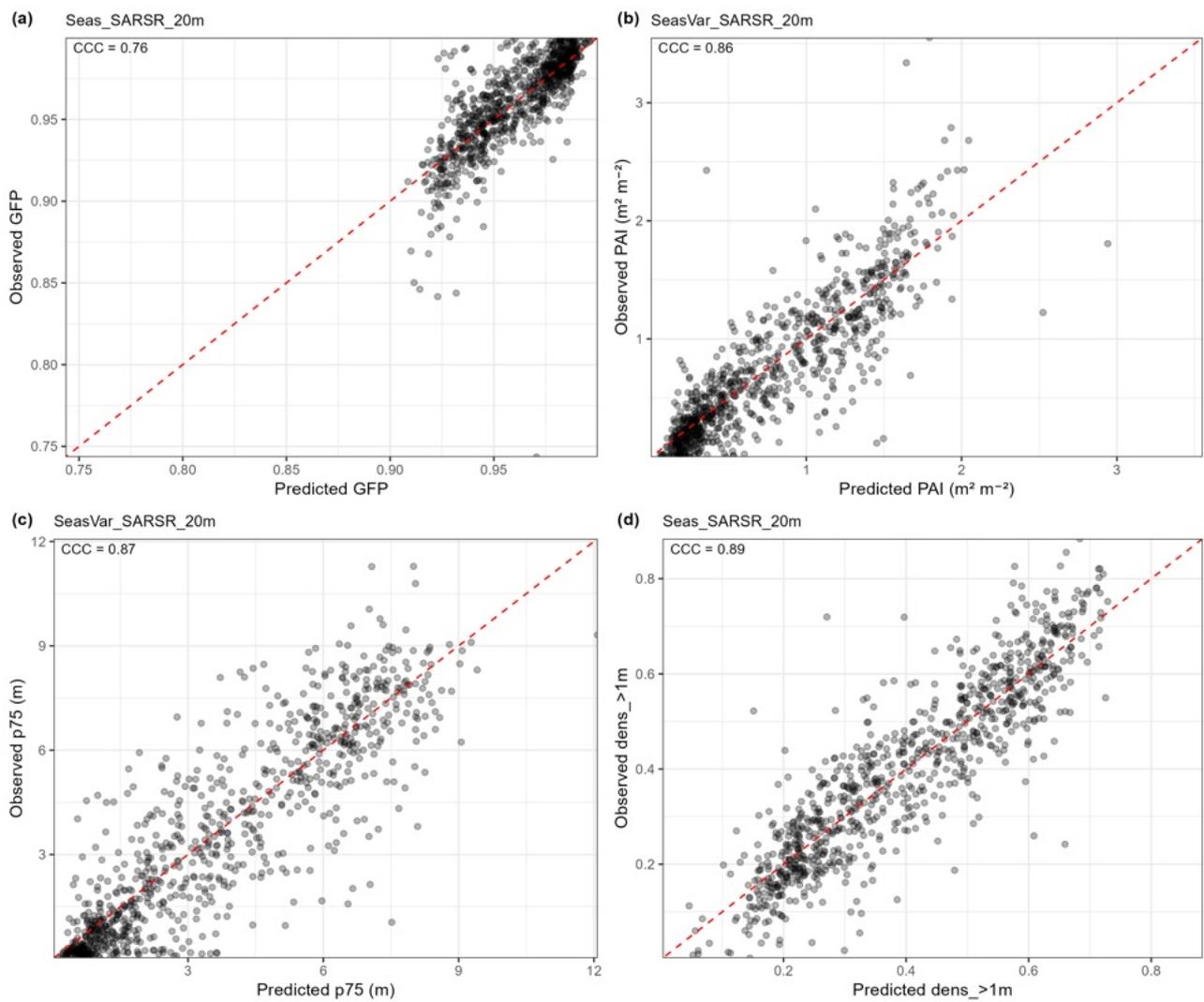


Figure 13 – Observed versus predicted (a) Gap Fraction Probability (GFP), (b) Plant Area Index (PAI), (c) 75<sup>th</sup> percentile height (p75), and (d) density above 1 m height (dens\_>1m) at held-out point locations ( $n = 880$ ) when modelled using both SAR and SR variables. The red dashed line shows the 1:1 relationship between observed and predicted values. Model names are specified above each plot. Lin's Concordance Correlation Coefficient (CCC) is reported in the top left corner of each subplot.

### 5.3 Extrapolating predictions of vegetation 3D structure over a larger extent

Spatial predictions of the Gap Fraction Probability (GFP), Plant Area Index (PAI) and density above 1 m ( $\text{dens}_{>1\text{m}}$ ) are illustrated in Figure 14, Figure 15, Figure 16, alongside the prediction intervals given by the quantile regression forests and comparisons against the SAR only alternatives. The prediction intervals (i.e., Figure 14b, Figure 15b, Figure 16b) can be used to assess areas with the highest predictive uncertainty, and can be used to help guide the collection of additional training data by targeting specific features where the model is uncertain. These prediction intervals may also be compared against alternative models to explore the range of uncertainty remaining when applying different modelling approaches. The addition of SR as explanatory variables considerably decreased the range of uncertainty present in each model (Figure 14d, Figure 15d, Figure 16d). This is likely due to the additional information provided by the multi-spectral imagery that improves the distinction between dense and more sparsely vegetated parts of the landscape, as is visible in the differences shown in Figure 14c, Figure 15c, and Figure 16c.

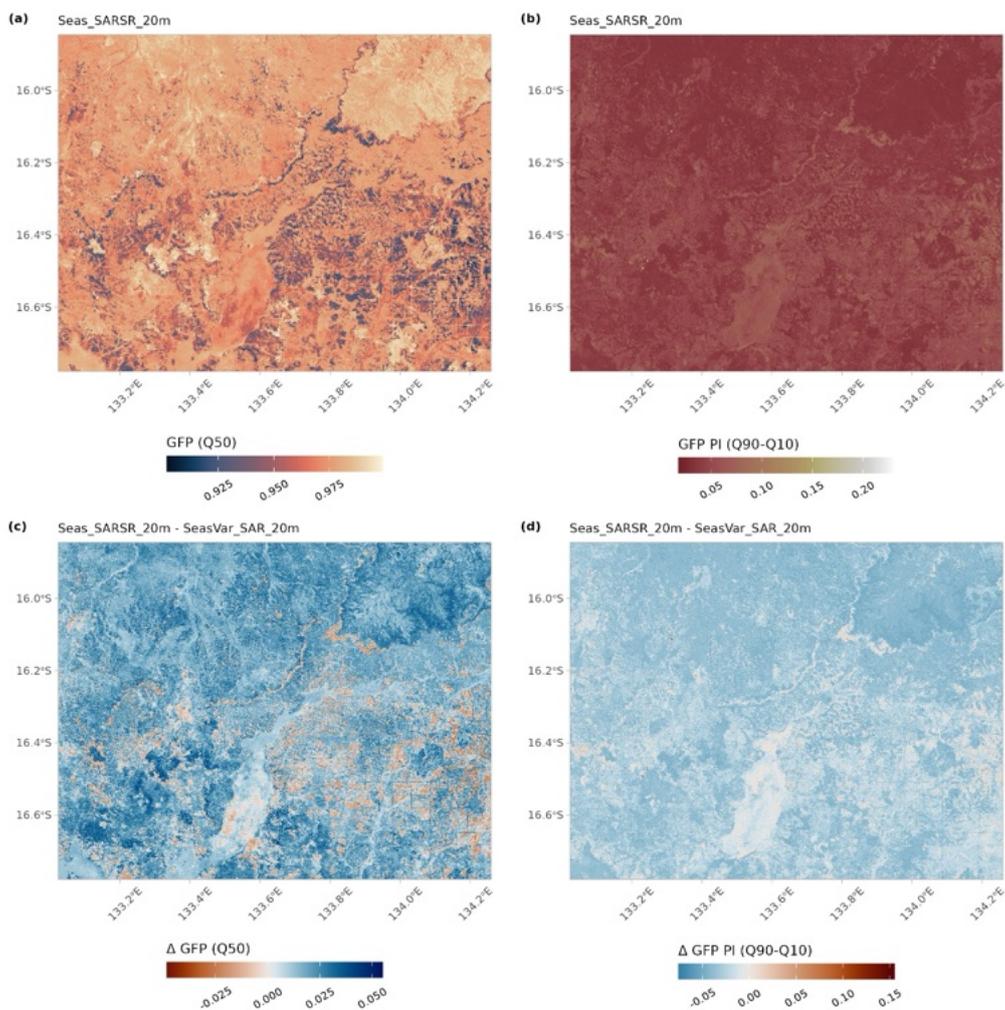


Figure 14 – Modelled (a) Gap Fraction Probability (GFP), (b) width of the GFP prediction interval (PI), (c) difference in modelled GFP when using surface reflectance (SR) in addition to SAR for prediction, and (d) difference in the prediction interval width when using both SR and SAR for prediction (blue indicates reduced uncertainty). Q10, Q50 and Q90 represent the 10<sup>th</sup>, 50<sup>th</sup> and 90<sup>th</sup> percentile of predictions, respectively. Model names and their order for the purposes of calculating differences are specified above each plot.

