

Calibration of satellite-borne sensors & Validation of satellite-derived information

Randall Donohue
CSIRO Environment



Sensor calibration

Testing the accuracy of a sensor's outputs and adjusting them before they are used

We'll look at how satellite-borne sensors are calibrated

Model training

Using observations to adjust how an information product is derived from sensor outputs (model training/fitting)

Product validation

Using observations to test the accuracy of sensor-derived information products

We'll look at examples of how observations have been used to validate satellite-derived products, and things learned in the process

Calibration of radiometers

Compare measurements to objects with stable and known properties

- Can be used to convert a relative digital signal to real values *in absolute units*
- Can be used to check how accurate a sensor is (has it changed from factory settings?)
- Calibration error = difference between what is measured and what *should* be measured
- Ideally uses end-members (to bracket extreme values)
- Black bodies (1), reflectance panels (2), lamps with specific & accurate irradiance (3)

2



1



3



Calibration of satellite-borne sensors

Once sensors are in space, how can they be calibrated?

What if the sensor's accuracy changes over time?

Launch is a traumatic event for even the bravest of sensors



Satellite-borne sensor calibration

Pre-launch calibration

- lab-based sensor calibration
- no different from ground radiometer calibration



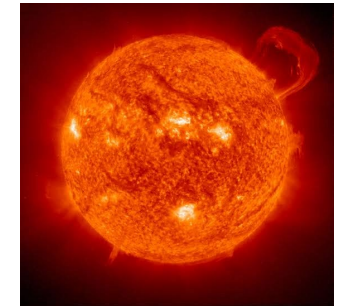
On-board calibration

- various cool tricks in space



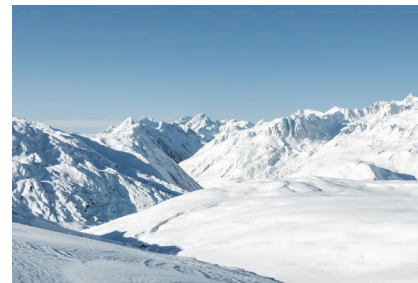
Deep dark
space

A darker-than-
black panel kept
at a very stable
temperature



Post-acquisition calibration

- calibration of downloaded imagery
- pseudo-invariant features (PIFs)



AVHRR as an example

AVHRR is an old satellite program that is still incredibly useful but is plagued with on-going calibration issues

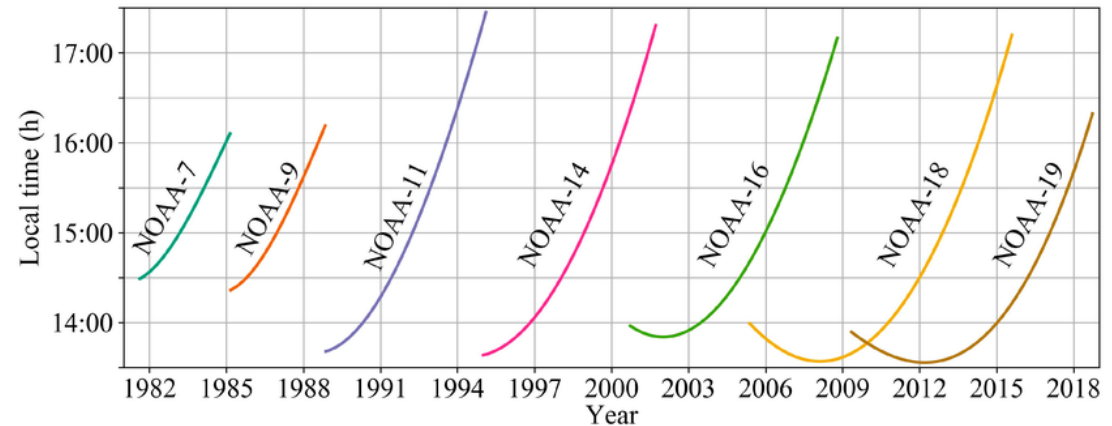
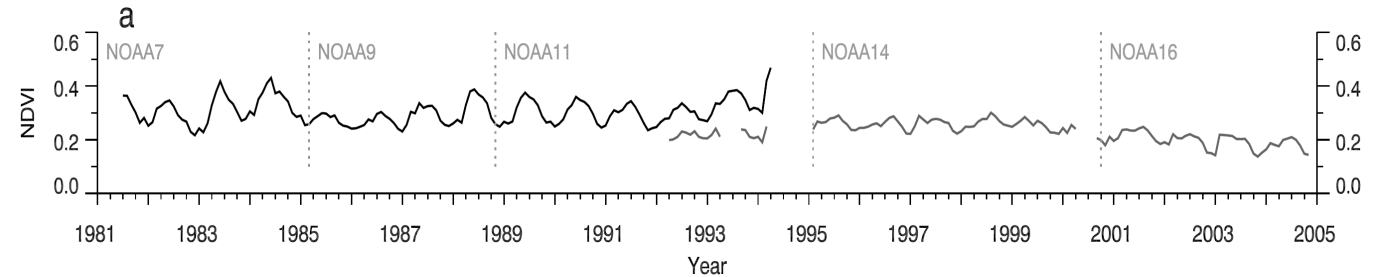
Pre-launch calibration

~~On-board calibration~~

Nup

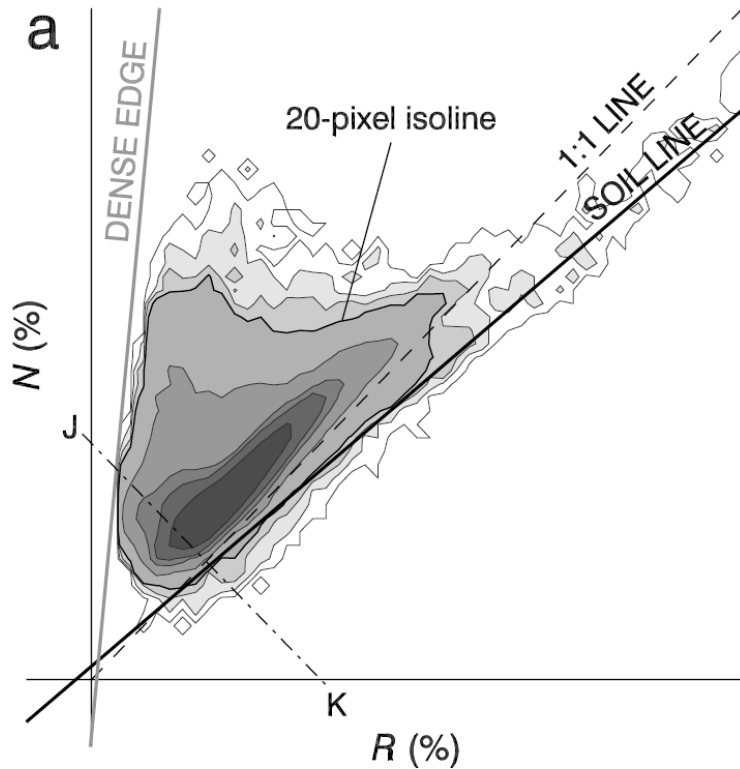
Post-acquisition calibration

Multiple sequential sensors that have calibration drift

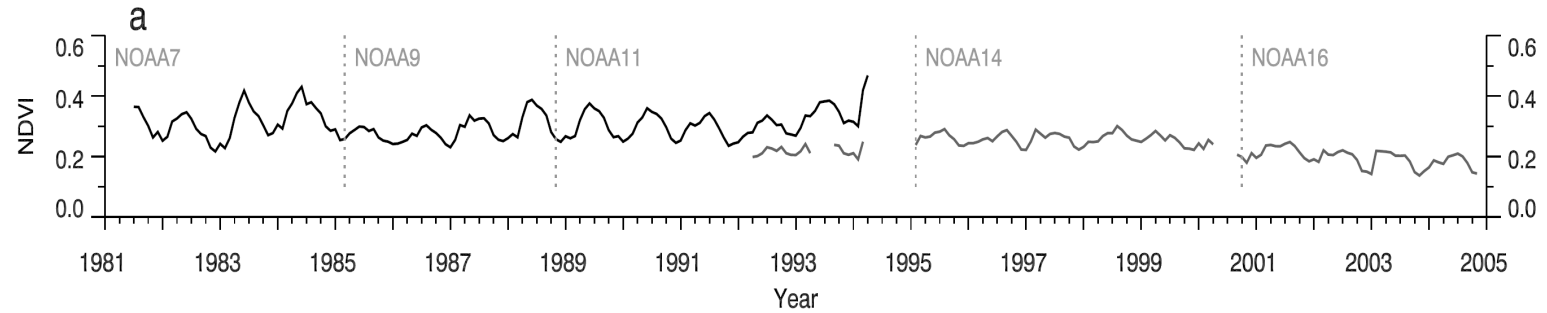


Changes in equatorial crossing time caused by orbital drift

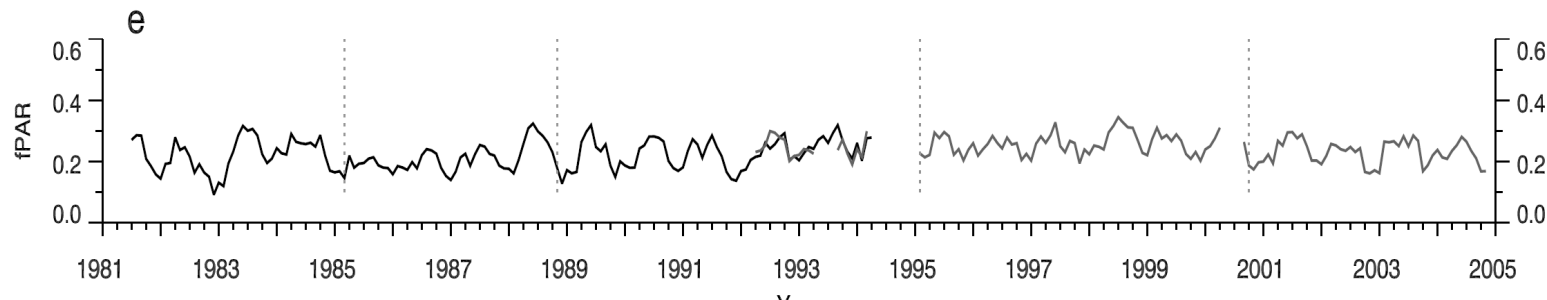
AVHRR as an example



Before cover triangle stabilisation

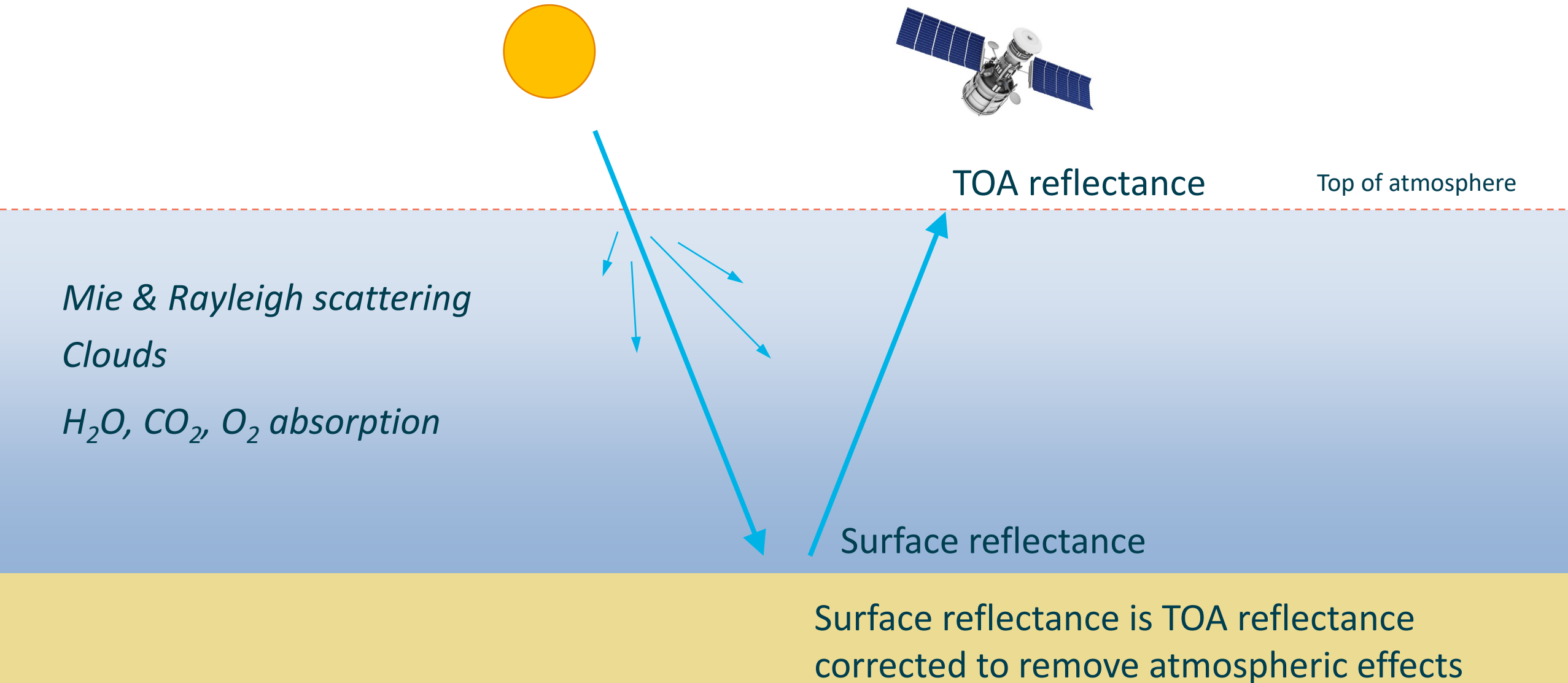


After cover triangle stabilisation

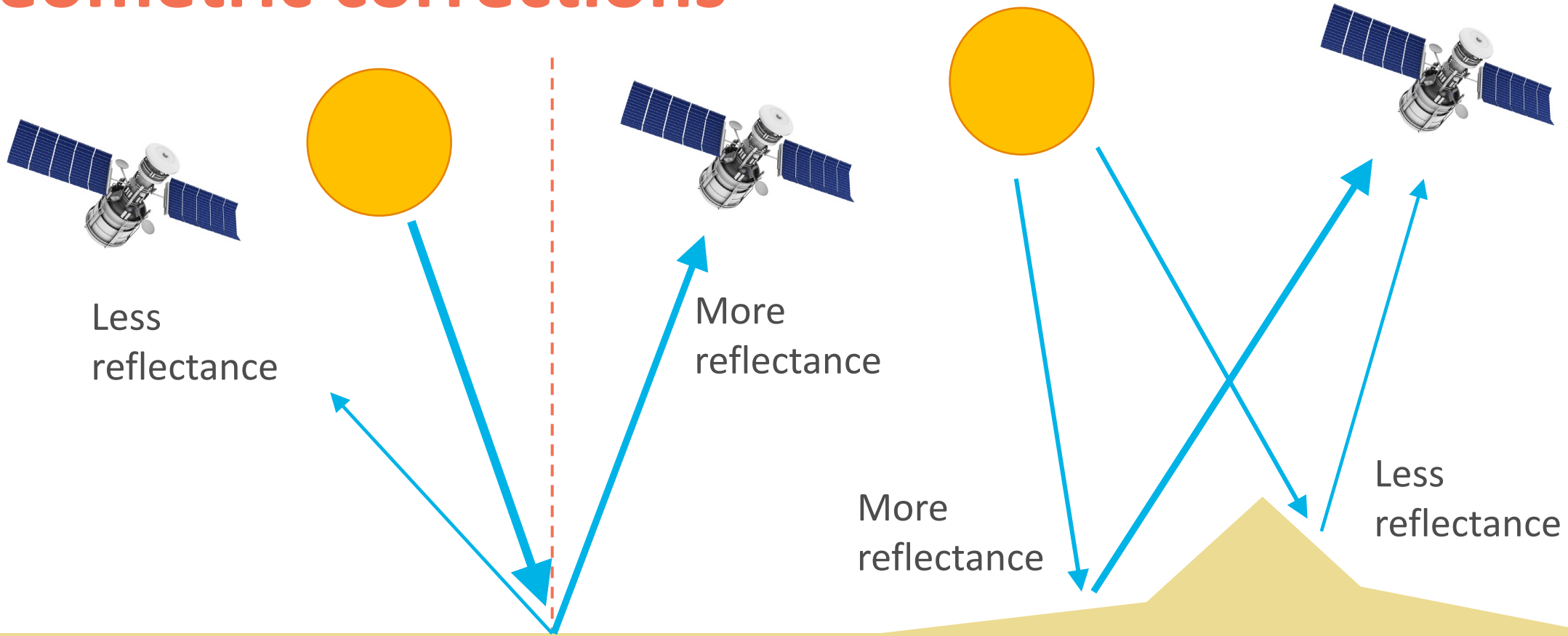


An invariant cover triangle calibration method

Atmospheric corrections



Geometric corrections



Bidirectional reflectance distribution functions (BRDF) correct for differences in sun and viewing angles