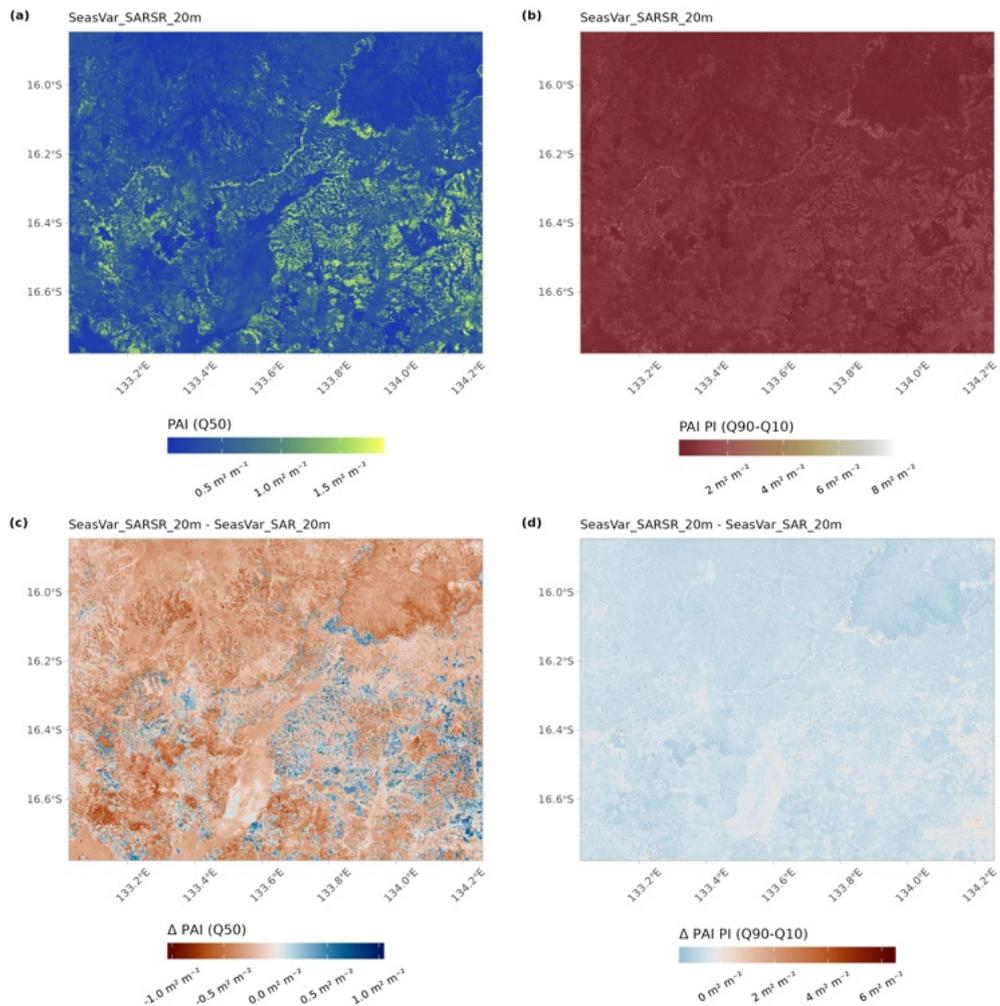


Figure 15 – Modelled (a) Plant Area Index (PAI), (b) width of the PAI prediction interval (PI), (c) difference in modelled PAI when using surface reflectance (SR) in addition to SAR for prediction, and (d) difference in the prediction interval width when using both SR and SAR for prediction (blue indicates reduced uncertainty). Q10, Q50 and Q90 represent the 10<sup>th</sup>, 50<sup>th</sup> and 90<sup>th</sup> percentile of predictions, respectively. Model names and their order for the purposes of calculating differences are specified above each plot.

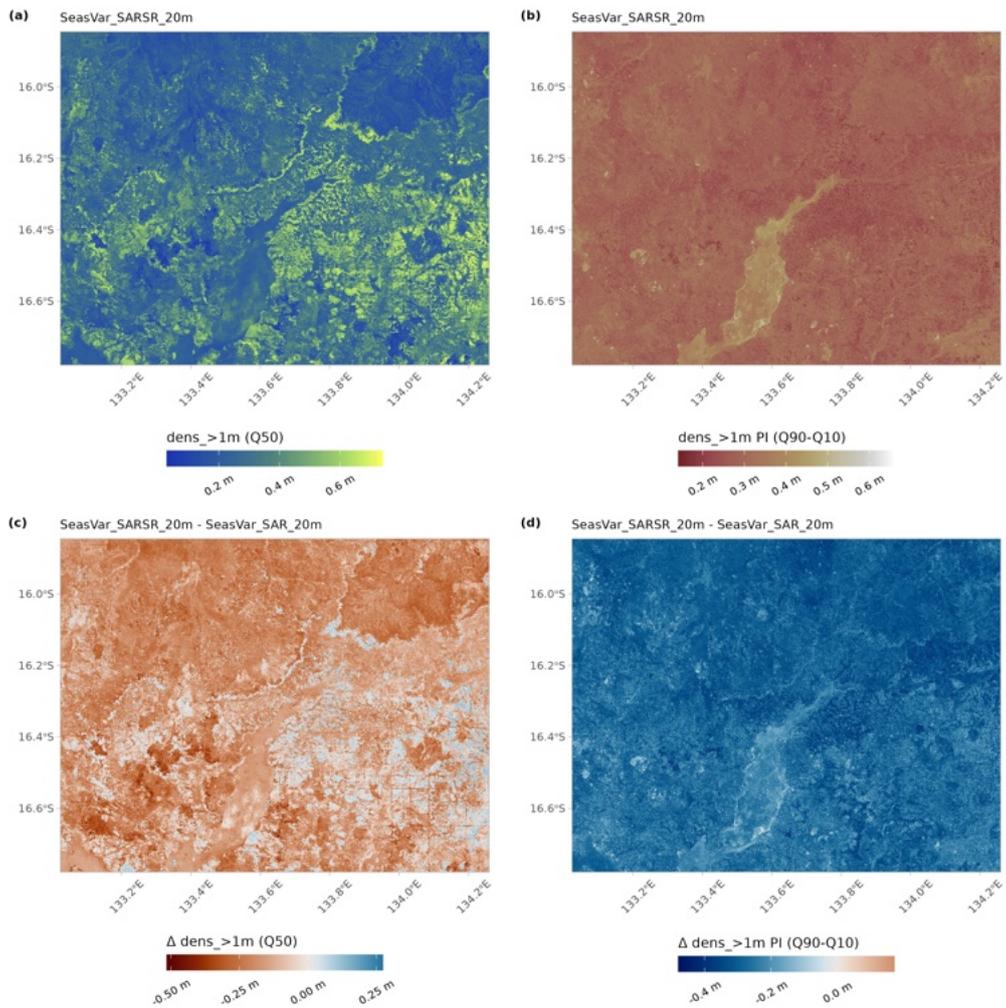


Figure 16 - Modelled (a) return density above 1 m ( $\text{dens}_{>1\text{m}}$ ), (b) width of the prediction interval (PI) for density above 1 m, (c) difference in modelled density above 1 m when using surface reflectance (SR) in addition to SAR for prediction, and (d) difference in the prediction interval width when using both SR and SAR for prediction (blue indicates reduced uncertainty). Q10, Q50 and Q90 represent the 10th, 50th and 90th percentile of predictions, respectively. Model names and their order for the purposes of calculating differences are specified above each plot.

## 5.4 Quantifying temporal dynamics

The models calibrated using data from 01/06/2023 to 31/05/2024 were then applied to imagery from previous years to evaluate changes in vegetation 3D structural metrics over time. The changes in Plant Area Index (PAI) and density above 1 m ('dens\_>1m') are illustrated in Figure 17. The trends observed indicate that vegetation cover has increased since 2019 (Figure 17e,f). Note that while the uncertainty bands are wide (i.e., red shading), these values are calculated across the whole scene and should be interpreted carefully, as they reflect the uncertainty in the model predictions from the calibrated models, and therefore are influenced by the training data used. While increases in vegetation structural metrics are observed across the majority of the study region, there is a smaller region in the centre of the study region (i.e., red pixels in Figure 17c,d) where decreases were observed.

Further investigation of these areas (see subset regions in Figure 18) shows varying patterns of decrease in vegetation density. For example, between 2019 and 2021 there are increases in vegetation structural attributes (Figure 18c,e) associated with growth over exposed soils (Figure 18a), and disturbances (i.e., reduced structure) that display a response to existing roads and infrastructure in the region. Low to moderate changes in vegetation cover are observed across the study region between 2022 and 2024, with strong decreases clearly observable where vegetation has been cleared for roads and development infrastructure. Time-series plots at three different point locations demonstrate the ability of the predictive models to track changes in vegetation structure over time on a per-pixel basis, with quantified uncertainty (Figure 19).

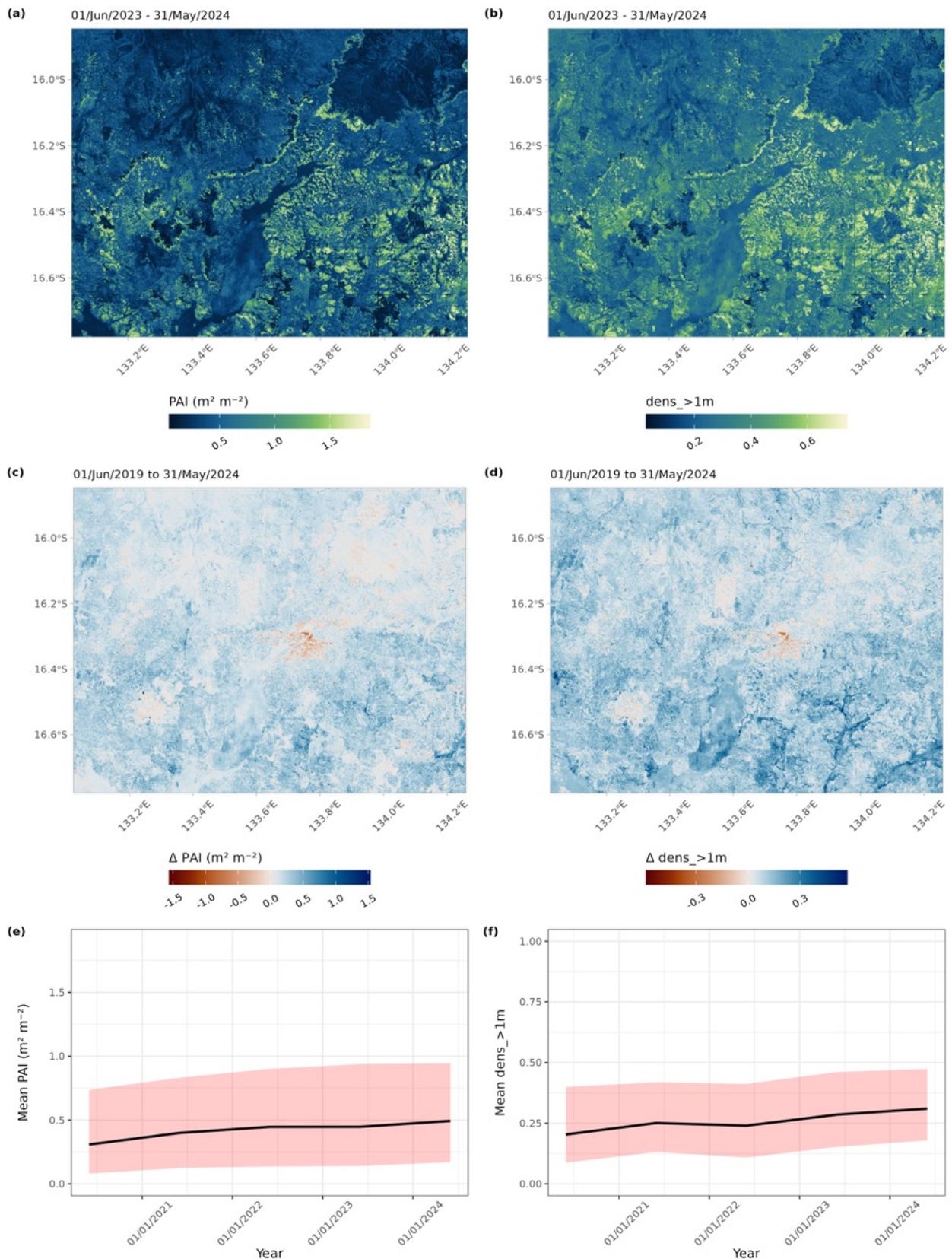


Figure 17 – Modelled predictions of (a) PAI and (b) density above 1 m height for year beginning 01/June/2024, the difference in (c) PAI and (d) density above 1 m height since the year beginning 01/June/2019, and the annual trend in mean (e) PAI and (f) density above 1 m height calculated across the whole study region. The decrease in PAI and density above 1 m height (c, d) is associated with development activity, including roadworks. Red shading (e, f) shows the prediction interval given by the 10<sup>th</sup> and 90<sup>th</sup> percentile.

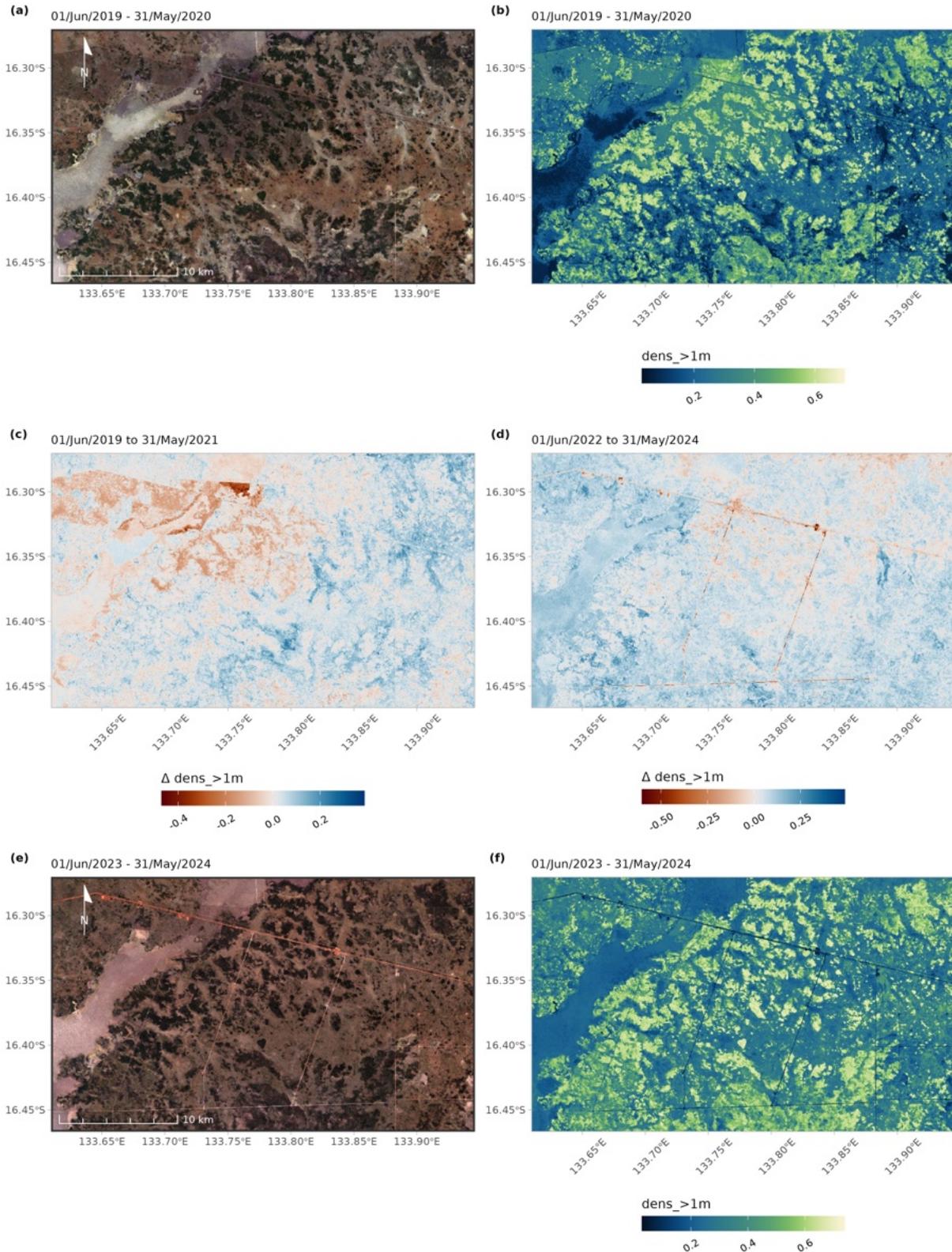


Figure 18 – Visualisation of change in density above 1 m (dens\_>1m) over time in a subset of the study region. Images show annual RGB composites for (a) the year beginning 01/June/2019 and (e) year beginning 01/June/2023, return density above 1 m height for (b) the year beginning 01/June/2019 and (f) year beginning 01/June/2023, and change in density above 1 m between (c) 01/June/2019 and 31/May/2021, and (d) 01/June/2022 and 31/May/2024.