In all but flat landscapes, terrain corrections are also important

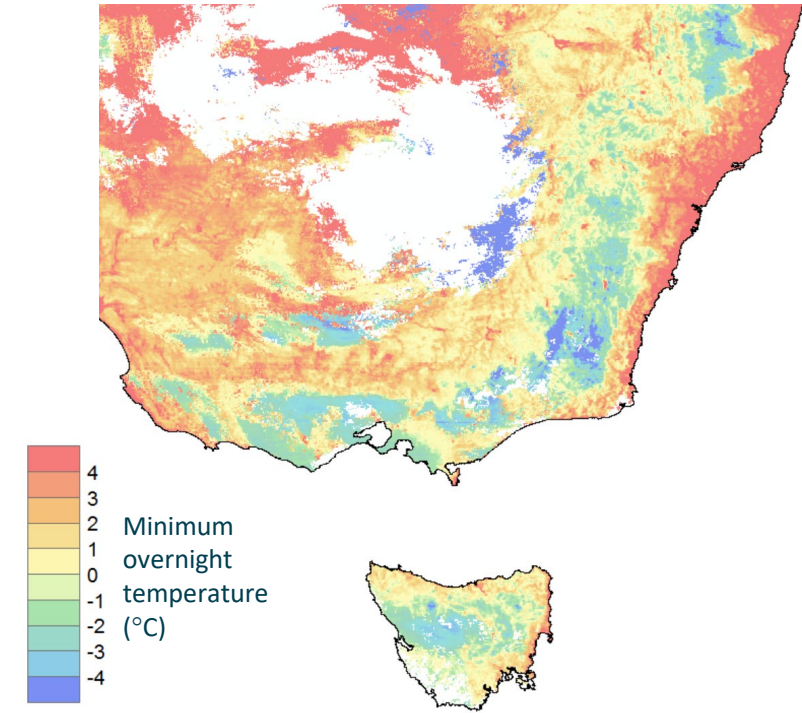
An example from a frost mapping project

Frost causes wide-spread damage to cereal crops each year

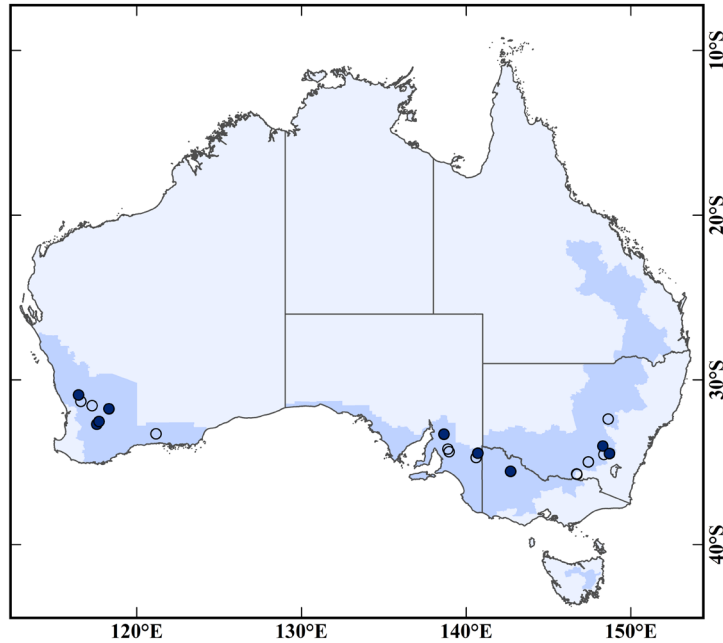
We have no way of knowing where, when and how severely frost occurs (across large areas)

But we can use satellites to measure overnight surface temperatures (at 2 km spatial resolution)

We had to find out how accurately these satellite products are

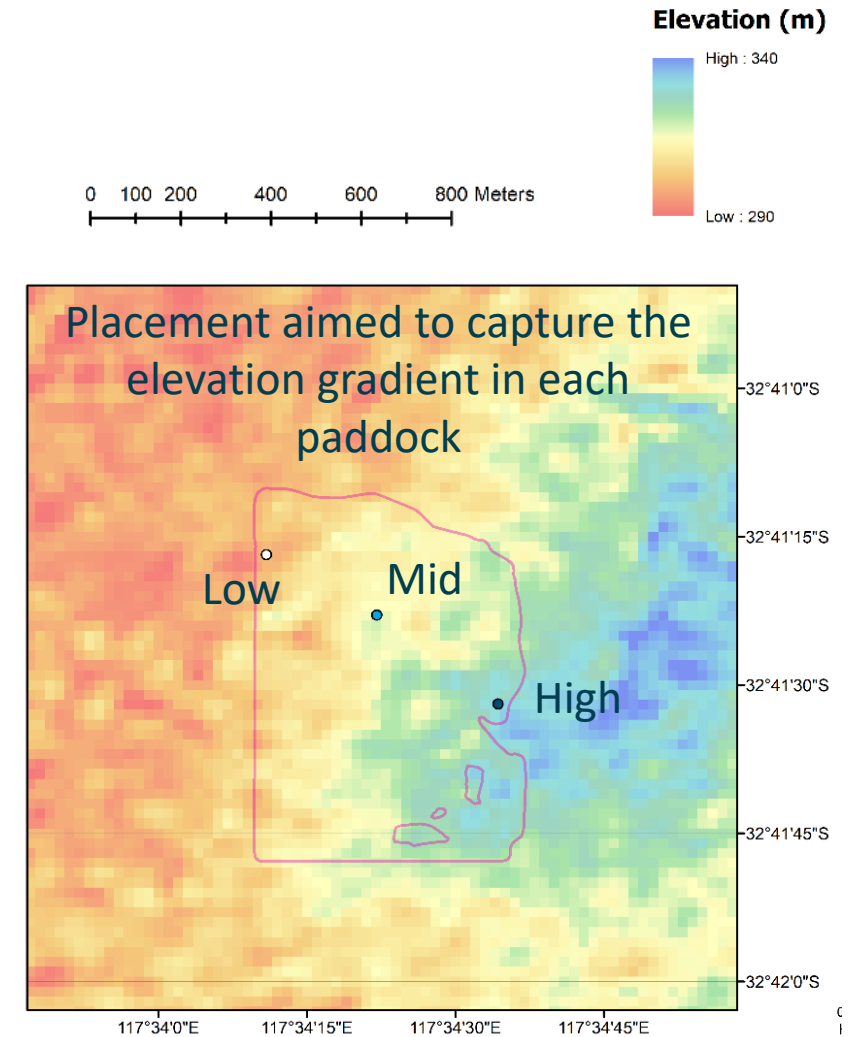


Surface temperature field campaign



We conducted a field campaign in 2024 aimed at testing the relationship between surface temperature and frost damage

Thermal cameras were installed at 57 locations in 23 wheat paddocks across 16 farms



We had to calibrate the thermal cameras



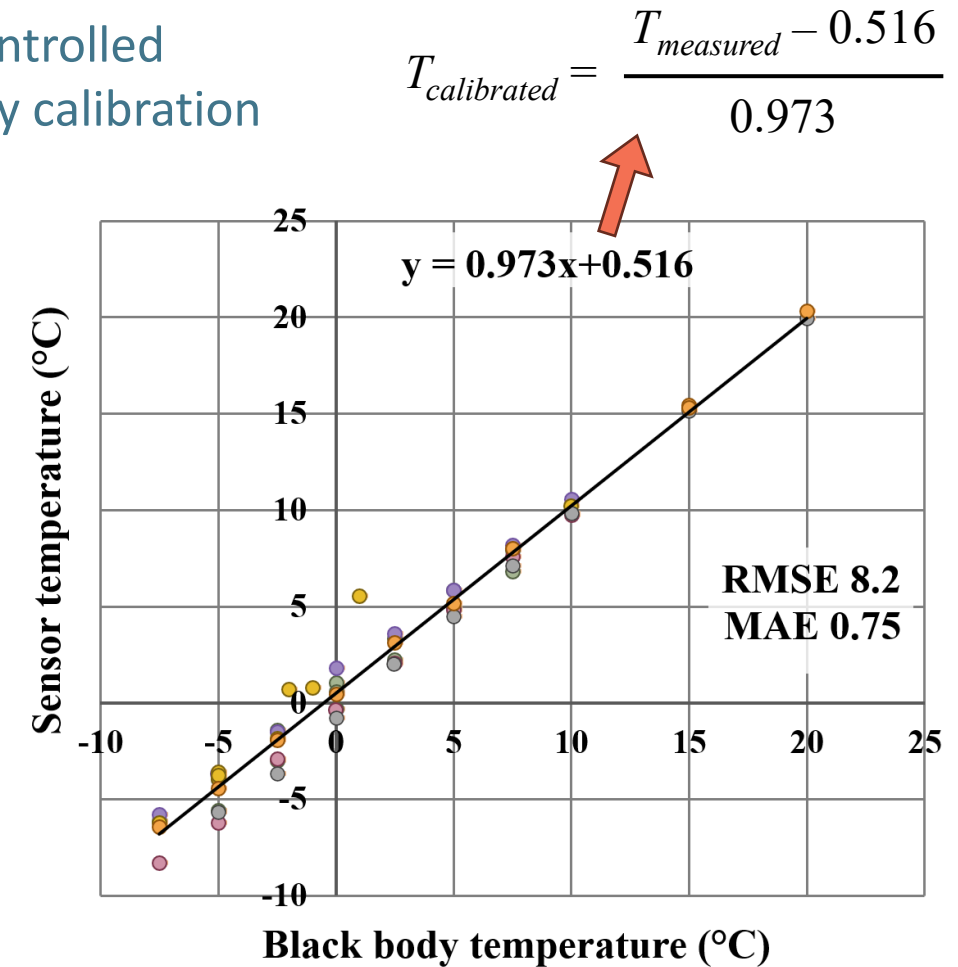
← Temperature-controlled black body cavity calibration devices

Calibrated at temperatures between -7.5 and +10 °C in 2.5° increments

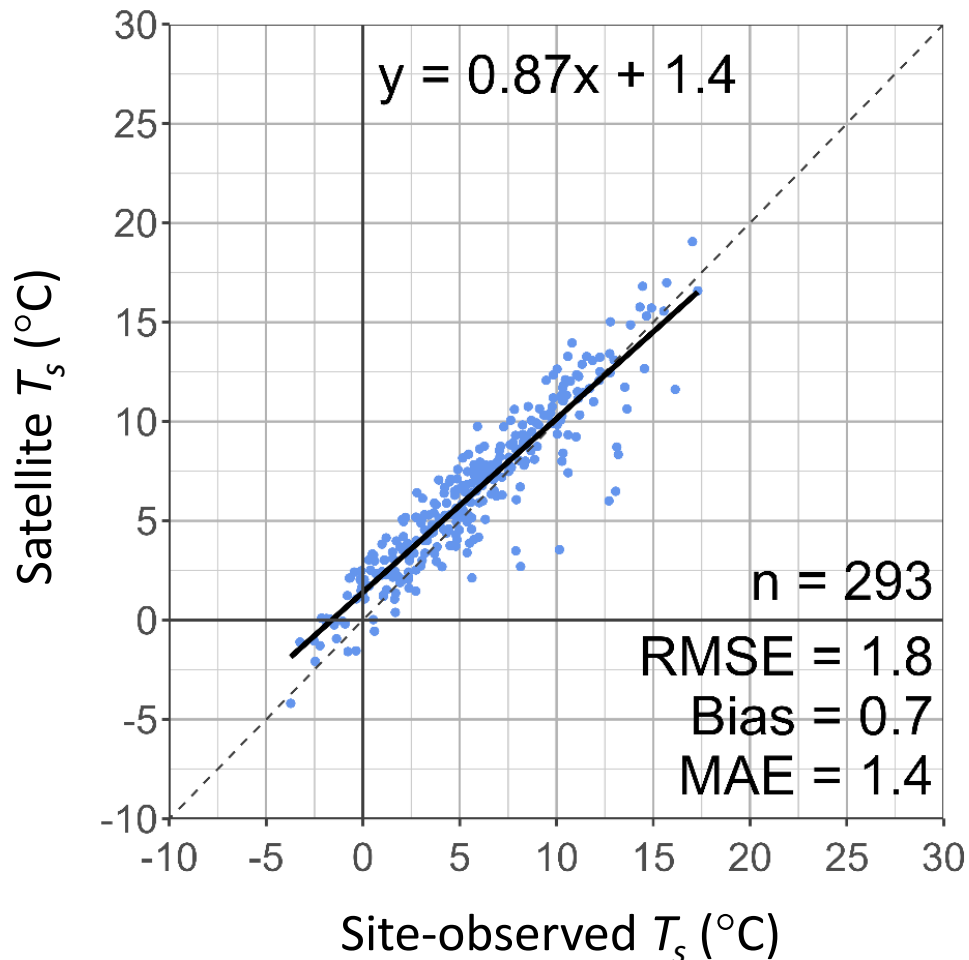
What an utter pain in the ...

...each sensor took about 5 hours to do...

...and it drove me (and my colleagues) bonkers



Surface temperature validation results



Two classic problems

1. The field campaign (sensor placement) wasn't designed for validating 2 km imagery

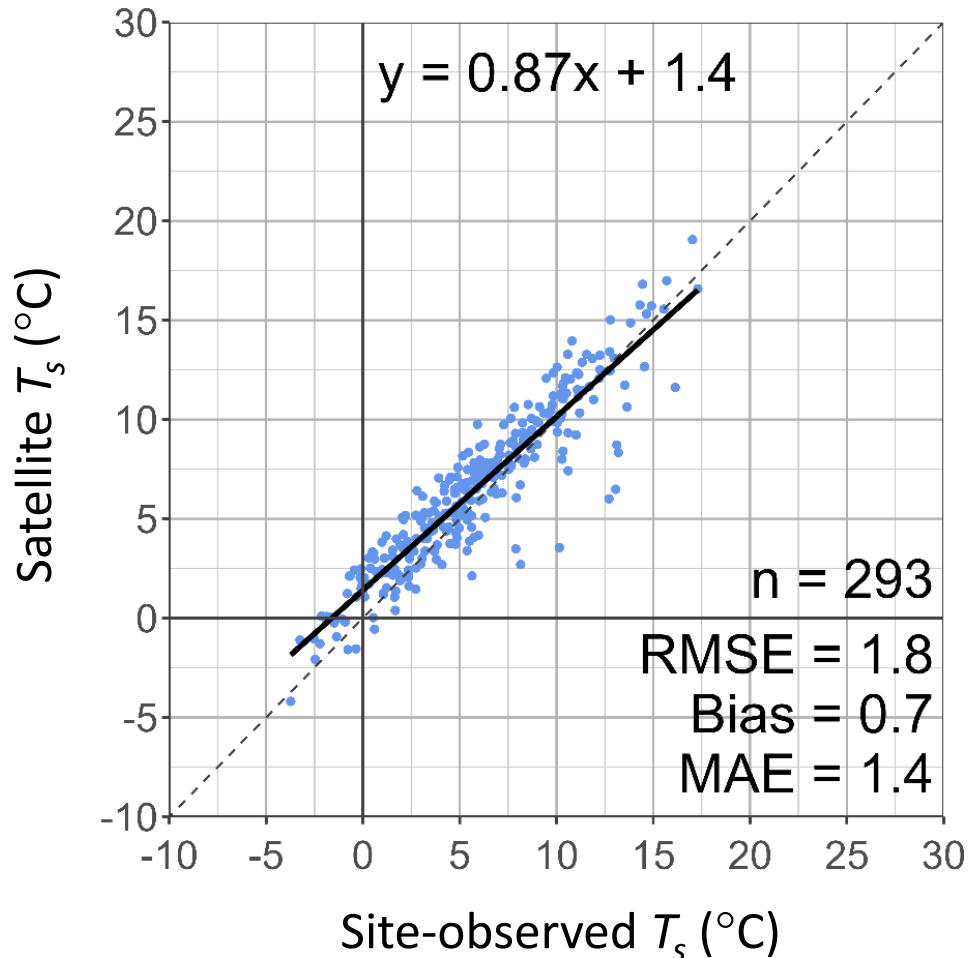
Often field data are used for purposes they weren't intended for

We were able to use data from only 9 of the 23 paddocks

And even in these the spatial representativeness of the sensor locations was poor



Surface temperature validation results



Two classic problems

2. What is truth?

Informal expectation is that surface temp is $\sim 2^\circ\text{C}$ lower than air temp on cold nights.

This isn't what our measurements showed.