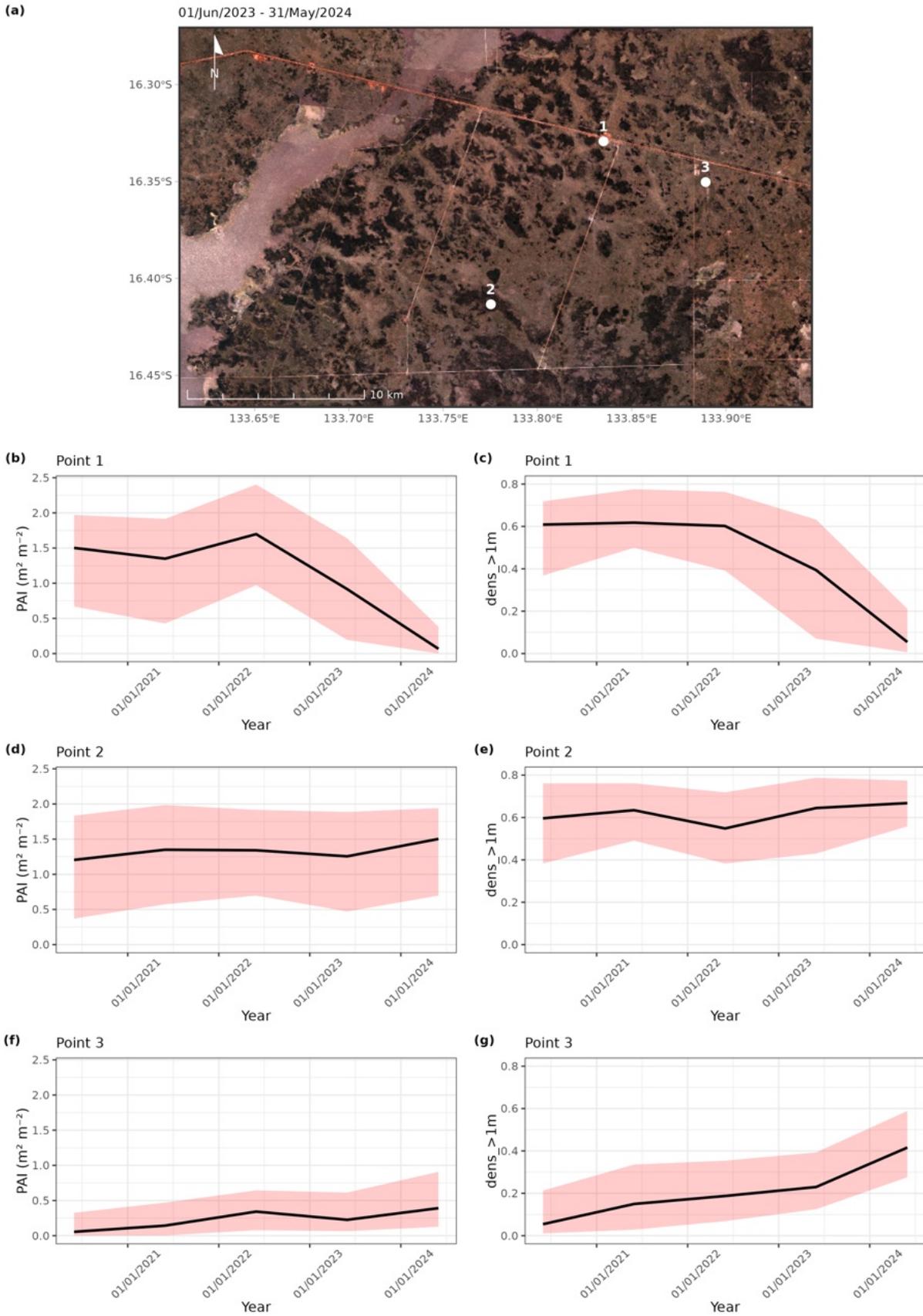


Figure 19 – Time-series of Plant Area Index (PAI) and density above 1 m height ( $\text{dens}_{>1\text{m}}$ ) at (a) three point locations where vegetation cover (b, c) decreases, (d, e) remains stable, and (f, g) increases. Red shading shows the prediction interval given by the 10th and 90th percentile.

## 6 Outlook for large area structural monitoring

### 6.1 Key findings from this study

This study has shown that key attributes of vegetation 3D structure in the Beetaloo region can be mapped at scale with open-source satellite imagery and processing tools. Vertical and horizontal components of woody vegetation, as quantified from UAV-LiDAR, were strongly correlated with imagery from spaceborne sensors at 20 m resolution. The strength of these relationships varied with time of year, being stronger in the dry season months than in the wet season months. These findings lay the foundation for systematic on-going monitoring of vegetation structure over the broader region.

Gap Fraction Proportion (*GFP*), Plant Area Index (*PAI*), canopy density (*dens\_>1m*), 75<sup>th</sup> percentile height (*p75*), and standard deviation of height (*std*) were among the best performing metrics in our predictive models. Combined, these five variables provide a comprehensive representation of vegetation 3D structure. The *dens\_>1m* metric used here is derived from the same equation used by Fisher et al. (2020) for their the *d<sub>all</sub>* metric, which was well correlated with field-measured canopy data across diverse vegetation types in Australia. The ability to detect the mere presence of woody vegetation is the first step in monitoring its extent, but the ability to map patterns in gap fraction and leaf area distribution over large area provides an avenue for understanding changes in condition. Vegetation height is a fundamental ecological attribute, and whilst maximum canopy height can be mapped with cm accuracy and precision from UAV-LiDAR (Bohn Reckziegel et al., 2025), our satellite-based predictions at 20 m resolution were most reliable for the 75<sup>th</sup> percentile height (CCC = 0.87). Models of the higher height percentiles (90<sup>th</sup>, 95<sup>th</sup>, 98<sup>th</sup>) did not perform as well (CCC = 0.72-0.76) but this dampening effect on height is to be expected when pixel sizes exceed the diameter of typical trees in the landscape.

Seasonal metrics performed best in our models, with the dry season months providing the strongest correlations with the UAV-LiDAR metrics. Sentinel-1 and Sentinel-2 data are collected systematically across Australia, with a cadence of 12 and 10 days respectively. These frequent revisit times enable flexibility in the temporal compositing of imagery, and deciding on the most appropriate compositing window depends upon the desired application. Over the five years' worth of data included in our analysis, the dry season months provides a more stable baseline for the SAR characterisation of vegetation structure, by avoiding scattering effects of saturated soils and standing water in the landscape. These months were also optimal from a multispectral optical perspective, as the contrast between dry herbaceous material and evergreen woody canopies was most pronounced. Important to note however is that late dry season imagery can be subject to large fire scars and smoke interference. Annual composites with measures of intra-annual variability also performed very well, and could provide a more stable avenue for assessing change when infrequent processes like fires or floods are present in the landscape. Models built upon annual composites and measures of intra-annual variability are also likely to be more transferable to regions with different rainfall seasonal patterns.

The fusion of SAR (Sentinel-1) and SR (Sentinel-2) performed better than SAR alone. The SAR only models performed well for multiple structural metrics, but the extra dimensionality added by the multi-spectral sensor improved accuracy and reduced bias in the spatial predictions.