So we spent about a year trying to figure out what we were doing wrong. Ans: nothing, we proved the expectation wrong :)

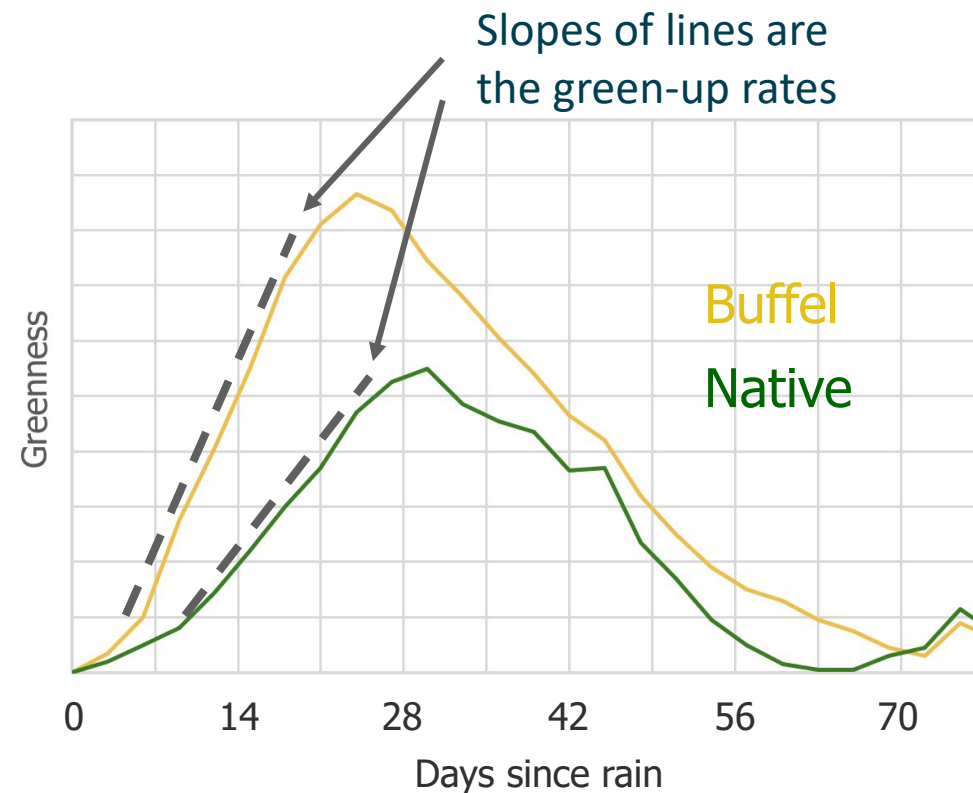
Validation of satellite-derived information

Scenario:

Buffel grass is invading Australia's arid and semi-arid landscapes. After rain, it greens up faster, and peaks higher, than surrounding native grasses.

Maybe, if we can detect a difference in the rate of post-rain green-up between Buffel and natives, we might be able to develop a method for mapping Buffel distribution.

We have access to satellite foliage cover imagery that has a resolution (cell size) of 90 m and time-step of 7 days.

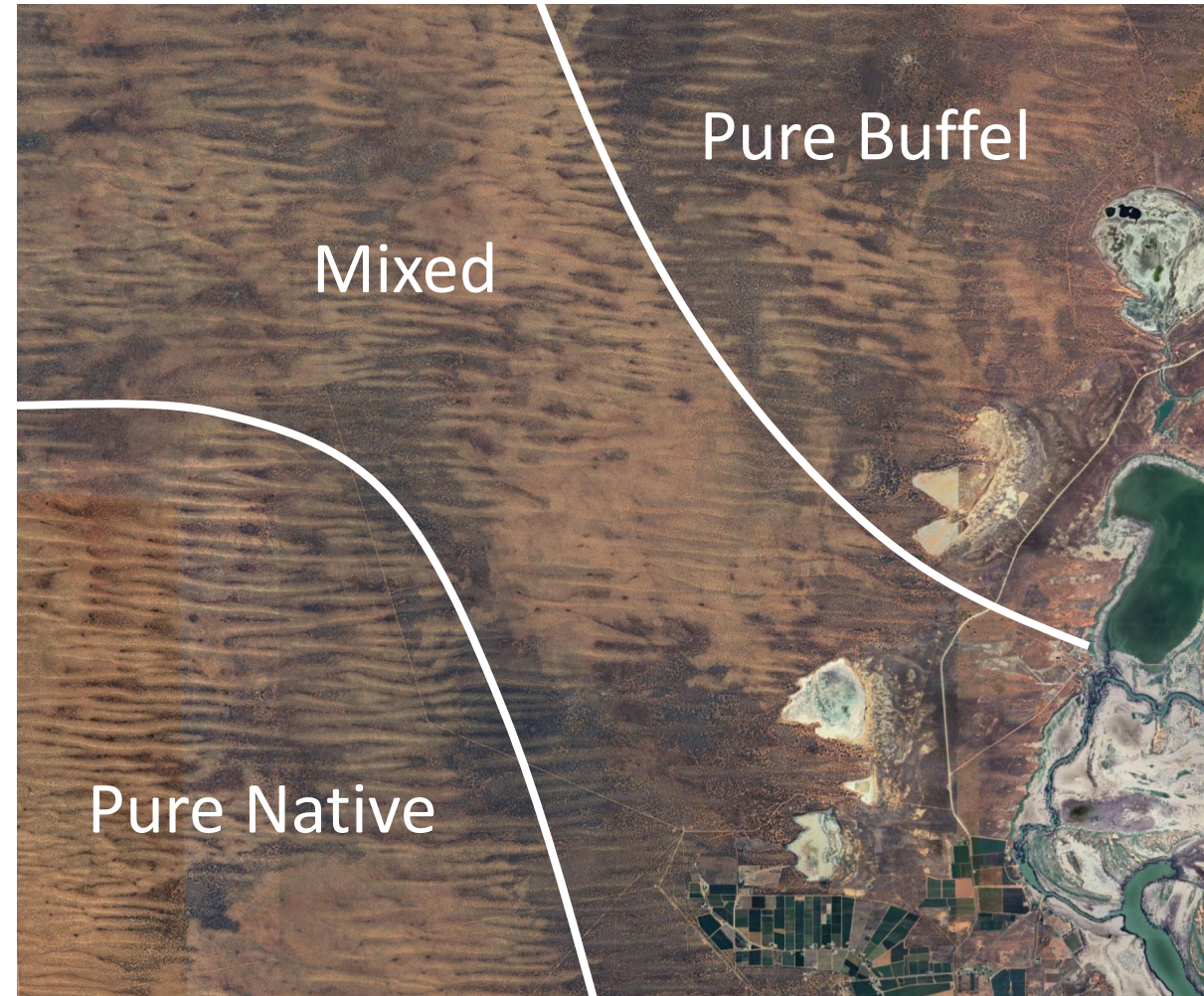


Your challenge:

Design a field campaign that will generate observations suitable for testing how well this imagery can differentiate between Buffel grass and native grasses based on rates of green-up.

1. What will you measure?
2. Where will you put your plots?
How many?
3. What sample design will you use within your plots?
4. How many samples? When? How often?

Study area



5 km

Validation of satellite-derived information

Field data curation - *every modeller's nightmare*

Can take up to 1/3 of model development time

Can make or break the defence of the model



tern
Ecosystem Research Infrastructure



esa

ECOSPHERE

Dynamics of necromass in woody Australian ecosystems

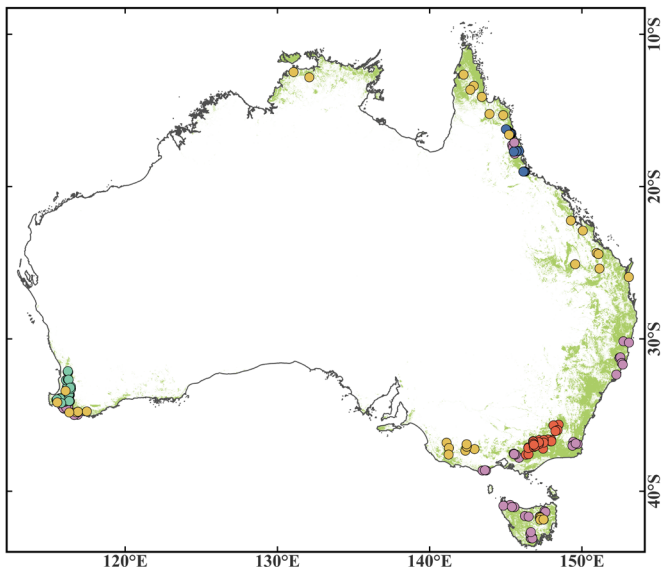
MATHIAS NEUMANN ^{1,2,†} JOHN TURNER,³ TOM LEWIS ⁴ LACHIE MCCAW,⁵ GARY COOK,⁶ AND MARK A. ADAMS ¹

Neumann & co compiled a massive database of litter mass observations from 498 published and unpublished works, yielding 5000 observations. *An incredible piece of work.*

We used these data to validate a litter fuel load model. We had to filter the observations, using only those that:

- use the same definition of litter
- have published study details
- reported location coordinates, and month-of-observation
- collected after 2001 and from within forest;
- contained at least ten sites per dataset; and
- represented a spatial footprint of quadrats at sites (that is, multiple, widely spread quadrats rather than a single quadrat).

This gave us 225 observations to work with



- Bowman-A (wet forest)
- Bowman-B (tall open forest)
- McCaw (Jarrah forest)
- Parsons (tropical rainforest)
- Prior (multiple types)
- Forest

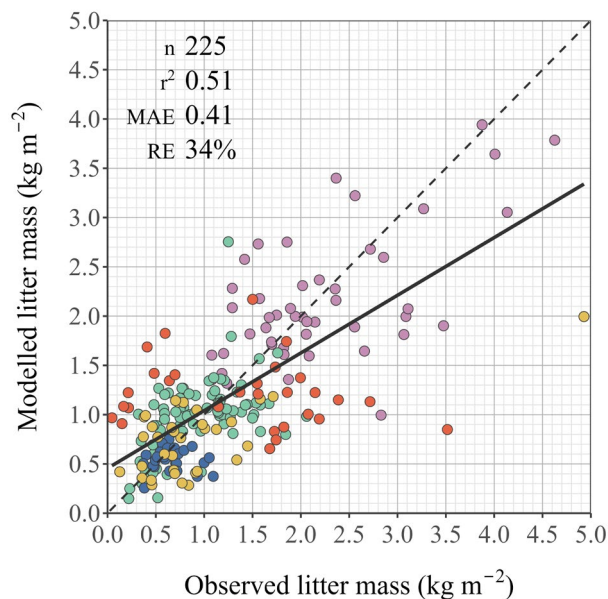
Got to get down in the weeds



Spatial arrangement and replication and size of quadrats was never the same ...and never matched the scale of the model

Can't tell if variations in model performance are tied to

- geography
- vegetation type, or
- study methods



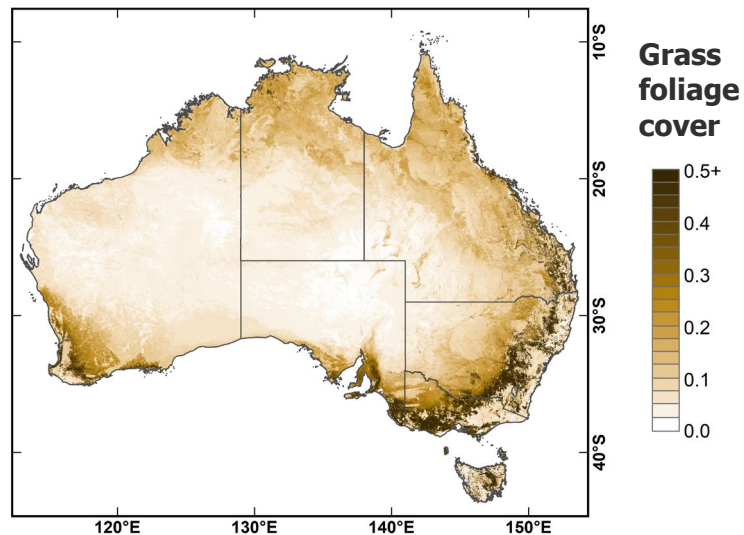
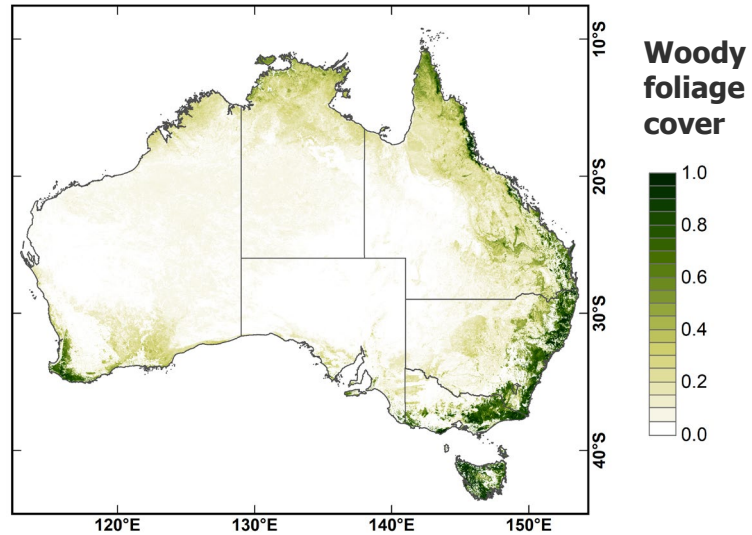
- Bowman-A
- Bowman-B
- McCaw
- Parsons
- Prior

Data sources in Neumann et al

ref	count	My comment	What is litter?
Grigg 2009, 2010	218	No reference in original spreadsheet. Relates to mine site rehab plantations and difference management effects. Not that useful.	
Prior 2017 (AU.334.)	204	Average of 3 30 m transects separated by 100 m	Grass + litter (<6 mm) CONFIRMED
McCaw 2011ab (AU.274.)	116	Jarrah forests. Samples taken every 10 m along two 100 m transects from 1x1 m2 plots.	Leaf litter and FWD (<25 mm)
DAF unpublished	109		
Lewis 2012a	108	Pine plantations, SE Qld	
Bowman 2017, Wood 2014 (AU.039.)	48	TERN Ausplots forest network	Leaves + FWD (<6 mm) + grass CONFIRMED
Bennett 2017	42	Plots 30 x 30 cm, running along a 25 m transect. Only one location given for all sites	
Dalgleish 2015, Knuckey 2016	40	20 m transects	
Selkirk 2017	40	Thesis - couldn't find	
Whitford 2003	37	Internal report	
Cunningham 2012, 2015	36	30 x 30 m plots in farm plantations.	
Keith 2014	31	Marysville post-fire biomass inventory. Litter depth taken at 25 points in a 1 ha site. Calibrated to three sub-plots where litter mass was measured Only one location given for all sites	All material < 25 mm
Prior 2017, Bowman 2014b	30	No different from prior above?? There is no Bowman 2014b	
Gilroy 2009, 2005 (AU.151.)	29	SE Qld. Five 25x25 cm plots "at each site". No info on how these are spread out.	< 6 mm all material, including standing live veg up to 50 cm high
Parsons 2014 a,b	29	FNQ trop rainforest. Plots 5 x 5 m at 200 m intervals along a 1 km transect. Too far apart to represent a 250 m cell No month of collection given. Litter are the averages collected over a 24 month period (May 2007 - Aug 2009).	All litter < 2 cm
Lewis 2020	28	Private forestry	
Wentzel 2010, 2018 (AU.439)	24	24 100 x 100 m sites. Five 30x30 cm plots at each. No month of collection given, but these are not recently burned and have low litter mass (and so seasonal variability prob low)	All litter but diam not specified
Cook 2012 (AU.094.)	23	Transects located to be representative of large blocks. Each plot had 6 sites. Each site had three 1 x 1 m quadrats. These are rehabilitated mine sites and have "substantially greater [loads] than those of the surrounding unmined landscape". More typical values are 6	Not defined
Lewis 2012b	21	Private forestry	
Lewis 2016	20	Private forestry	

An assessment of the accuracy of satellite-derived woody and grass foliage cover estimates for Australia

Randall J. Donohue^{A,*} and Luigi J. Renzullo^B



Euc tall open forest (TAS)

Temperate rainforest (NSW)

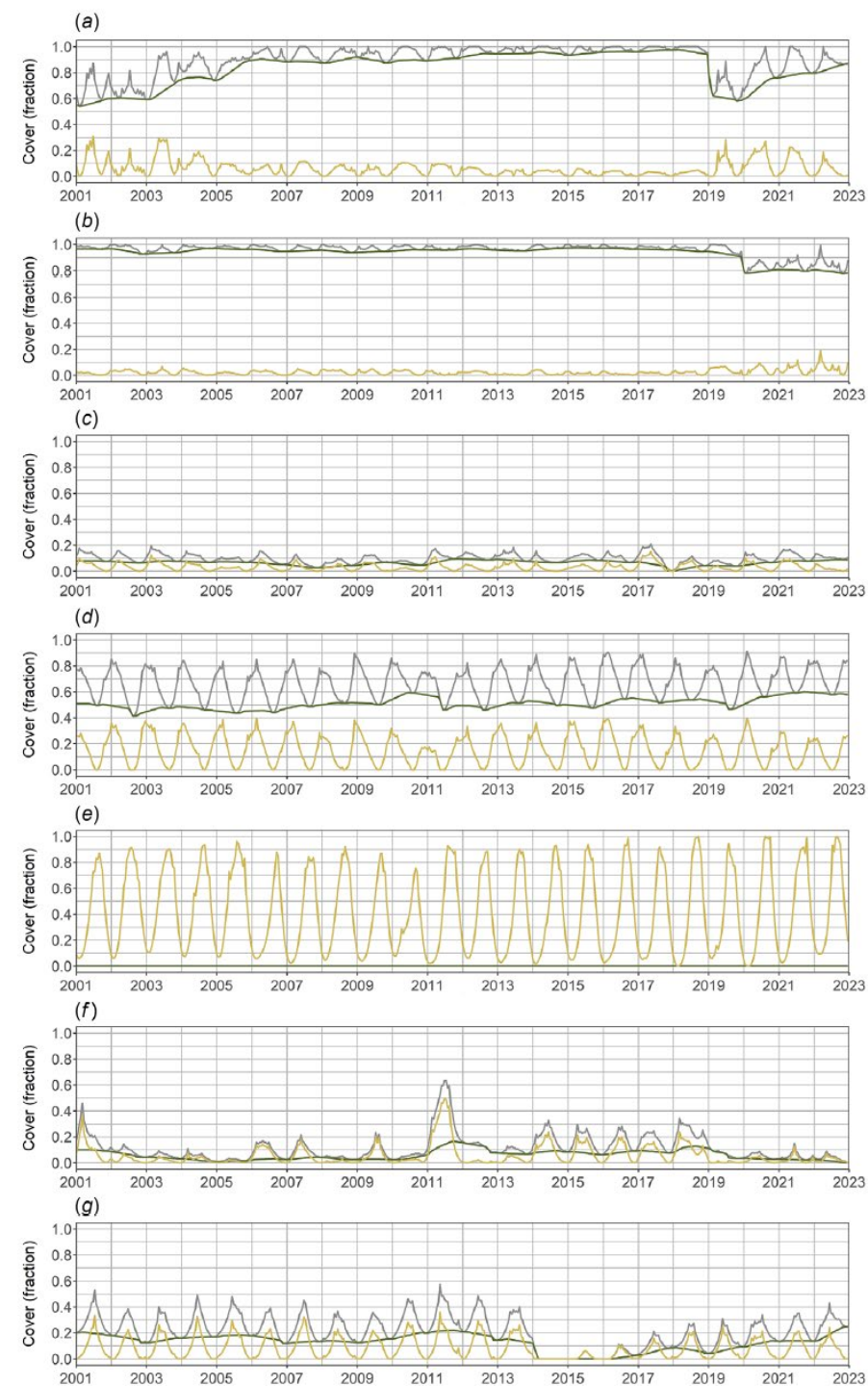
Hummock grassland (WA)

Savannah woodland (NT)

Winter cereal crop (WA)

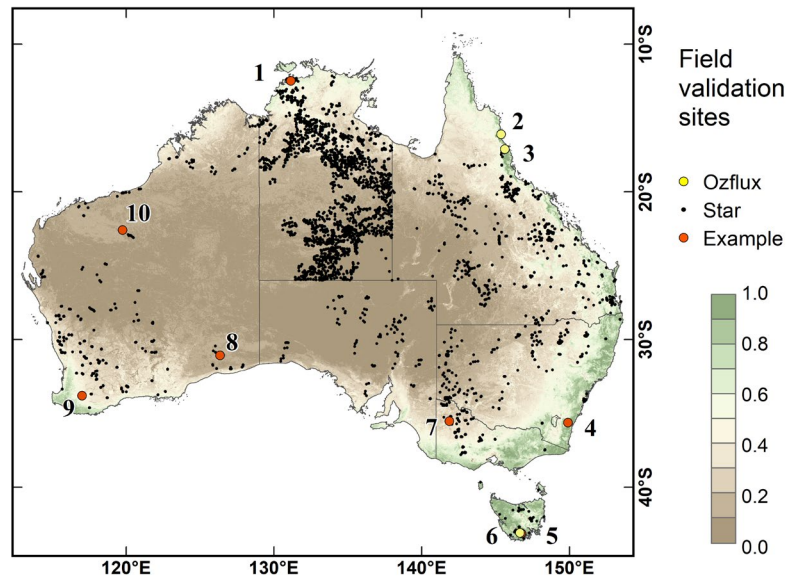
Acacia woodland (WA)

Mallee woodland (VIC)

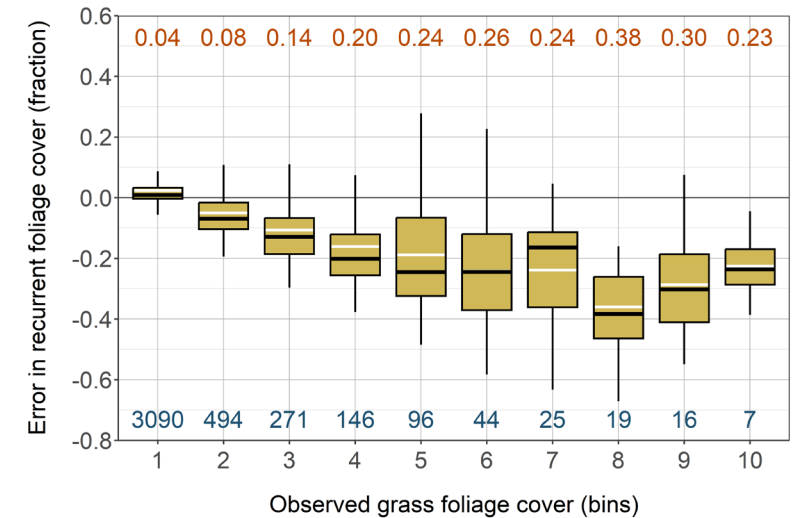
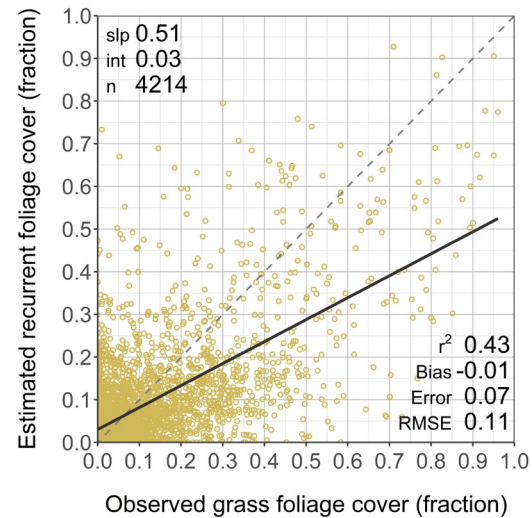
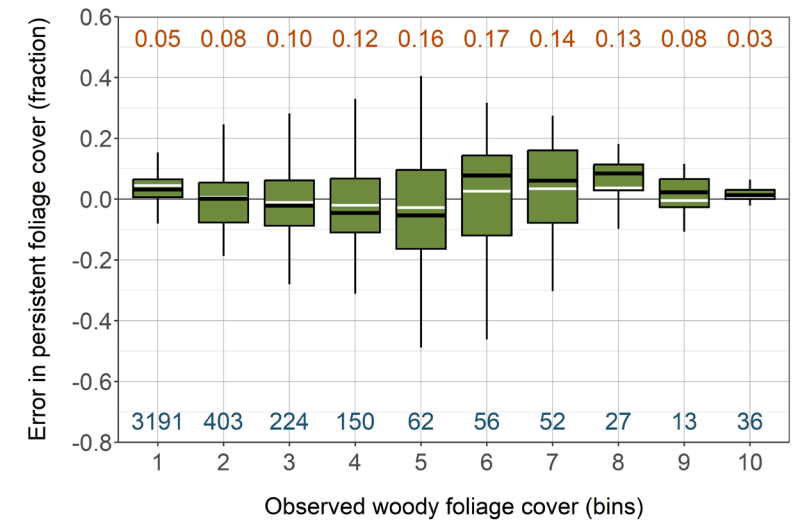
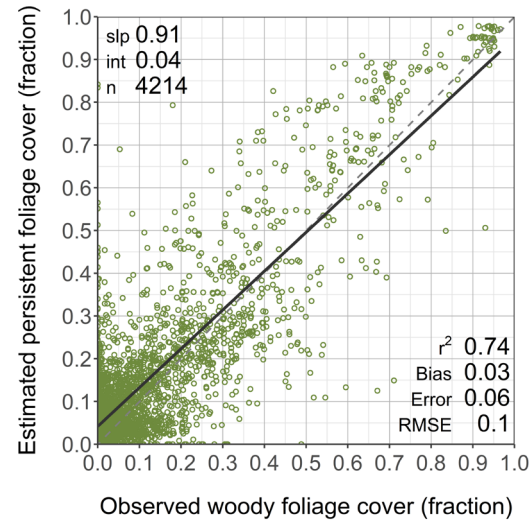


Validated with Australia's ground cover database

SLATS Stars ground cover data
4214 observations at 3559 sites
A single, standard field methodology (hooray!)

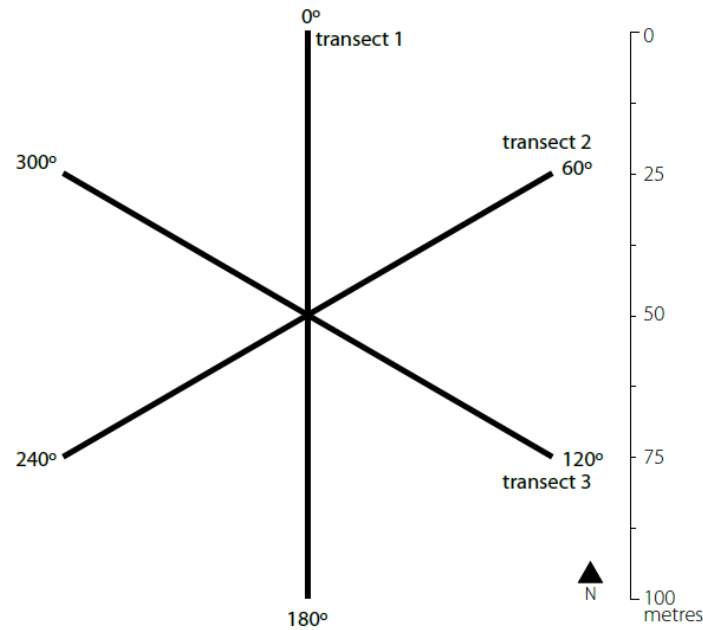


But what's wrong here?



Ground cover measurements

SLATS star plot layout designed by Qld government
for validating fine-scale satellite imagery



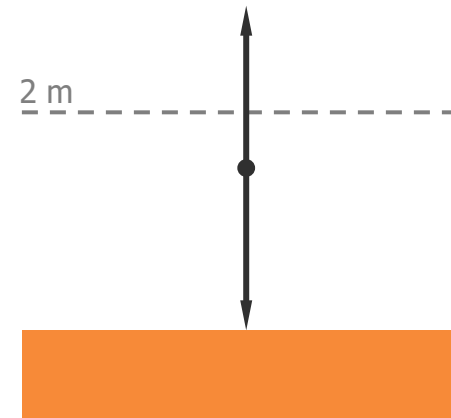
Point-intercept measurements made every 1 m
along each transect
Excellent for generating spatially representative
data

Field measurement of fractional ground cover

A technical handbook supporting
ground cover monitoring for Australia

J Muir, M Schmidt, D Tindall,
R Trevithick, P Scarth and JB Stewart

Point-intercept method



Woody vegetation > 2 m

Green leaf

Dry leaf

Branch

Within-canopy clear-sky

Woody vegetation < 2 m

Green leaf

Dry leaf

Branch

Ground cover

Green leaf

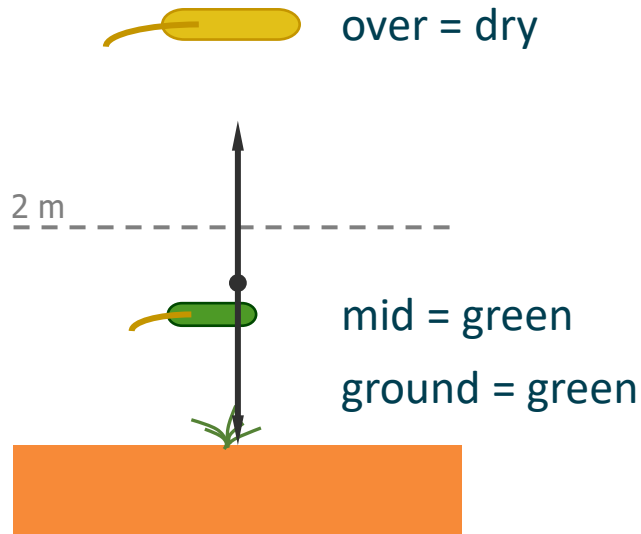
Dead leaf

Other litter

Soil/rock/crust/crypto/disturbed

How are these measurements to be interpreted?

As recorded in field notes

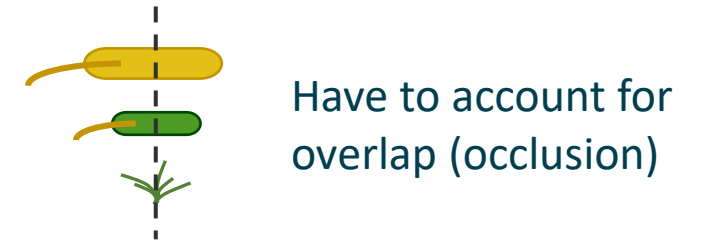


As summarised per site

Stratum	Type	Cover (%)
Overstorey	crown	4
	green	1.33
	dry	0
	branch	0.67
Midstorey	-	94
	green	8.67
	dry	4
	branch	2.33
Ground	-	85
	crust	36.49
	dist	5.74
	rock	12.5
	green	1.01
	dead	0
	litter	20.27
	crypto	23.99

Summarised from
3 x 300 measurements

But what would a satellite see?



Can't do this from the site summaries so
must go back to raw field data and
reprocess them

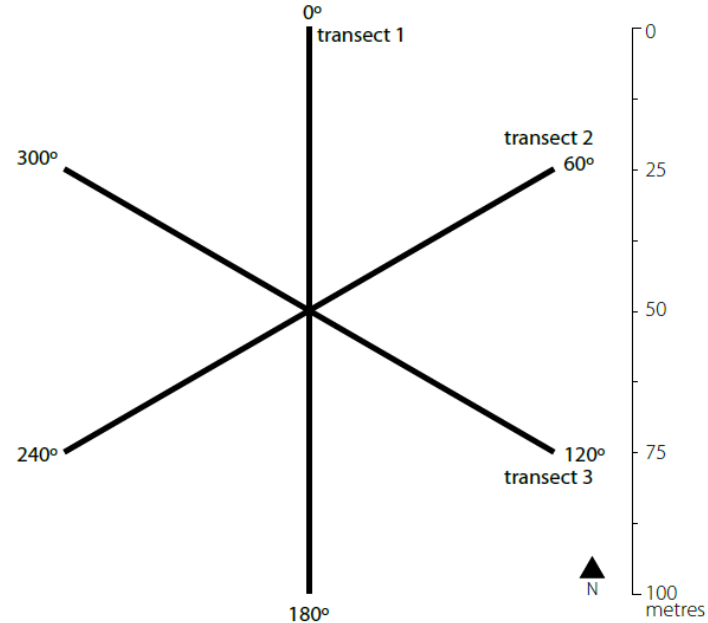
This is a pain cause the raw data aren't
published

Modellers need to understand *exactly*
what the data are describing

Field operators must *comprehensively*
document data and methods

Ground cover measurements

SLATS star plot layout designed by Qld government
for validating fine-scale satellite imagery

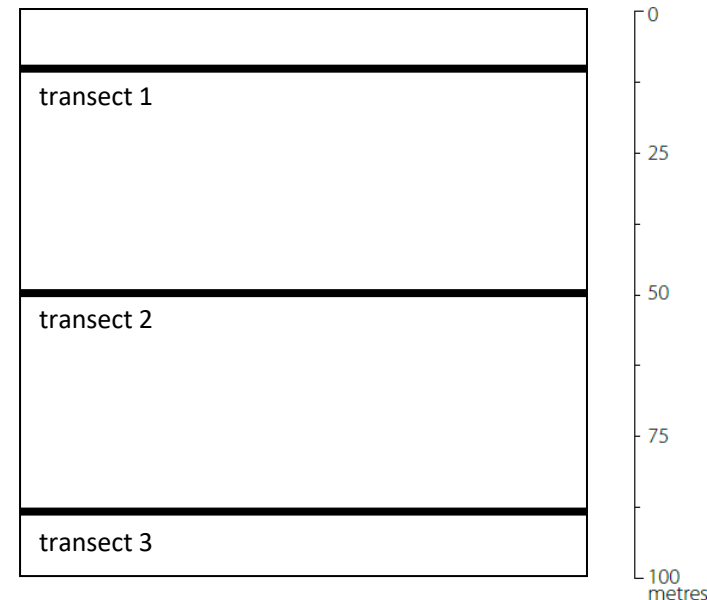
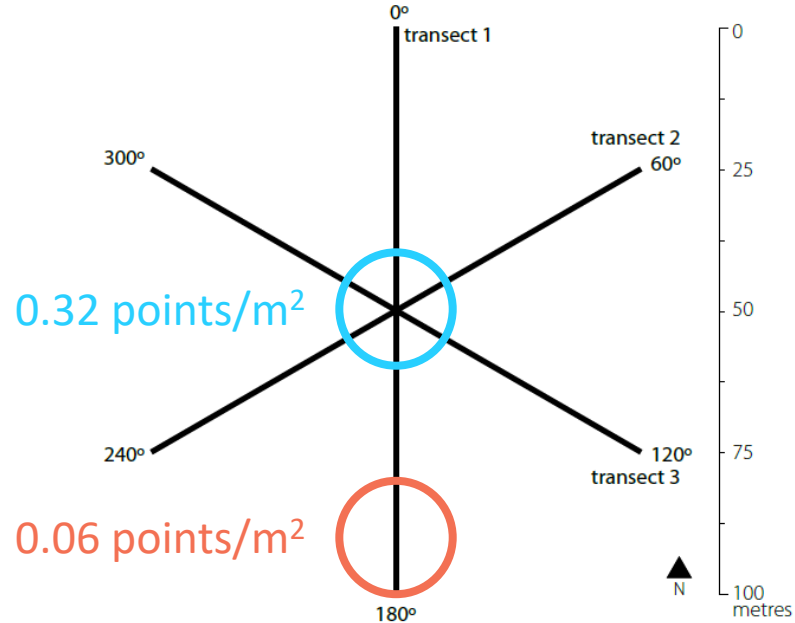


What are the limitations with this design?