A limitation of this study is one that is inherent to UAV remote sensing as a whole – spatial extent. UAV systems offer good flexibility, but they are typically operated over much smaller spatial extents than crewed aircraft, due to technical (battery life) and regulatory (line of sight) restrictions. The Acecore NOA hexacopter with Riegl VUX-120 system used in this study is at the forefront of commercially available systems, with class leading coverage and data quality. Nonetheless, the training and validation of our models was based on a spatially concentrated 1,000 ha of point-cloud data. Whilst this coverage did span good variation in canopy density and vegetation types that are representative of the region, it was geographically focused to the southern side of the Carpentaria highway. This concentration of data collection was not by design but arose through access limitations imposed by highway roadworks and wet season flooding. The density and precision of the UAV-LiDAR point-clouds used in this study exceeds what can be achieved from traditional airborne laser scanning (ALS), and approximates the quality obtained from terrestrial laser scanning (TLS) (Bohn Reckziegel et al., 2025). As such there is a trade-off between data quality and quantity that needs to be balanced, and we prioritised quality so that we could fully explore the relationships between UAV-LiDAR structural metrics and the satellite data products. The spatial extrapolation of metrics to larger areas was a secondary objective.

## 6.2 Future directions

The interpretation of radar imagery presented in this study relies on temporal aggregation of imagery products at seasonal scales. Given the non-stationarity and the importance of major planetary climate patterns affecting Australia (ENSO, SAM and IOD), seasonal amplitude and timing of temperature and precipitation can be highly variable every year even when considering similar seasons. Hence seasonal aggregation, based on fixed periods of several months, might not be representative of the extreme contrasts between dry and wet seasons. Processing of full, multi-year time-series (Figure 18) allows identifying how the radar response fluctuates inter-annually. Comparing these fluctuations with data reflecting the availability of water to plants such as GRACE data (Figure 18a) allows understanding how the vegetation reacts to water abundance or scarcity. We note, for example, a decrease of InSAR coherence of +/- 0.4 (Figure 18b) and a synchronous increase of +/- 3 dB radar backscatter intensity in both VV and VH bands (Figure 18c) during dry periods (Figure 18a). As soil moisture tends to impact VV and VH similarly, the use of band combinations such as the VH/VV ratio and the Radar Vegetation Index (RVI; based on the same bands) decreases the influence of bare soils and moisture, making them particularly suitable for vegetation monitoring applications (Figure 18d). Finally, the use of long time-series along with in situ measurement or LiDAR-derived vegetation metrics allows understanding the normal variation of the radar backscatter over healthy vegetation populations. This is a critical step for identifying deviations from the normal dynamics of vegetation structures, such as the effects of fires or dieback events related to floods, pests or droughts.

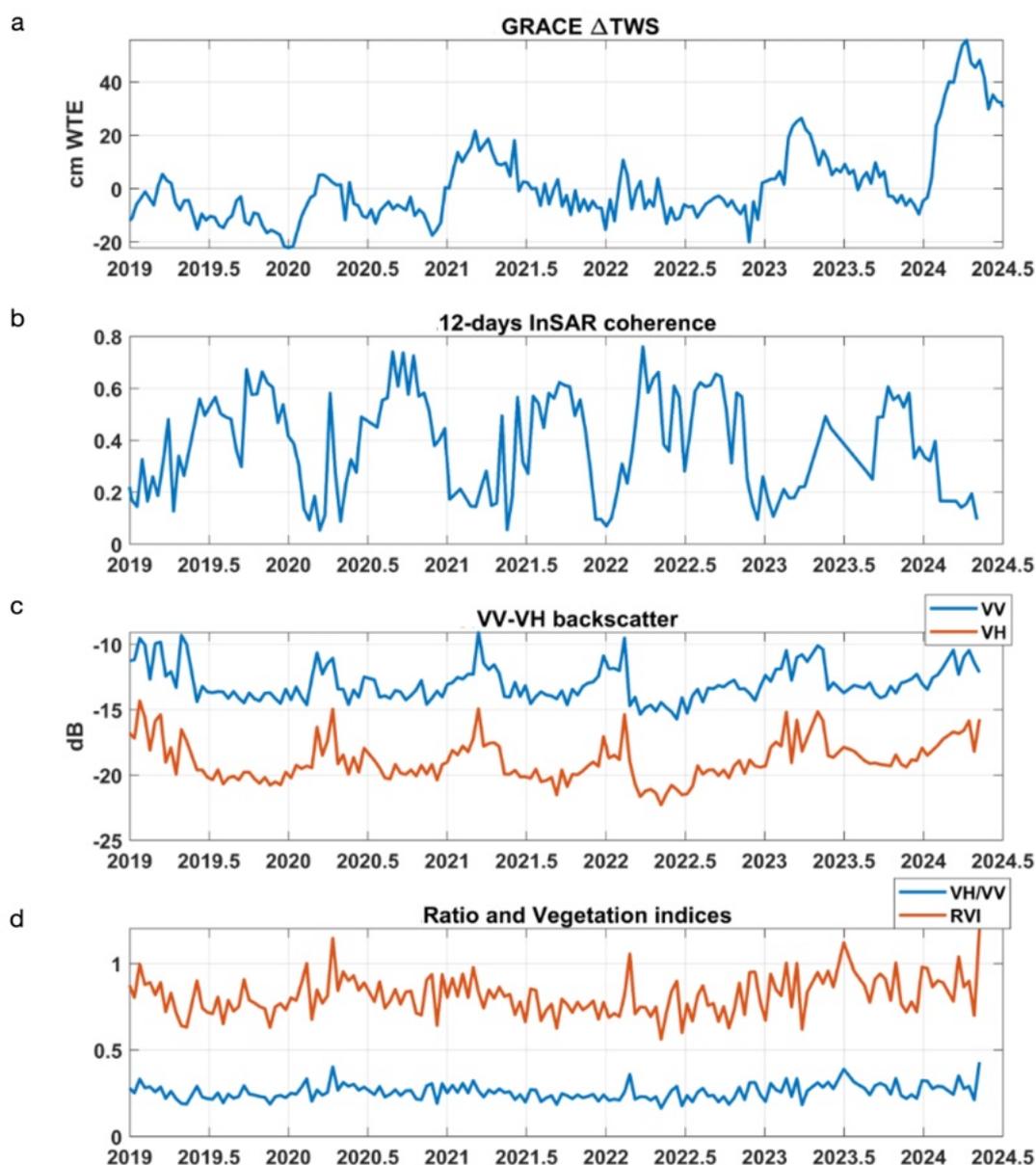


Figure 18 – Changes in the water storage across the landscape and subsurface from GRACE-FO data (changes in Terrestrial Water Storage, or  $\Delta TWS$ , in cm of Water Thickness Equivalent; A), 15-m resolution InSAR coherence time-series using all possible Sentinel-1 image pairs separated by 12 days (B), 15-m resolution VV-VH amplitude time-series (C), VH/VV intensity ratios and RVI (D) over a location covered with Lancewood (Lon/Lat: 133.569961/-16.297862).

## 6.3 Conclusion

This study has shown that key structural attributes of vegetation communities can be modelled with Sentinel-1 SAR imagery, and that predictive performance is improved through the inclusion of Sentinel-2 multi-spectral imagery. Time-series analysis of the Sentinel-1/-2 archive can support large-scale, automated vegetation structure monitoring applications, and model uncertainty can be reduced through the inclusion of additional LiDAR-based calibration and validation datasets. Fusion of Sentinel-1 with longer-wavelength SAR sensors, such as the recently launched P-band BIOMASS mission (European Space Agency), and the upcoming L-band NISAR mission (NASA/JPL), will likely improve the characterisation of vegetation 3D structural attributes even further, providing greater confidence in the assessment of changes that might impact biodiversity.