Ground cover measurements

The density of observations is not uniform

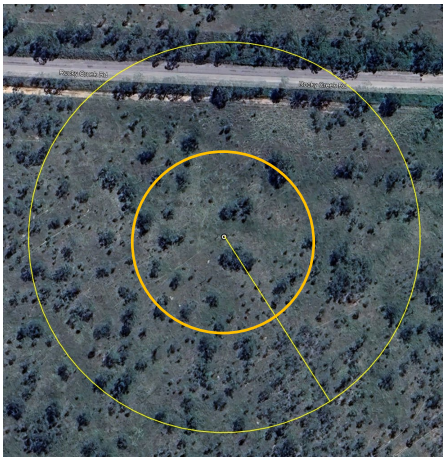
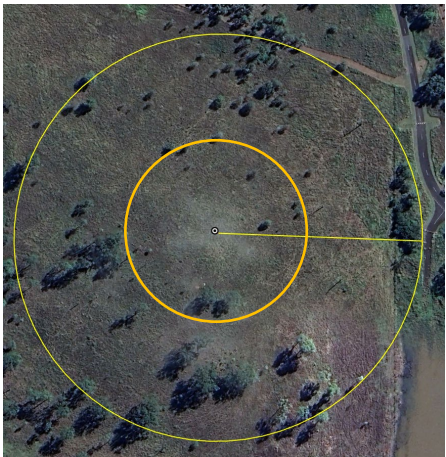
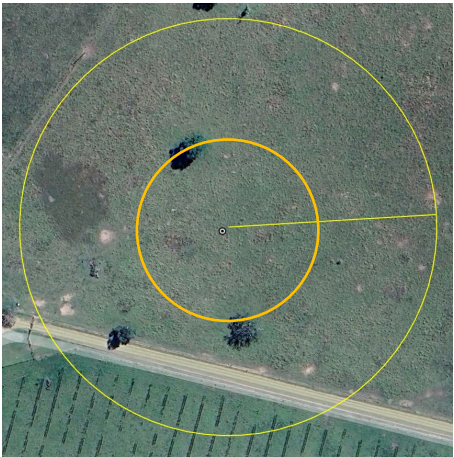
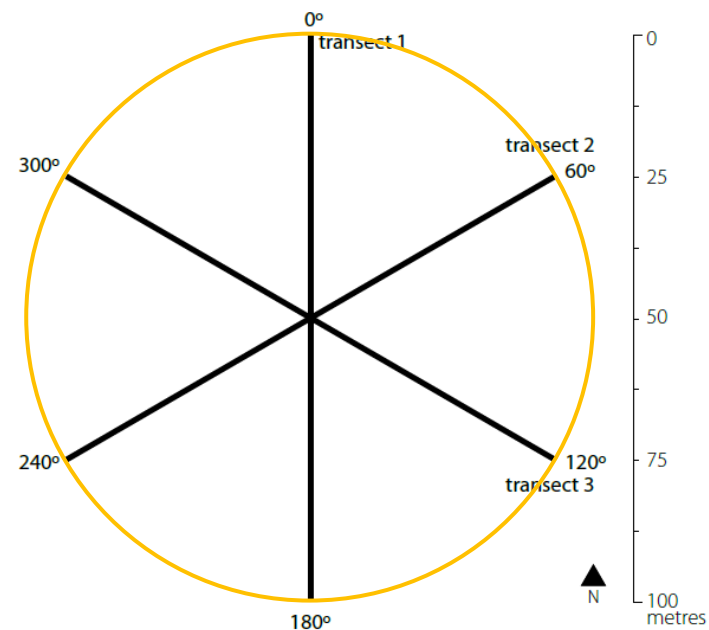
- this isn't a huge problem, especially if the area is homogenous
- be careful where you centre the plot



Is this any better?

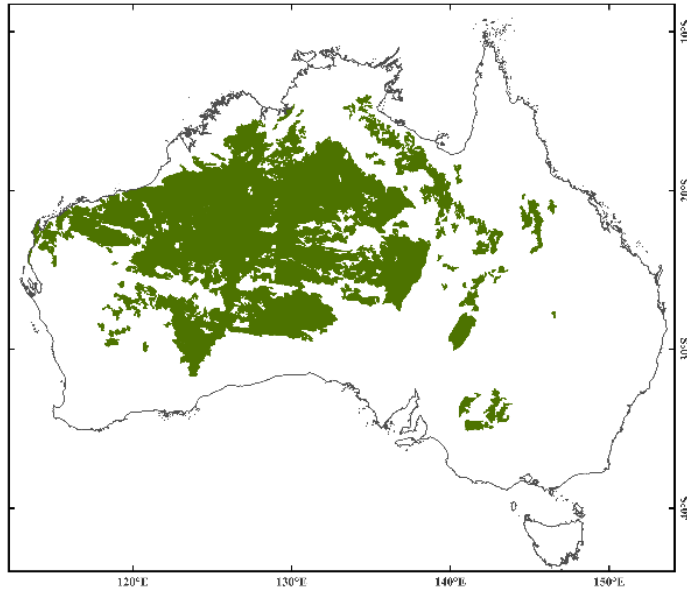
Pure grassland sites according to the field observations...

It was designed for
validating fine resolution
(30 m) imagery



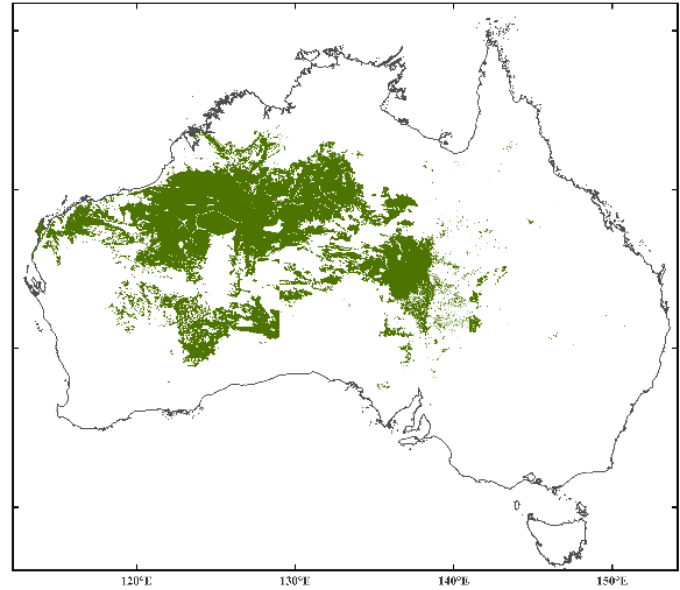
Status of national *Spinifex* grassland mapping

Carnahan 1990



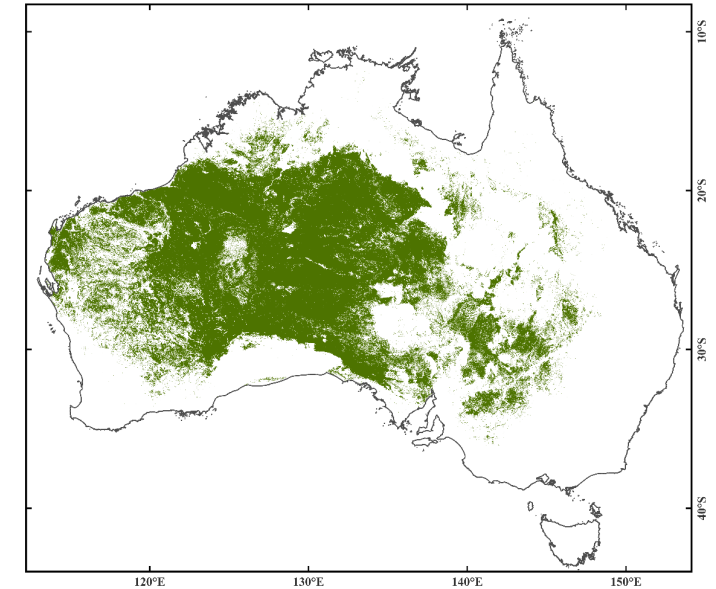
2.23 million km²

National Vegetation
Information System 2024



1.37 million km²

Spikey - a new satellite-
based map

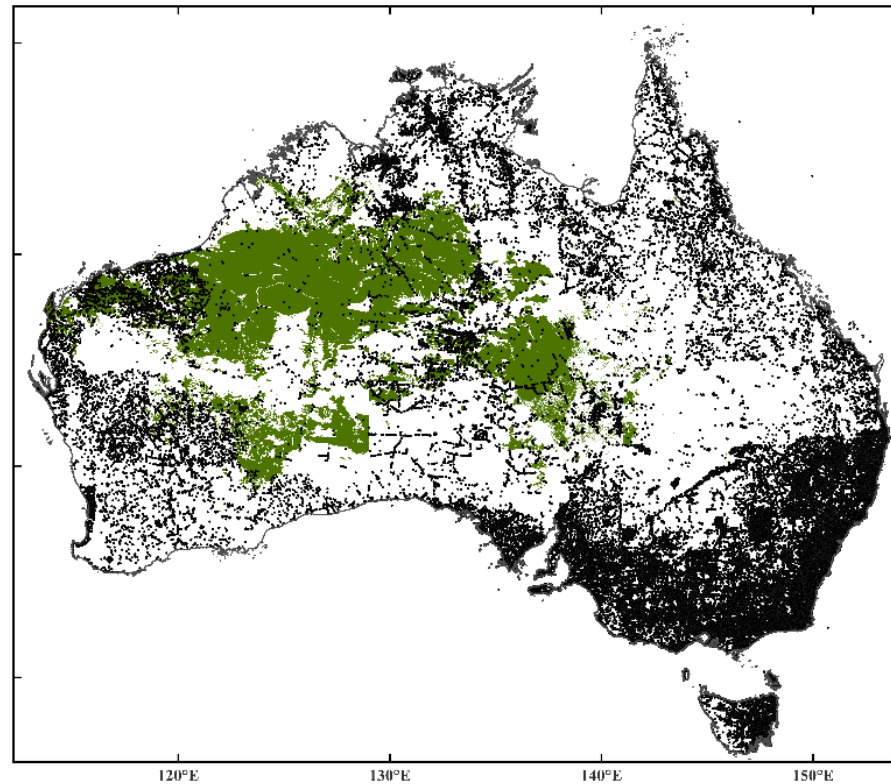


2.64 million km²

Which is correct?

Observations available for Spinifex validation

Available
observations



212,363 field
observations

Atlas of Living
Australia
+
Harmonised
Australian
Vegetation
Plot dataset
+
Ausplot
Rangelands

Issues with Spinifex observations

Compilation of data from multiple sources

- diverse methods, meanings and quality

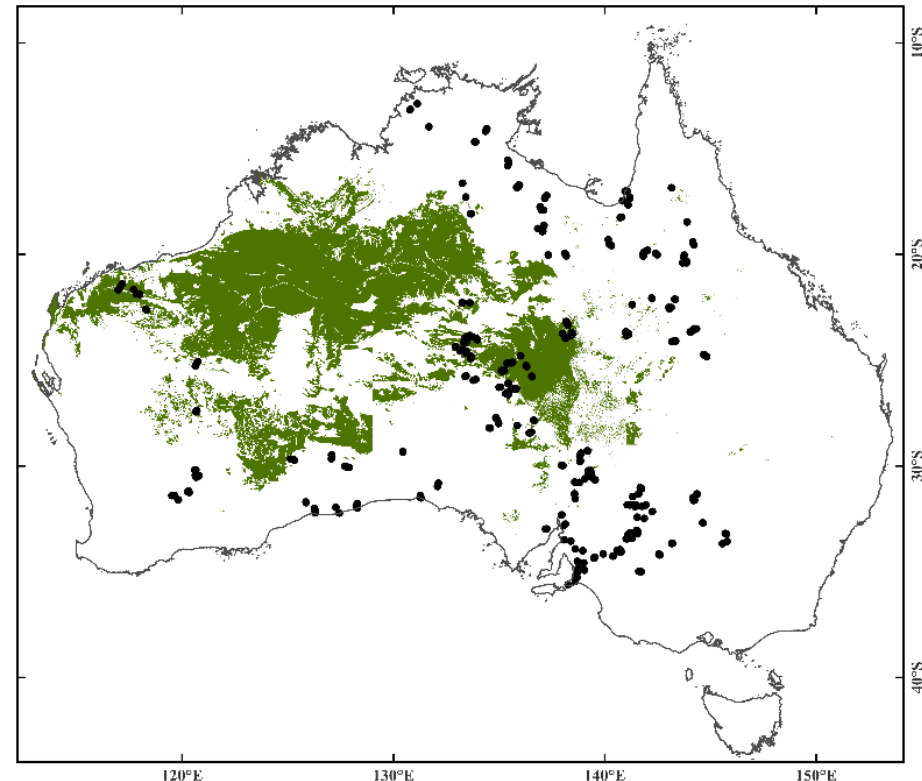
Presence only data

- easy to show your map is 100% accurate!

Poor spatial representativeness of observations

- remote areas are always data-sparse

Usable observations



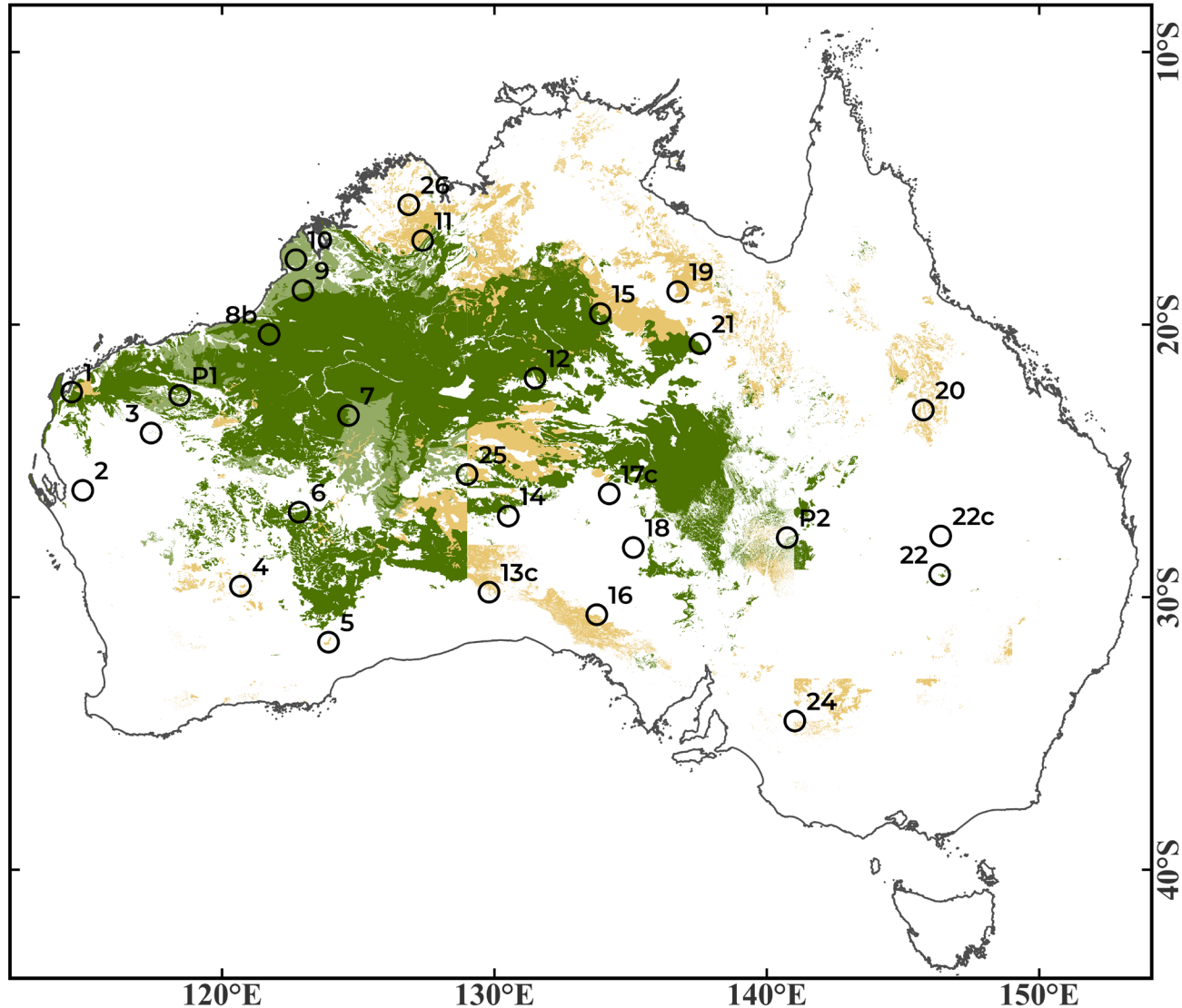
Known (and consistent) methods

Abundance not presence-only

Reliable representation of Spinifex

(these are the Ausplot Rangeland sites!)

So how can we build and validate Spinifex grassland maps??



We are testing whether it is possible to use super-high-resolution satellite imagery to generate **pseudo-observations**

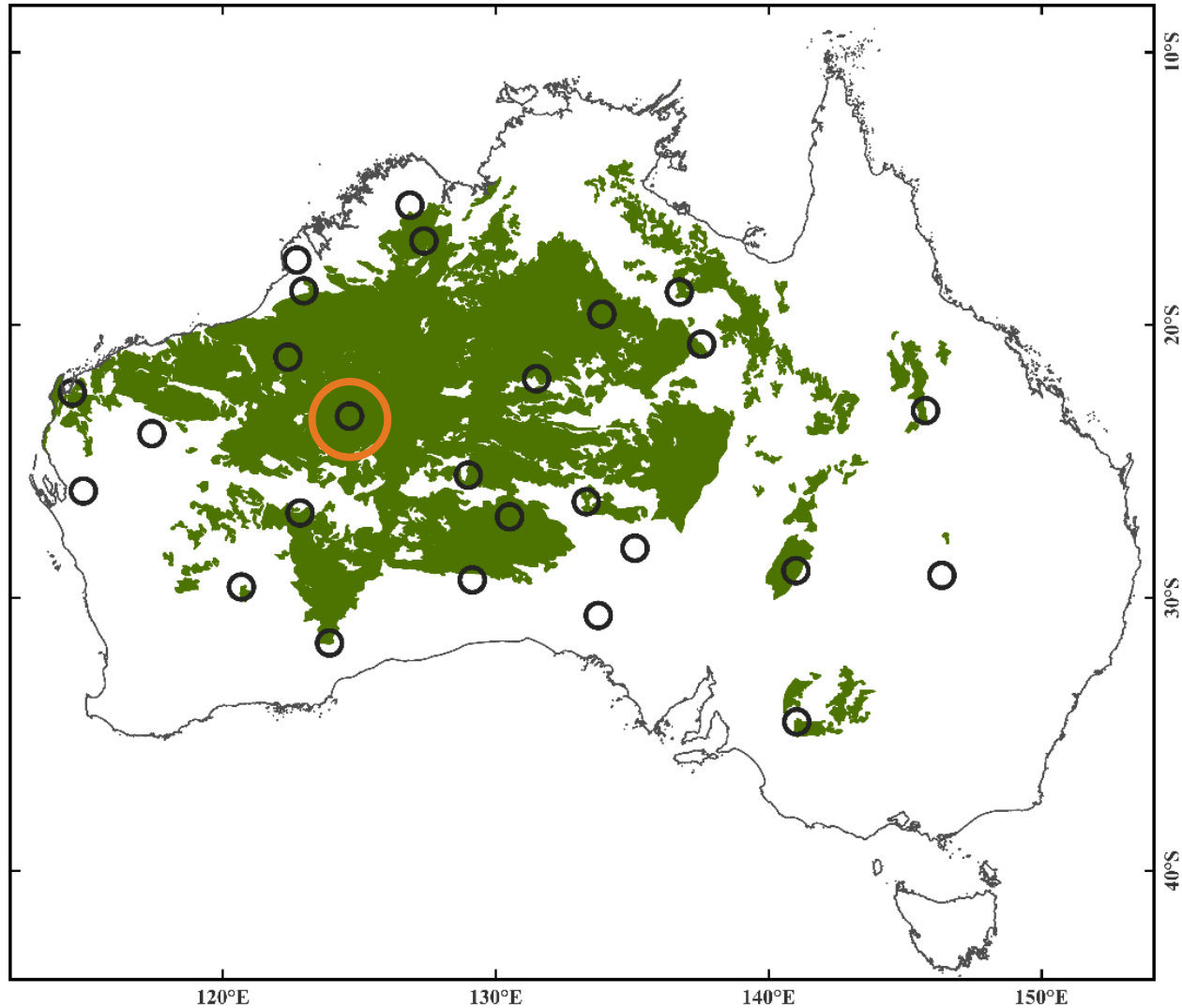
5 x 5 km ultra-high-resolution images at 27 locations

Red, green, blue and near-infrared

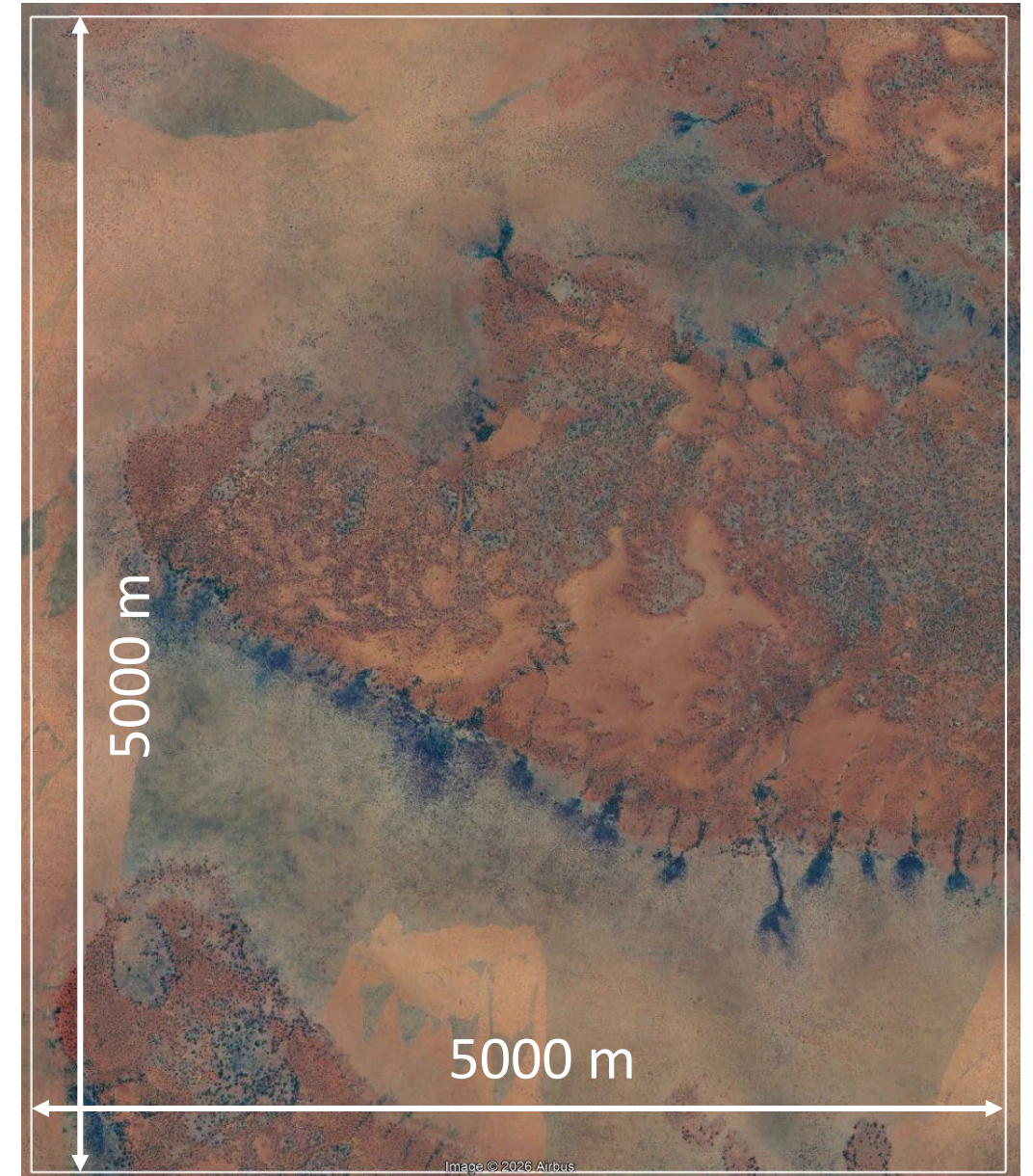
15 cm panchromatic (grey-scale) and 30-60 cm multi-spectral (R,G,B,NIR)

Dates chosen to be at the end of dry periods when grasses are browned off

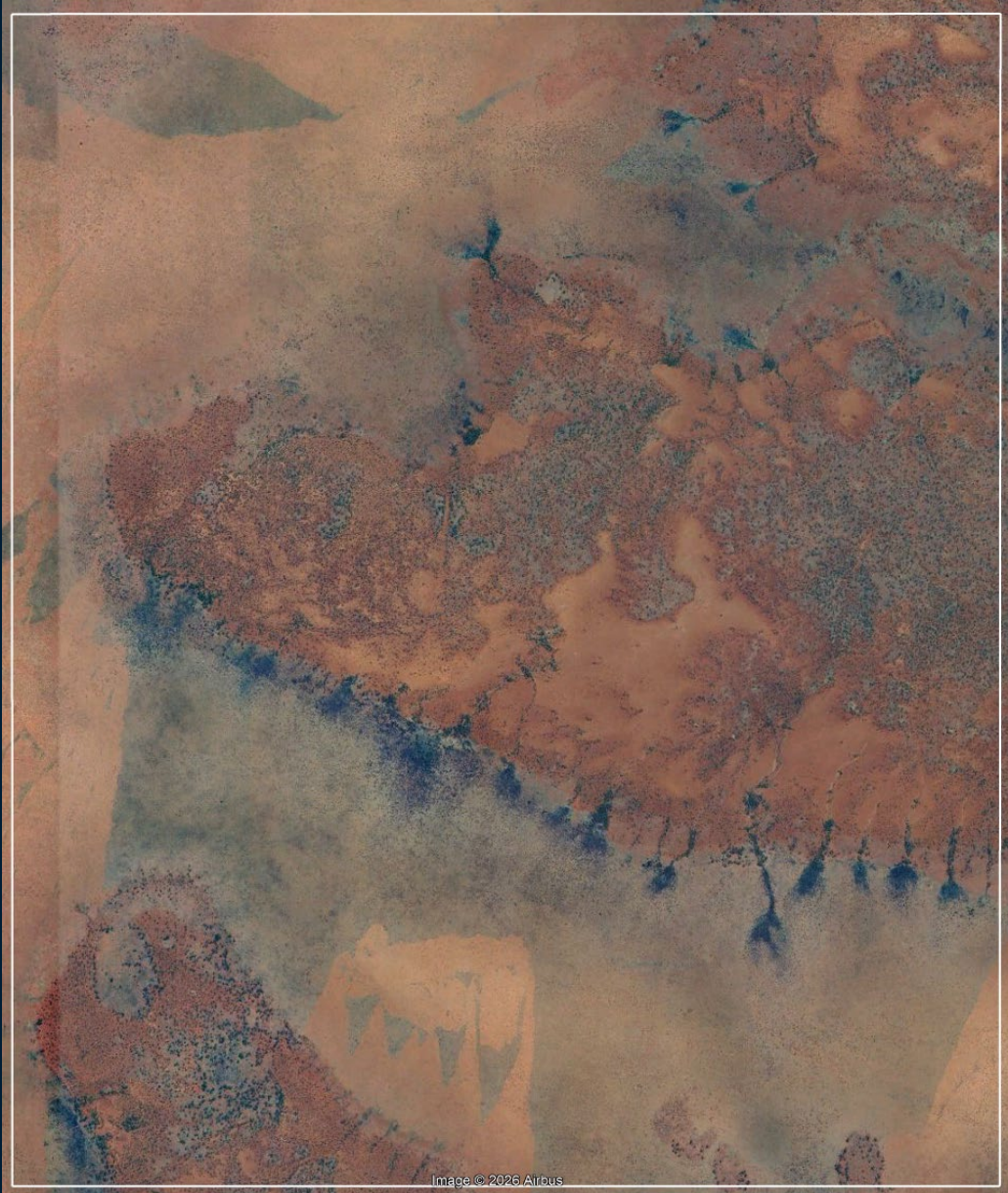
An example from the edge of the Gibson Desert



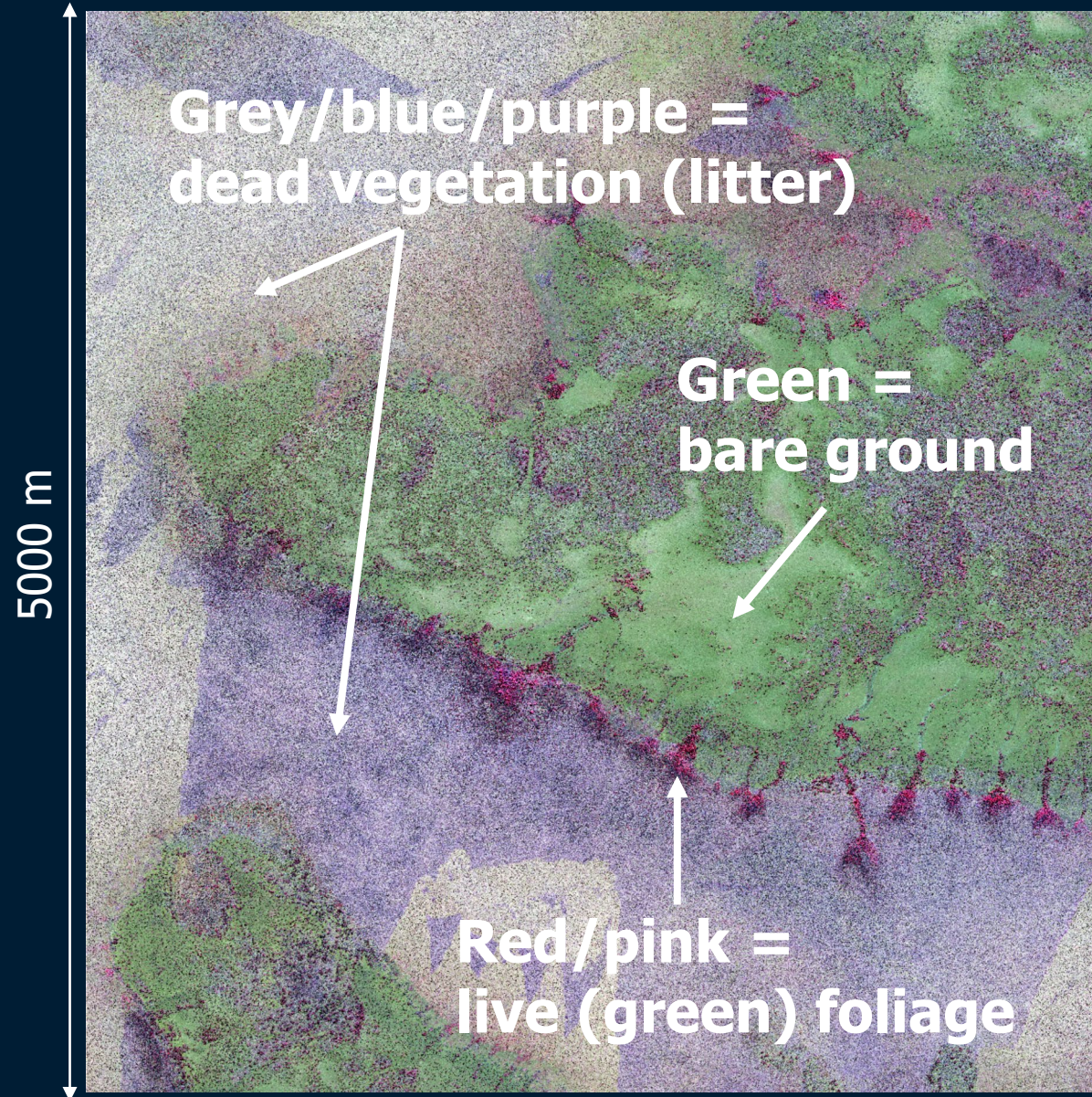
True-colour composite (R G B)