## 7 References

- Albino, F., Biggs, J., Yu, C., & Li, Z. (2020). Automated Methods for Detecting Volcanic Deformation Using Sentinel-1 InSAR Time Series Illustrated by the 2017–2018 Unrest at Agung, Indonesia. *Journal of Geophysical Research: Solid Earth*, *125*(2), e2019JB017908. <https://doi.org/https://doi.org/10.1029/2019JB017908>
- Bae, S., Levick, S. R., Heidrich, L., Magdon, P., Leutner, B. F., Wöllauer, S., Serebryanyk, A., Nauss, T., Krzystek, P., Gossner, M. M., & others. (2019). Radar vision in the mapping of forest biodiversity from space. *Nature Communications*, *10*(1), 4757.
- Bohn Reckziegel, R., Lowe, T., Devereux, T., Johnson, S. M., Rochelmeyer, E., Hutley, L. B., Doody, T., & Levick, S. R. (2025). Assessing the reliability of woody vegetation structural characterisation from UAV-LS in a tropical savanna. *Science of Remote Sensing*, *11*, 100178. <https://doi.org/https://doi.org/10.1016/j.srs.2024.100178>
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). Classification and regression trees. In *Classification and Regression Trees*. <https://doi.org/10.1201/9781315139470>
- Castellazzi, P., Doody, T., & Peeters, L. (2019). Towards monitoring groundwater-dependent ecosystems using synthetic aperture radar imagery. *Hydrological Processes*, *33*(25). <https://doi.org/10.1002/hyp.13570>
- Castellazzi, P., Khan, S., Walker, S. J., Bartley, R., Wilkinson, S. N., & Normand, J. C. L. (2023). Monitoring erosion in tropical savannas from C-band radar coherence. *Remote Sensing of Environment*, *290*. <https://doi.org/10.1016/j.rse.2023.113546>
- Castellazzi, P., Longuevergne, L., & Feng, W. (2024). Bridging the Scale Gap Between Ground Deformation and Gravity: Tools for Sustainability. In *Remote Sensing for Characterization of Geohazards and Natural Resources* (pp. 493–505). Springer.
- Daniel Kissling, W., Shi, Y., Wang, J., Walicka, A., George, C., Moeslund, J. E., & Gerard, F. (2024). Towards consistently measuring and monitoring habitat condition with airborne laser scanning and unmanned aerial vehicles. *Ecological Indicators*, *169*, 112970. <https://doi.org/10.1016/J.ECOLIND.2024.112970>
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*. <https://doi.org/10.1016/j.rse.2011.11.026>
- Fisher, A., Armston, J., Goodwin, N., & Scarth, P. (2020). Modelling canopy gap probability, foliage projective cover and crown projective cover from airborne lidar metrics in Australian forests and woodlands. *Remote Sensing of Environment*, *237*. <https://doi.org/10.1016/j.rse.2019.111520>
- Hill, R. A., & Hinsley, S. A. (2015). Airborne lidar for woodland habitat quality monitoring: Exploring the significance of lidar data characteristics when modelling organism-habitat relationships. *Remote Sensing*, *7*(4). <https://doi.org/10.3390/rs70403446>
- Huddleston-Holmes, C. R., Frery, E., Wilkes, P., Bailey, A. H. E., Bernadel, G., Brandon, C., Buchanan, S., Cook, S. B., Crosbie, R. S., & Evans, T. (2020). *Geological and environmental baseline assessment for the Beetaloo GBA region. Geological and bioregional assessment program: stage 2. Department of the Environment and Energy, Bureau of Meteorology*. Australia: Department of the Environment and Energy, Bureau of Meteorology ....
- Huth, N., Llewellyn, R., Kuehne, G., Thomas, M., Ratcliff, C., & Bramley, R. (2019). *Understanding natural gas impacts and opportunities on agriculture in the South East of South Australia*. CSIRO, Australia. <https://gisera.csiro.au/wp-content/uploads/2019/12/Gas....>

- Jensen, J. R. (2014). Remote sensing of the environment: an earth resource perspective second edition. In *Pearson Education Limited, Harlow, England* (Vol. 1).
- Lefsky, M. A., Cohen, W. B., Parker, G. G., & Harding, D. J. (2002). Lidar Remote Sensing for Ecosystem Studies. *BioScience*, *52*(1), 19–30. [https://doi.org/10.1641/0006-3568\(2002\)052\[0019:lrsfes\]2.0.co;2](https://doi.org/10.1641/0006-3568(2002)052[0019:lrsfes]2.0.co;2)
- Lemoine, J. M. (2023). *RL05 GSM monthly coefficients*.
- MacArthur, R. H., & MacArthur, J. W. (1961). On bird species diversity. *Ecology*, *42*(3), 594–598.
- Massonnet, D., & Feigl, K. L. (1998). Radar interferometry and its application to changes in the Earth's surface. *Reviews of Geophysics*, *36*(4), 441–500. <https://doi.org/https://doi.org/10.1029/97RG03139>
- Meinshausen, N. (2006). Quantile regression forests. *Journal of Machine Learning Research*, *7*.
- Mutanga, O., Masenyama, A., & Sibanda, M. (2023). Spectral saturation in the remote sensing of high-density vegetation traits: A systematic review of progress, challenges, and prospects. *ISPRS Journal of Photogrammetry and Remote Sensing*, *198*, 297–309. <https://doi.org/10.1016/J.ISPRSJPRS.2023.03.010>
- Nagendra, H., Lucas, R., Honrado, J. P., Jongman, R. H. G., Tarantino, C., Adamo, M., & Mairota, P. (2013). Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecological Indicators*, *33*, 45–59. <https://doi.org/10.1016/J.ECOLIND.2012.09.014>
- Parker, A. L., Pigois, J. P., Filmer, M. S., Featherstone, W. E., Timms, N. E., & Penna, N. T. (2021). Land uplift linked to managed aquifer recharge in the Perth Basin, Australia. *International Journal of Applied Earth Observation and Geoinformation*, *105*, 102637. <https://doi.org/10.1016/J.JAG.2021.102637>
- Pavey, C. R., McDonald, S. M., Midgley, D. J., Tran-Dinh, N., Mariani, C., Vanderduys, E. P., Murray, B., Johnson, S., Levick, S. R. (2025) *Remote sensing and threatened species surveys to better understand risks of forest fragmentation from the Narrabri Gas Project*. CSIRO, Australia.
- Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H. G., Scholes, R. J., Bruford, M. W., Brummitt, N., Butchart, S. H. M., Cardoso, A. C., Coops, N. C., Dullo, E., Faith, D. P., Freyhof, J., Gregory, R. D., Heip, C., Höft, R., Hurtt, G., Jetz, W., ... Wegmann, M. (2013). Essential biodiversity variables. In *Science* (Vol. 339, Issue 6117, pp. 277–278). <https://doi.org/10.1126/science.1229931>
- Roberts, D., Dunn, B., & Mueller, N. (2018). Open Data Cube products using high-dimensional statistics of time series. *International Geoscience and Remote Sensing Symposium (IGARSS), 2018-July*. <https://doi.org/10.1109/IGARSS.2018.8518312>
- Roberts, D., Mueller, N., & McIntyre, A. (2017). High-Dimensional Pixel Composites from Earth Observation Time Series. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(11). <https://doi.org/10.1109/TGRS.2017.2723896>
- Roussel, J. R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Meador, A. S., Bourdon, J. F., de Boissieu, F., & Achim, A. (2020). lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. In *Remote Sensing of Environment* (Vol. 251). <https://doi.org/10.1016/j.rse.2020.112061>
- Tapley, B. D., Bettadpur, S., Watkins, M., & Reigber, C. (2004). The gravity recovery and climate experiment: Mission overview and early results. *Geophysical Research Letters*, *31*(9). <https://doi.org/10.1029/2004GL019920>
- Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B. Ö., Floury, N., Brown, M., Traver, I. N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., ... Rostan, F. (2012). GMES Sentinel-1 mission. *Remote Sensing of Environment*. <https://doi.org/10.1016/j.rse.2011.05.028>
- Toulemont, A., OLIVIER, M., CLERC, S., BELLOUARD, R., REINA, F., GASCON, F., LUCE, J.-F., MAVROCORDATOS, C., & BOCCIA, V. (2021). *Copernicus Sentinel-2C/D Multi Spectral Instrument full field of view spectral characterization*. <https://doi.org/10.1117/12.2593729>

- Veloso, A., Mermoz, S., Bouvet, A., Le Toan, T., Planells, M., Dejoux, J. F., & Ceschia, E. (2017). Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sensing of Environment*, *199*, 415–426. <https://doi.org/10.1016/J.RSE.2017.07.015>
- Vihervaara, P., Auvinen, A. P., Mononen, L., Törmä, M., Ahlroth, P., Anttila, S., Böttcher, K., Forsius, M., Heino, J., Heliölä, J., Koskelainen, M., Kuussaari, M., Meissner, K., Ojala, O., Tuominen, S., Viitasalo, M., & Virkkala, R. (2017). How Essential Biodiversity Variables and remote sensing can help national biodiversity monitoring. *Global Ecology and Conservation*, *10*, 43–59. <https://doi.org/10.1016/j.gecco.2017.01.007>
- Villarroya-Carpio, A., Lopez-Sanchez, J. M., & Engdahl, M. E. (2022). Sentinel-1 interferometric coherence as a vegetation index for agriculture. *Remote Sensing of Environment*, *280*. <https://doi.org/10.1016/j.rse.2022.113208>
- Wright, M. N., & Ziegler, A. (2017). Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, *77*(1). <https://doi.org/10.18637/jss.v077.i01>
- Zebker, H. A., & Villasenor, J. (1992). Decorrelation in interferometric radar echoes. *IEEE Transactions on Geoscience and Remote Sensing*, *30*(5). <https://doi.org/10.1109/36.175330>