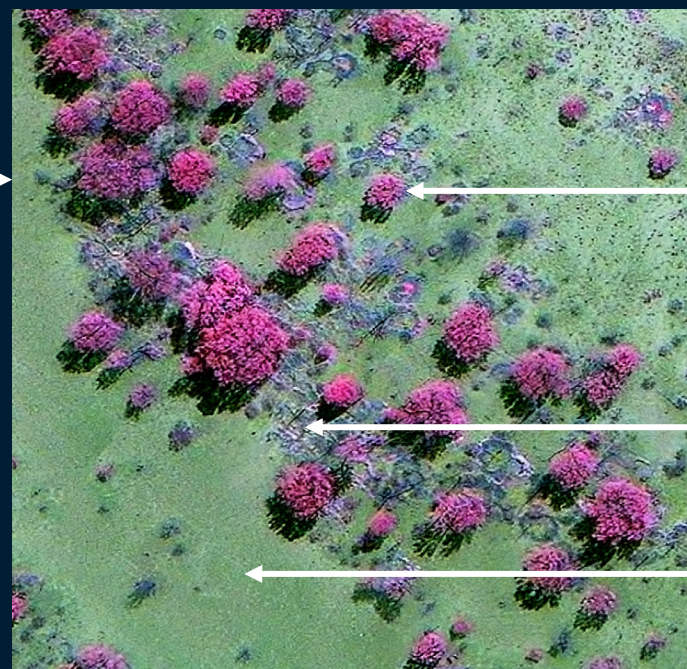
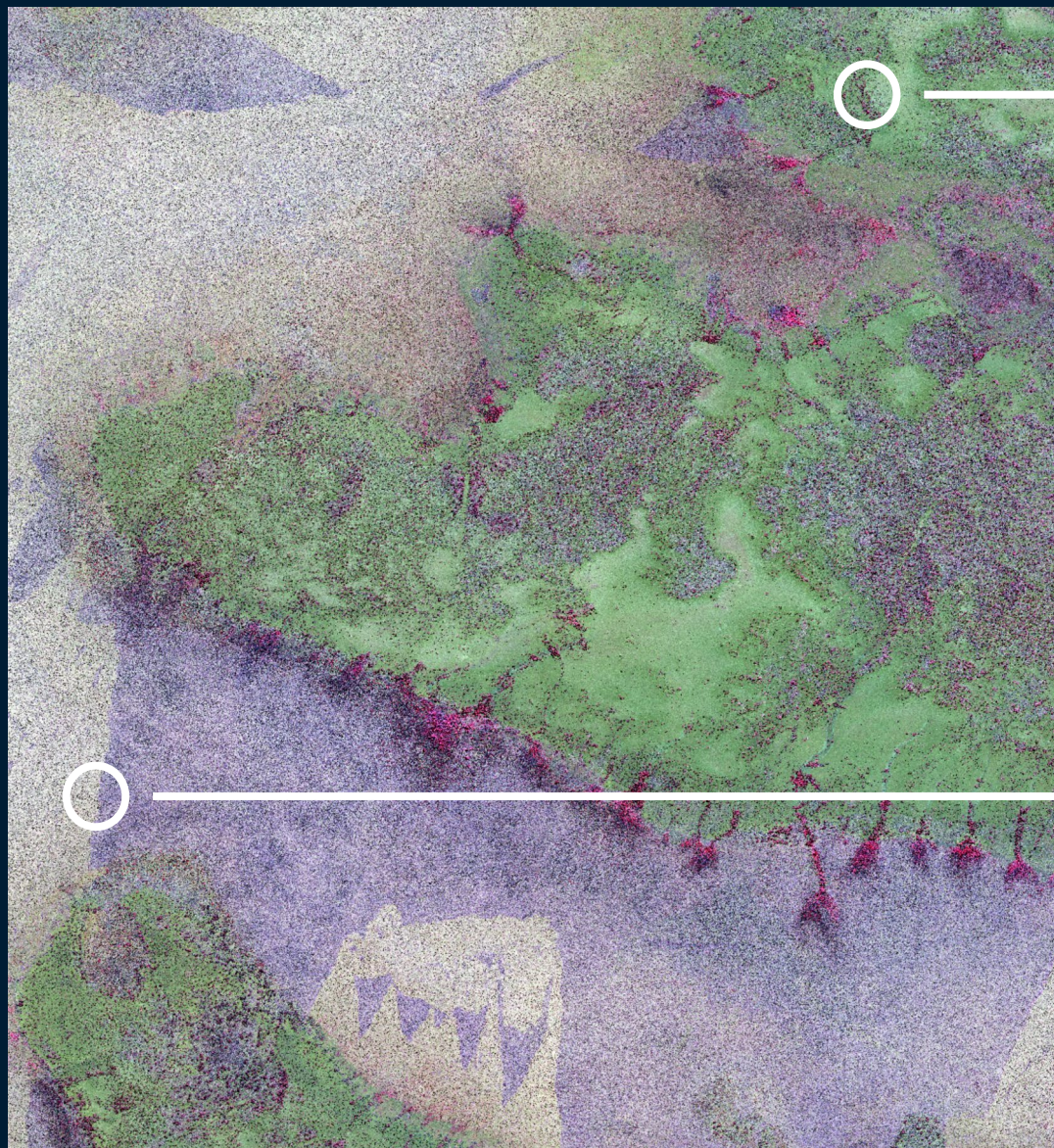
True-colour composite (R G B)



False-colour composite (NIR R G)



5000 m

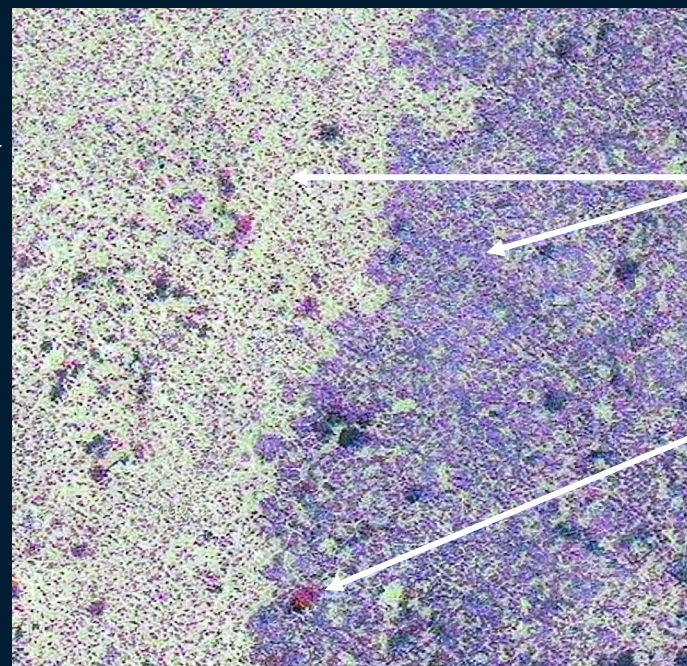


Live foliage

Litter

Bare ground

70 m



Litter

Live foliage

What does Spinifex look like?

Shape – old spinifex is (usually) easy to identify



45 m



55 m

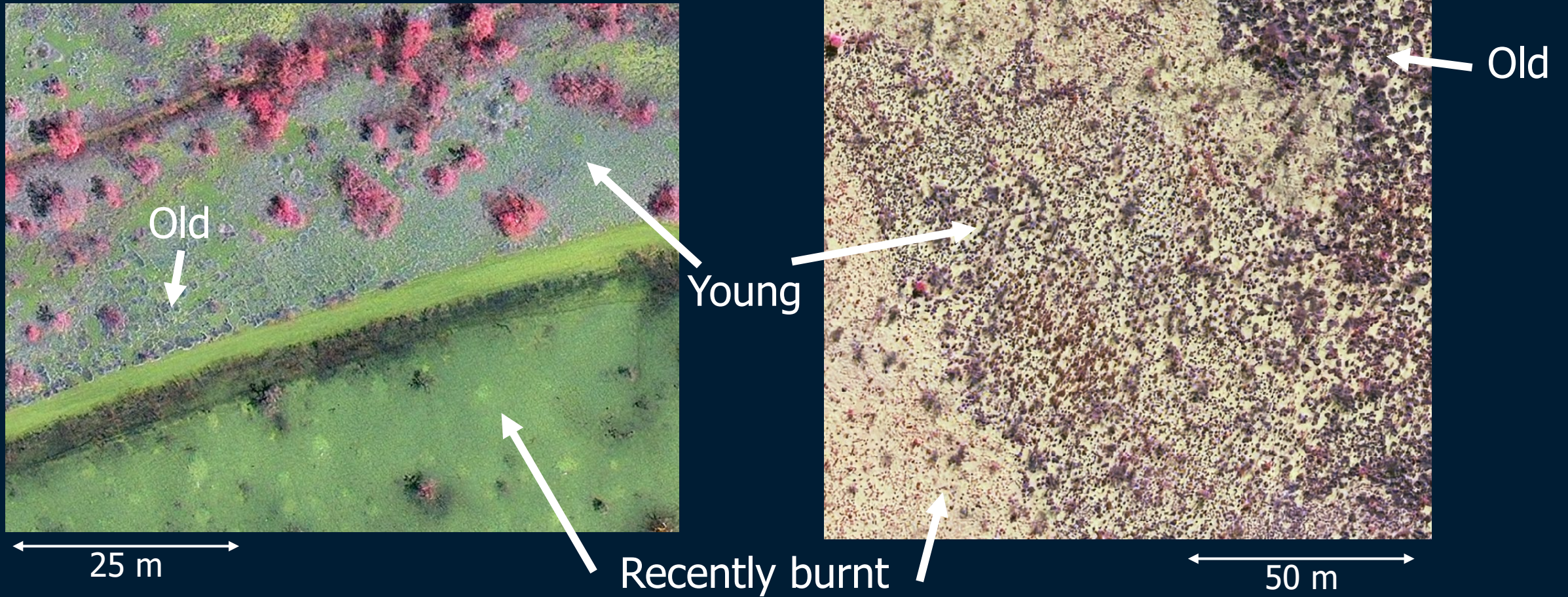


40 m



What does Spinifex look like?

Shape – young spinifex is harder to identify



What does Spinifex look like?

Shape – some regions don't get rings...



50 m



15 m



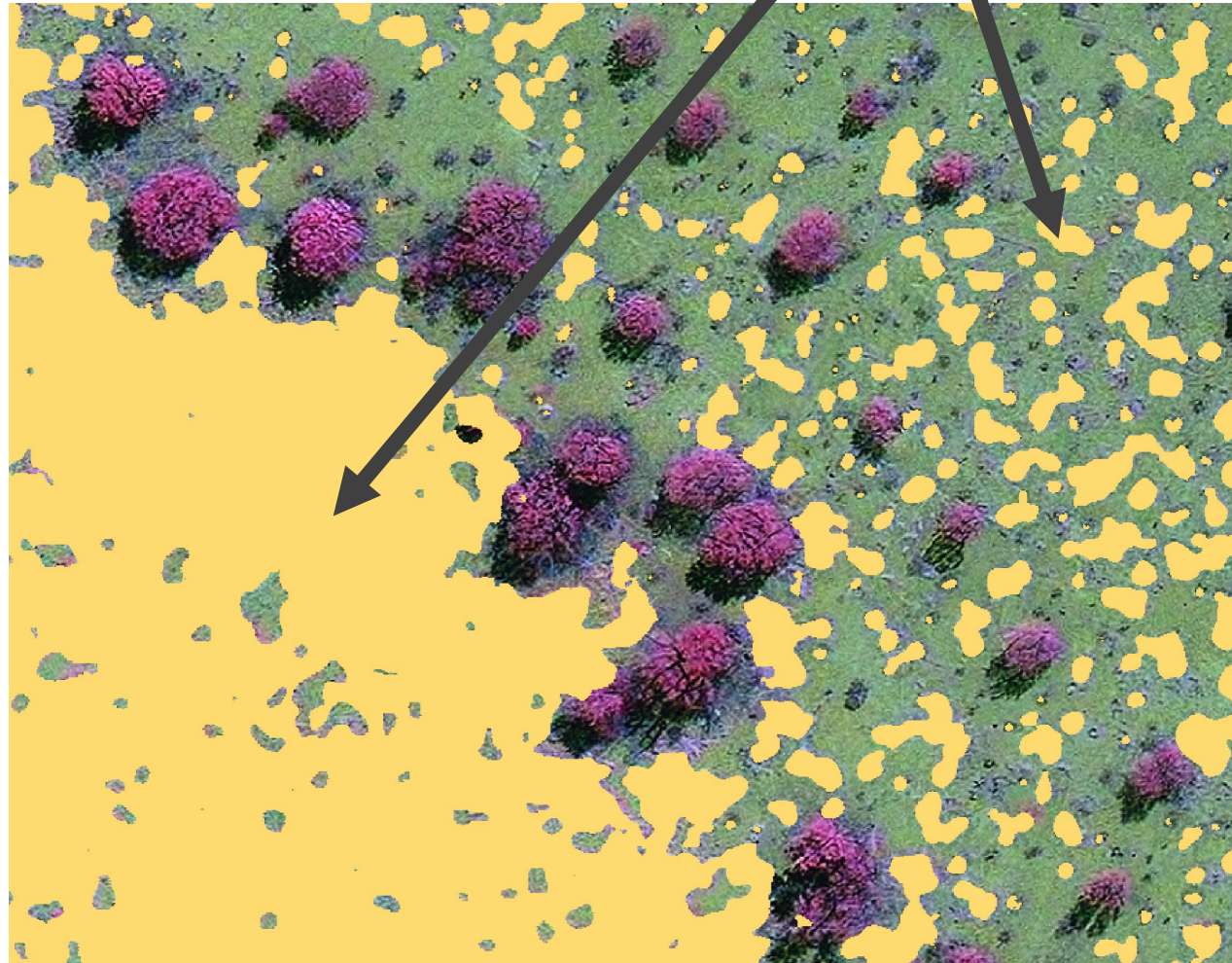
...and so context and expert opinion are needed