## 8 Appendix 1

### 8.1 Additional datasets and processing steps

#### 8.1.1 Interferometric radar coherence processing

Interferometric radar coherence, or InSAR coherence, is a radar imagery product that combines both the like-polarised band intensity (VV) and the phase information (Zebker & Villasenor, 1992). This similarity index aims at understanding if changes occurred in the ground targets which dominates the radar backscatter during the time interval between a pair of radar images. It ranges from 0 to 1, with 0 being allocated to areas of important changes, and 1 being allocated to areas where the phase and intensity information are identical between the two images of the pair. Vegetated areas typically show low to medium coherence, as movement and growth of the leaf, branch and canopy structures constantly changes the scattering mechanisms and the phase value recorded by the sensor. While InSAR coherence was initially developed to select areas where the phase information was sufficiently stable to be used for ground deformation mapping, it has been recently identified as having value for vegetation mapping purposes (Castellazzi et al., 2019, 2024; Villarroya-Carpio et al., 2022).

The GRD Sentinel-1 products are not suitable for creating InSAR coherence products, as the phase information is discarded in such format. Instead, the Single Look Complex (SLC) data format is used. It preserves both amplitude and phase information, allowing production of both intensity and coherence imagery products. This data type requires more processing steps and besides the possibility of producing coherence images, it also allows more flexibility for generating intensity products at different resolutions or with advanced filters.

The coherence time-series were generated using SARSCAPE 6.1, processing steps included: co-registration of each image pair, multilooking, coherence computation, resampling and projection on a 15-m resolution grid. Only pair separated by 12 days were processed, to maximise sensibility to vegetation and form a consistent dataset uninfluenced by varying temporal baselines. To visualise how the coherence changes over time, and for simplification, the date corresponding to each coherence image was set as the centroid of the two acquisition dates used to form the image pair.

#### 8.1.2 GRACE data retrieval and processing steps

The Gravity Recovery And Climate Experiment (GRACE) missions have been measuring the variations of the Earth's gravity field since 2002. The system comprises two satellites on the same low-altitude orbital track (~400 km), distant by approximately 200 km. Earth's gravity field affects the acceleration of the satellites. The first satellite on the orbital track is affected first while flying over a mass anomaly, the second is affected after a short time-delay. By tracking the changes of distance between the two satellites, the Earth's gravity field is measured. A highly precise ranging system retrieves that variation, which is inverted into gravity data during post-processing. The first

GRACE mission ended in 2017, the GRACE-Follow-On (GRACE-FO) mission was operational a year later and is still acquiring to this date.

While this sensing system does not show major advantages in mapping the gravity field, it offers the unique capability to precisely track its variations over time. Such temporal changes, at the timescales of days to decades, are mostly related to the constant changes in the spatial distribution of water masses on Earth (Tapley et al., 2004), in the oceans and on land, at the surface (glaciers, lakes, rivers) and within the sub-surface (vadose zone, aquifers). GRACE is the remote-sensing water monitoring system that is directly volumetric. It requires no calibration, as the relation between gravity, masses and volumes are direct.

GRACE data is expressed in Water Thickness Equivalent (WTE) and represents the aggregation of all changes of water masses occurring below and around the orbital track in reference to past data acquisitions. It is referred to as Total, or Terrestrial Water Storage Change ( $\Delta$ TWS) data. A decomposition strategy, accounting for the water masses occurring over a study area, is often performed to separate the water storage changes of a specific contributor (e.g. a large lake, an aquifer etc. – e.g. Castellazzi et al. (2024)). This strategy requires auxiliary data and is highly dependent on data availability, both in situ and from modelling.

To translate the GRACE measurements along its flight track into usable data expressed in WTE, the data must be processed. Several GRACE data processing approaches, or 'solutions' exist. They all rely on different spatial and temporal aggregation of the original GRACE measurements along its orbital track. They also rely on different approaches to filter noise patterns. For this study, we used the GRGS 10-days GRACE solution from GRGS (Lemoine, 2023). While most GRACE processing centres provide monthly GRACE data, this one is also available at a 10-days frequency, which conveniently matches the temporal frequency of Sentinel-1 and 2 over Australia (12 and 10 days, respectively).

The 'Future Directions' section of this study uses GRACE data to help interpreting time-series of radar backscatter and InSAR coherence. Because radar products are affected by soil moisture, vegetation moisture, and presence of water on the ground, an understanding of the water accumulated in the landscape is useful for interpretation. It is important to keep in mind that GRACE data are low resolution (hundreds of kilometres) and represent regional conditions. It does not represent the conditions specific to the location of radar, optical or UAV-LiDAR data acquisitions presented in this study.

# 9 Appendix 2

## 9.1 Supporting figures

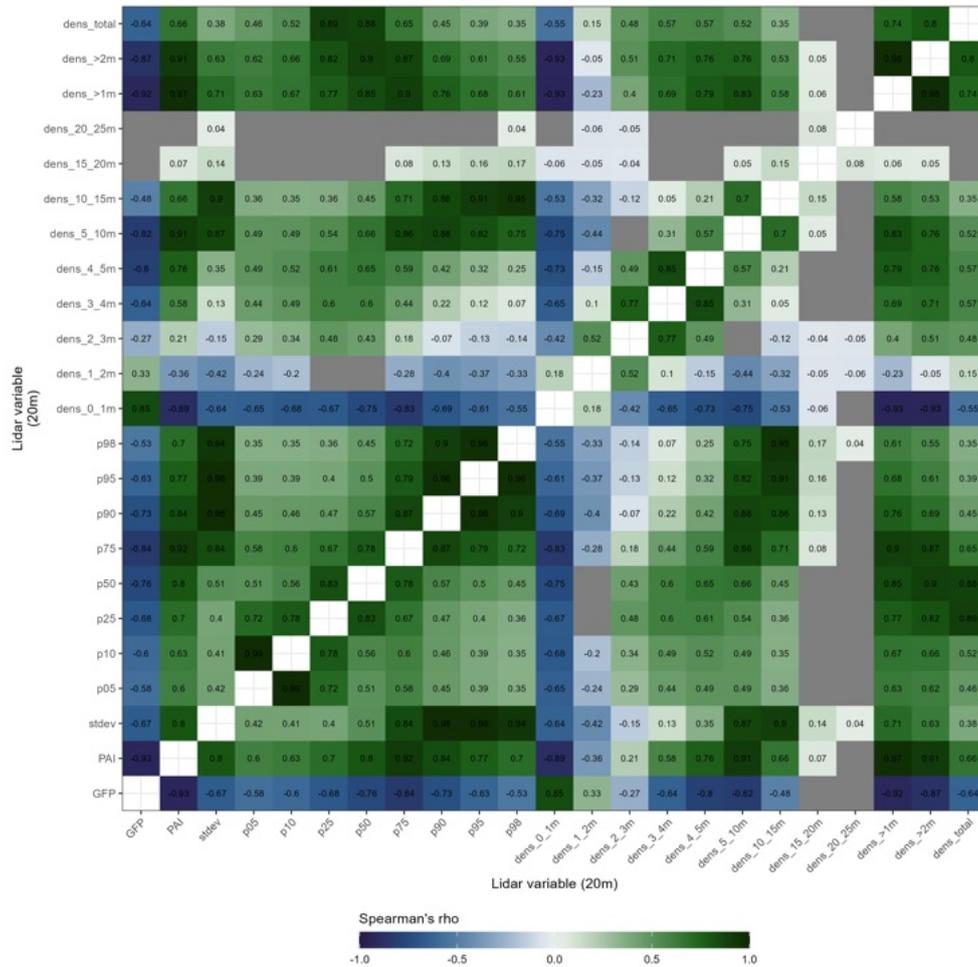


Figure 20 - Pairwise correlations between LiDAR structural metrics. All comparisons were conducted on a randomised spatial sample of point locations (n = 2,923) using Spearman's rank-based correlation coefficient (rho). Non-significant (p ≥ 0.05) correlations are not reported (grey cells).

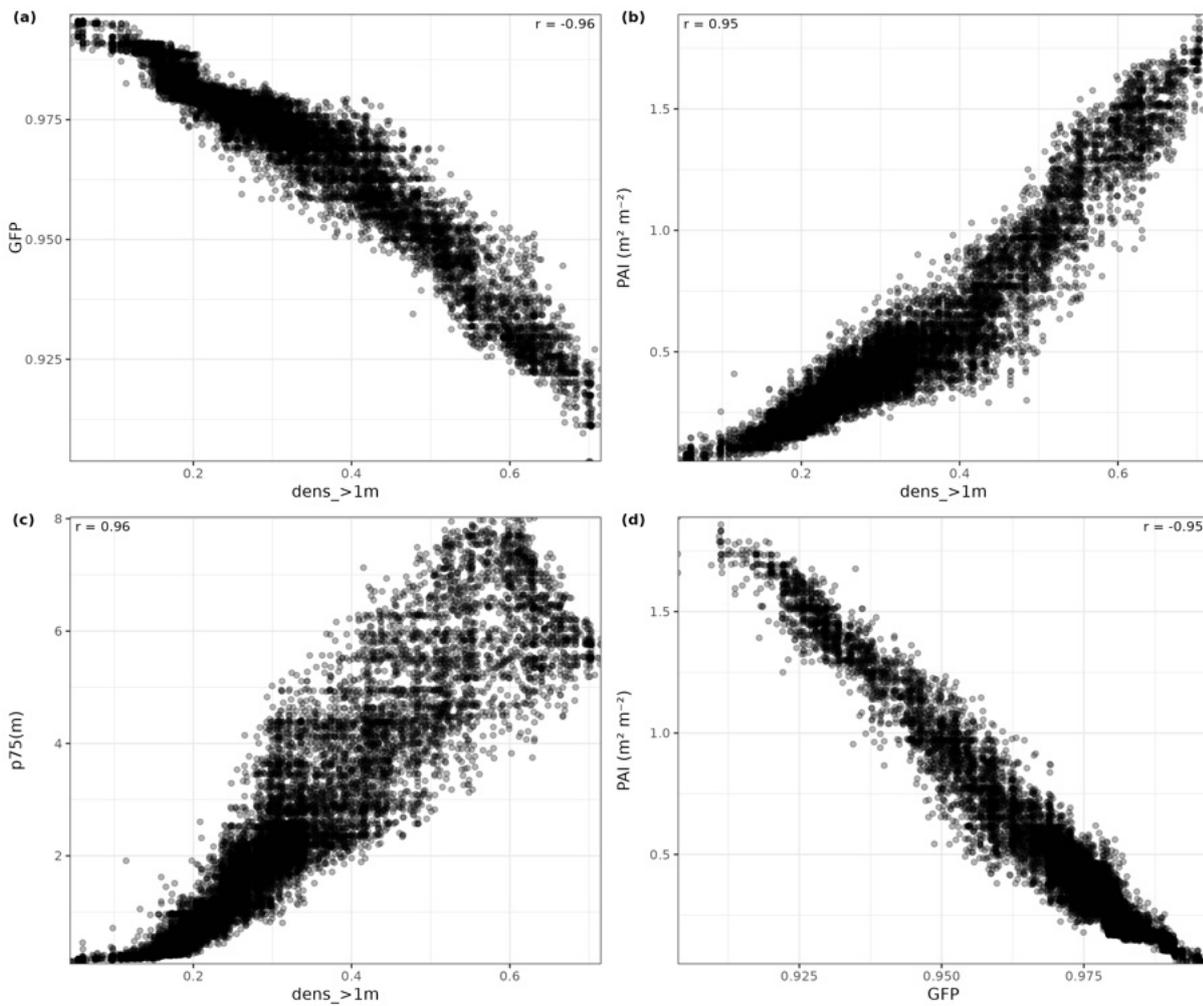
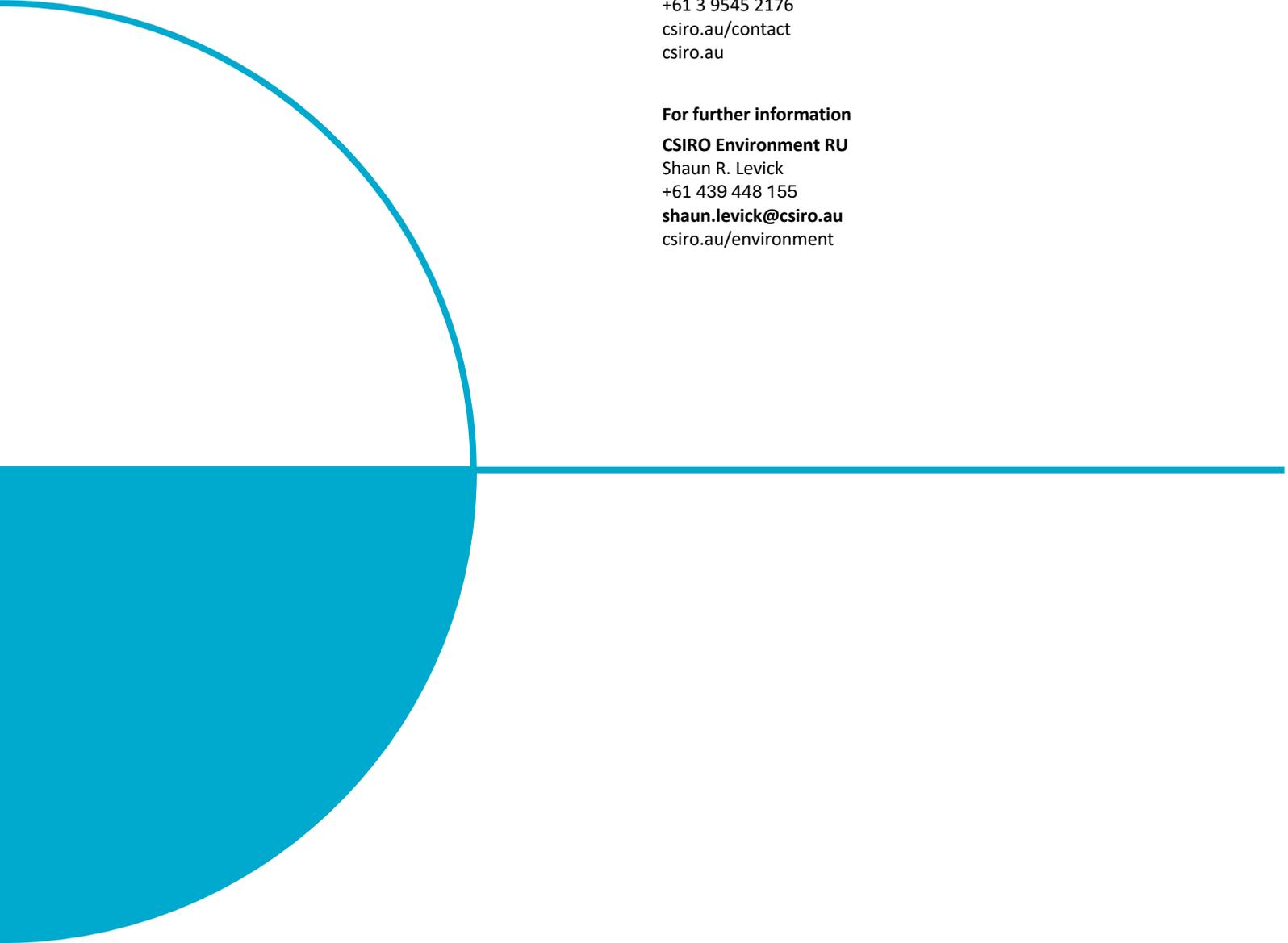


Figure 21 – Scatterplots of spatial random samples ( $n = 15,000$ ) of (a) Gap Fraction Probability (GFP) versus density above 1 m height ( $\text{dens\_>1m}$ ), (b) Plant Area Index (PAI) versus density above 1 m height, (c) 75<sup>th</sup> percentile height ( $\text{p75}$ ) versus density above 1 m height, and (d) PAI versus GFP. Point samples were extracted from extrapolated predictions from models fitted using SAR and SR as explanatory variables.



**As Australia's national science agency and innovation catalyst, CSIRO is solving the greatest challenges through innovative science and technology.**

CSIRO. Unlocking a better future for everyone.

**Contact us**

1300 363 400  
+61 3 9545 2176  
[csiro.au/contact](https://www.csiro.au/contact)  
[csiro.au](https://www.csiro.au)

**For further information**

**CSIRO Environment RU**  
Shaun R. Levick  
+61 439 448 155  
[shaun.levick@csiro.au](mailto:shaun.levick@csiro.au)  
[csiro.au/environment](https://www.csiro.au/environment)