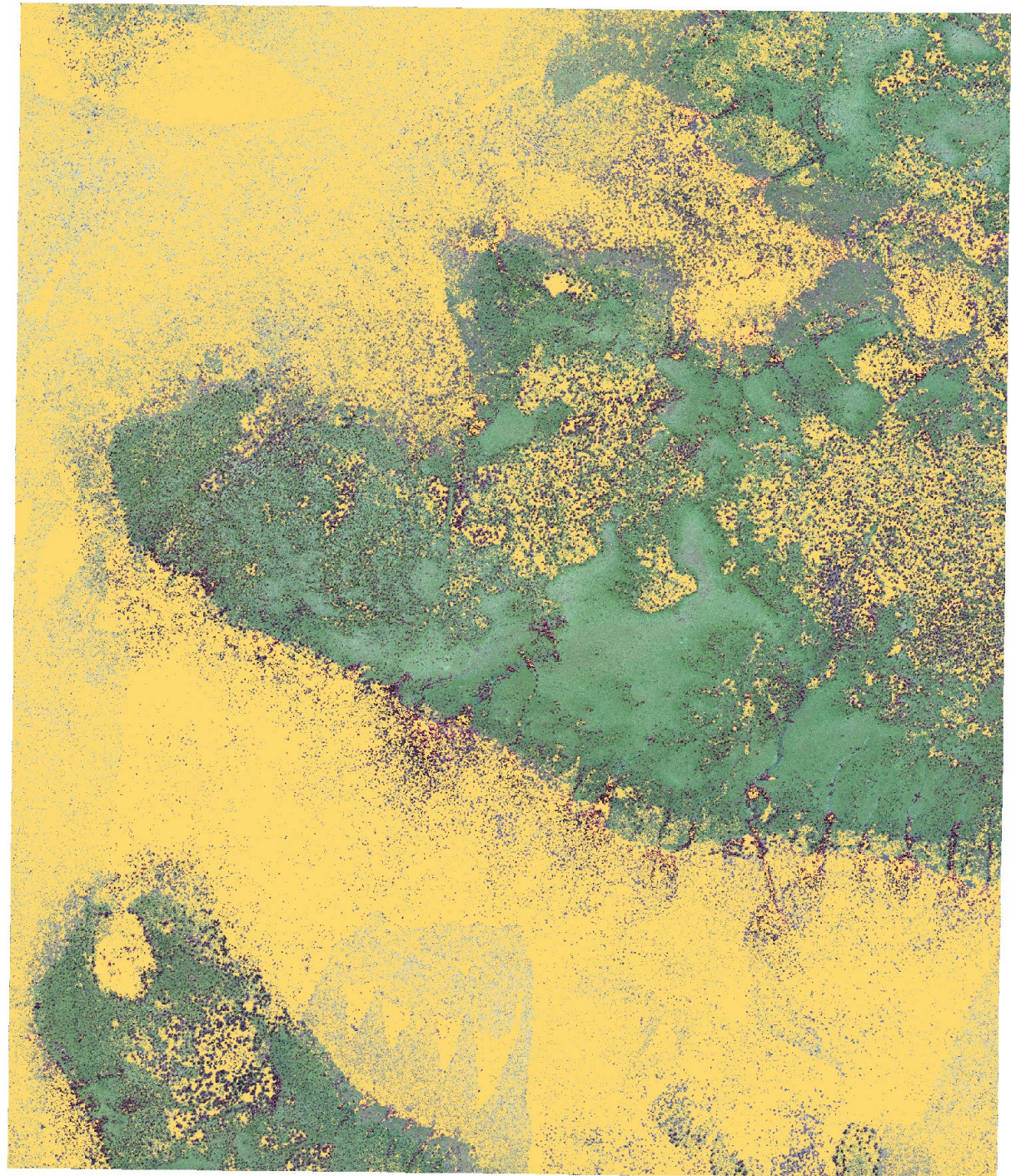
Classifying Spinifex hummocks

Colour and shape/texture-based
image classification

**Spinifex
hummocks**

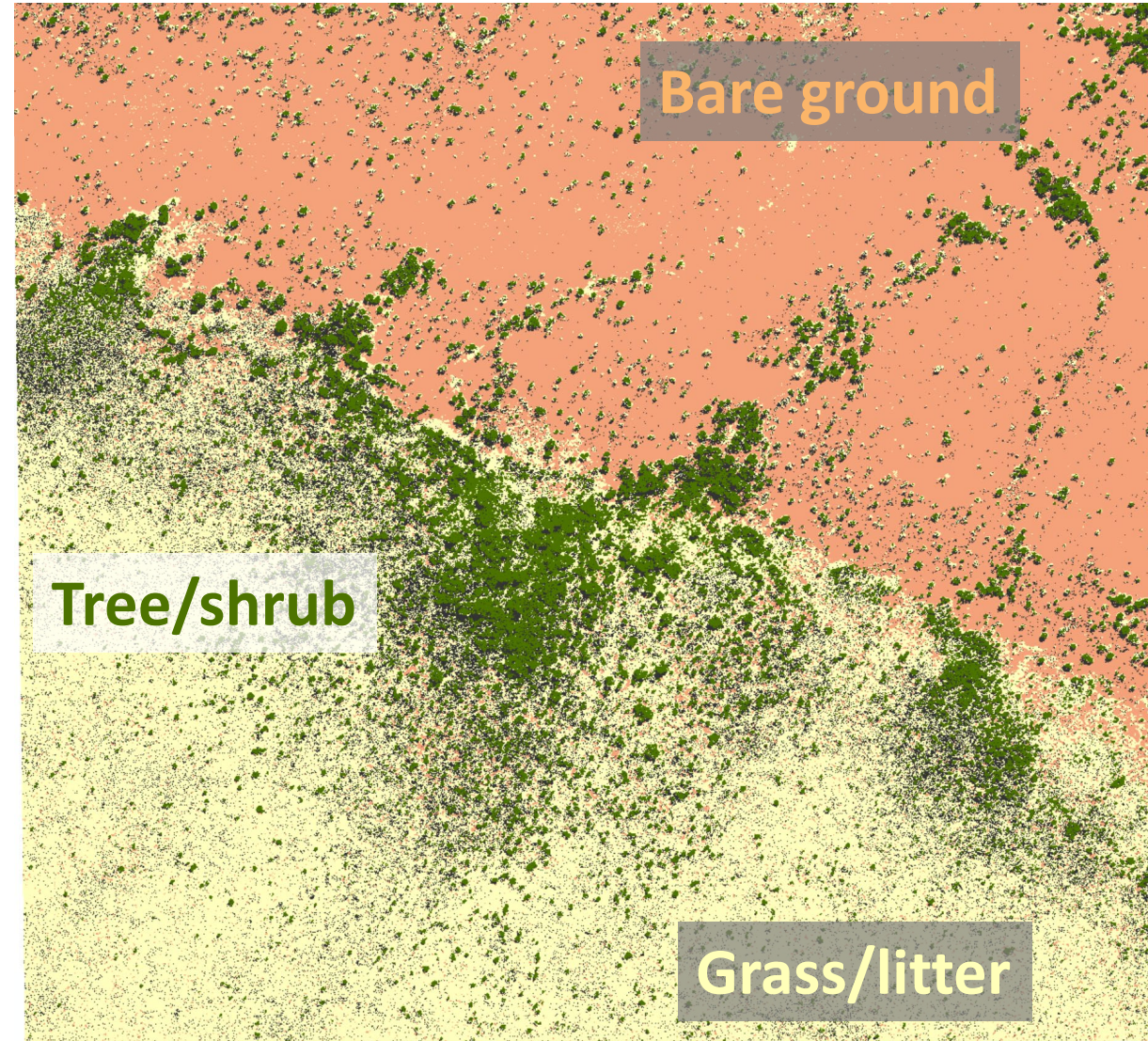
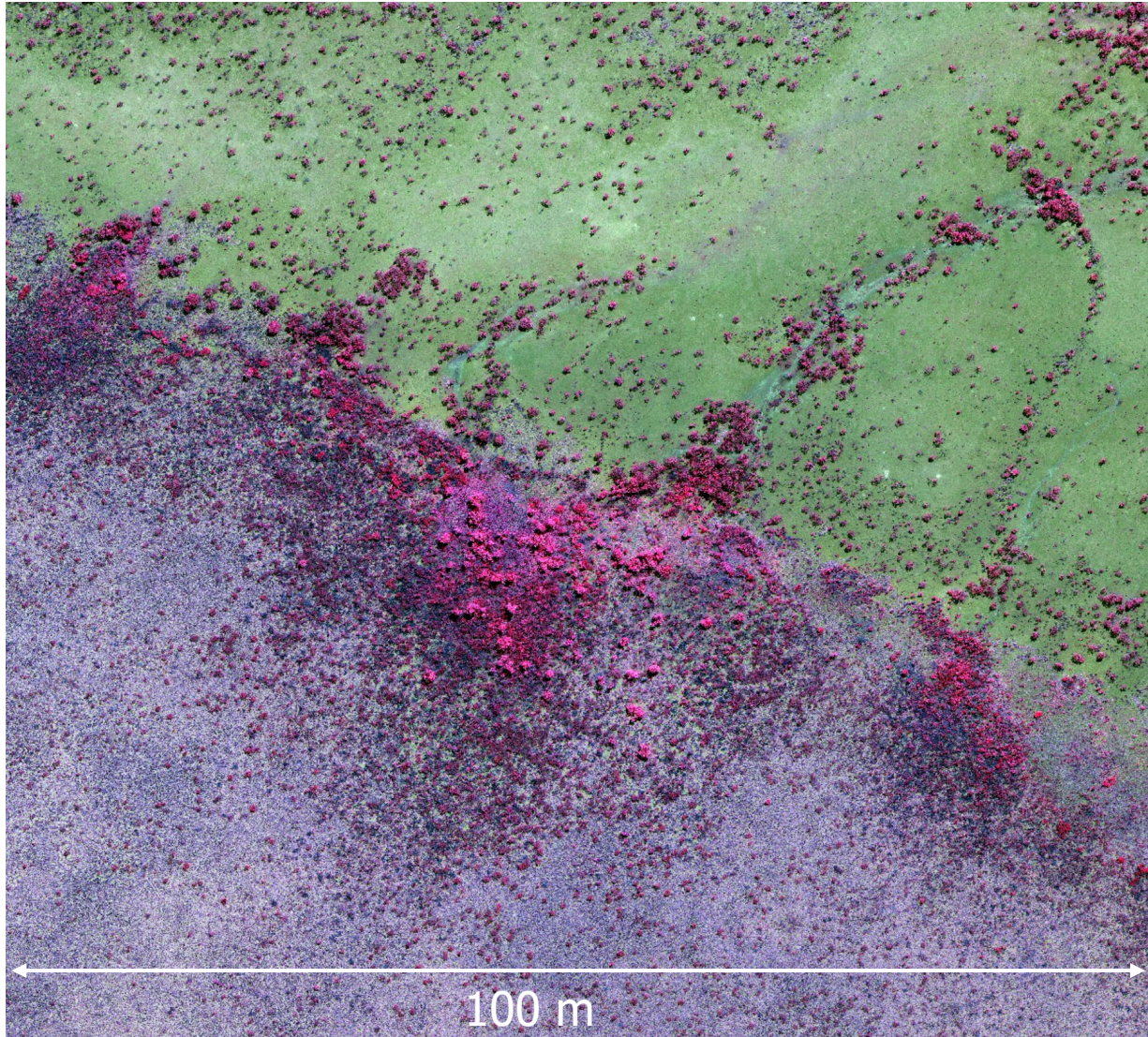
25 m

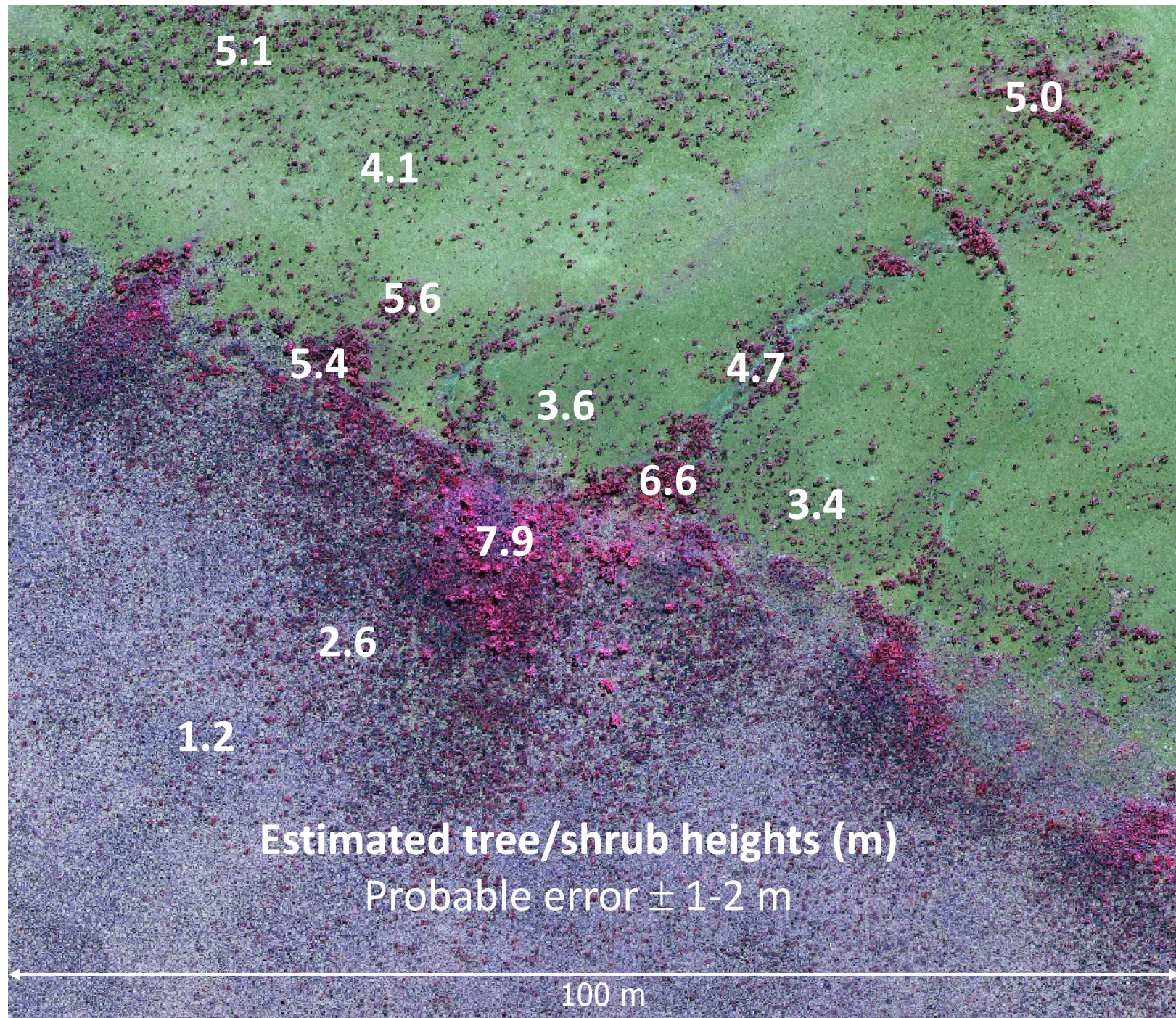
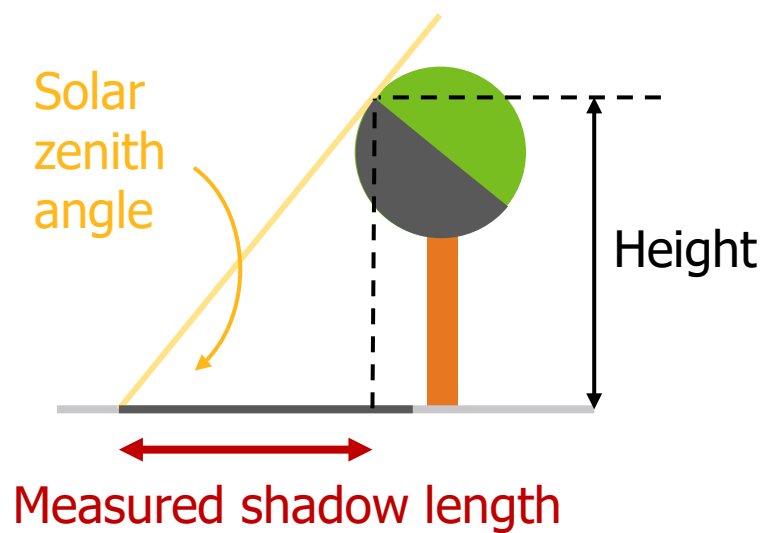
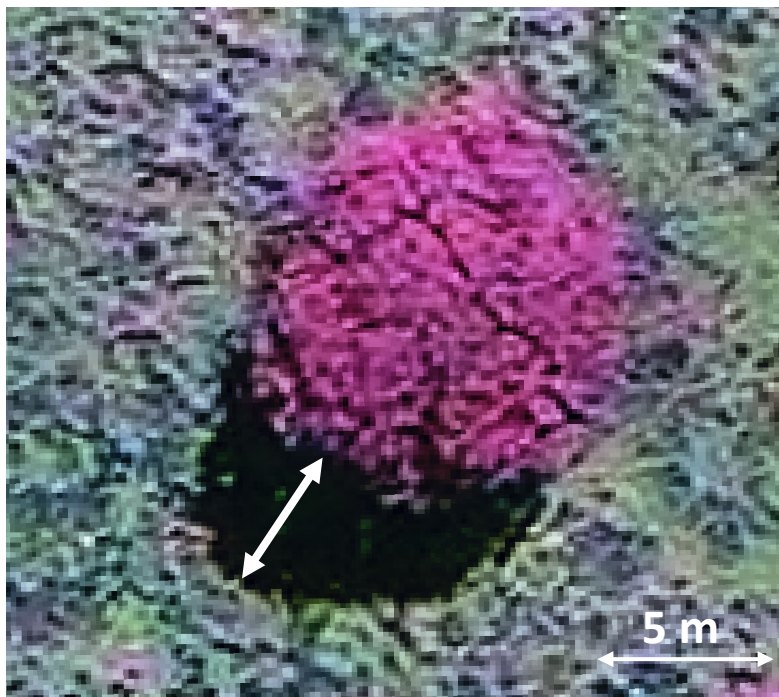




Deriving tree and shrub foliage cover

Spectral classification





Vegetation structure pseudo-observations

Attributes

- Spinifex (and other grass) cover
- Tree and shrub canopy cover
- Tree/shrub height
- Bare ground too?

4.59 km² of high-quality observations
(Ausplot Rangelands)

+

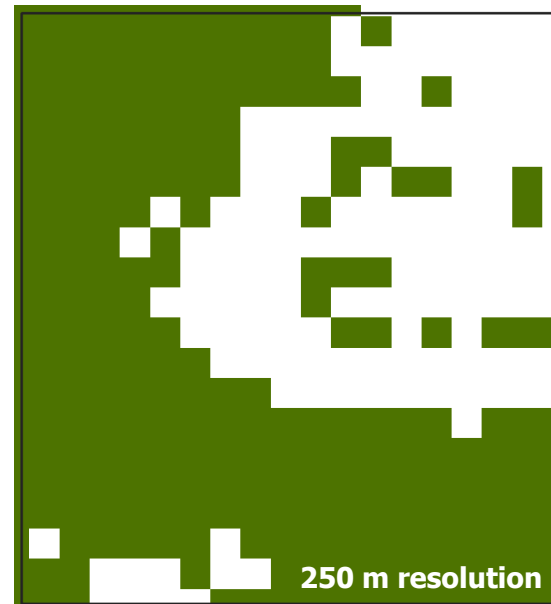
675 km² of pseudo-observations

With which we can build and validate
models of arid-region vegetation
structural type distributions

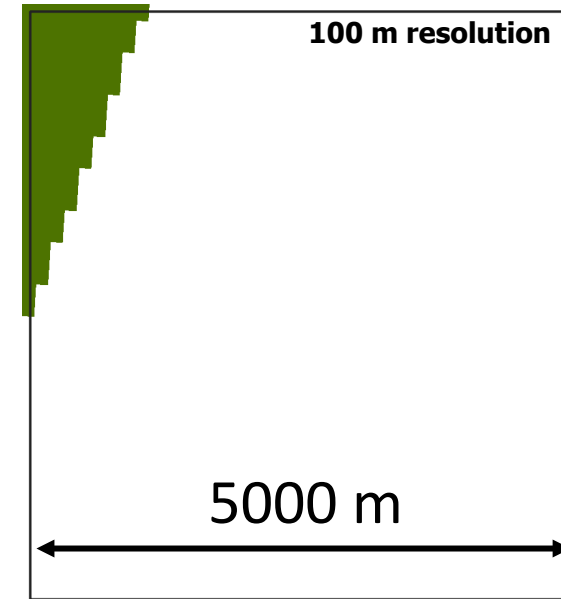
Pseudo-observed Spinifex



Spikey Spinifex



NVIS Spinifex



Carnahan Spinifex

