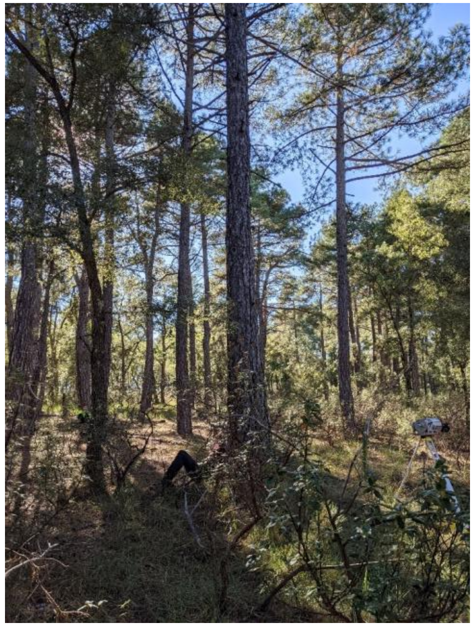
  

AI-enabled forest ecology: progress and challenges

Emily Lines   *ML4EO 2026, Exeter*

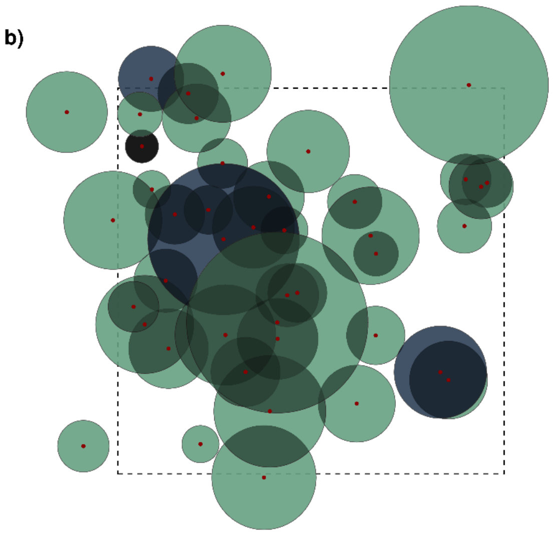
UNIVERSITY OF CAMBRIDGE

Imagine a tree

Pinus nigra ssp. salzmannii
Cambridge University Botanical Gardens

Imagine a tree like a forest ecologist

Figure 1.4—Crown diameter measurement.

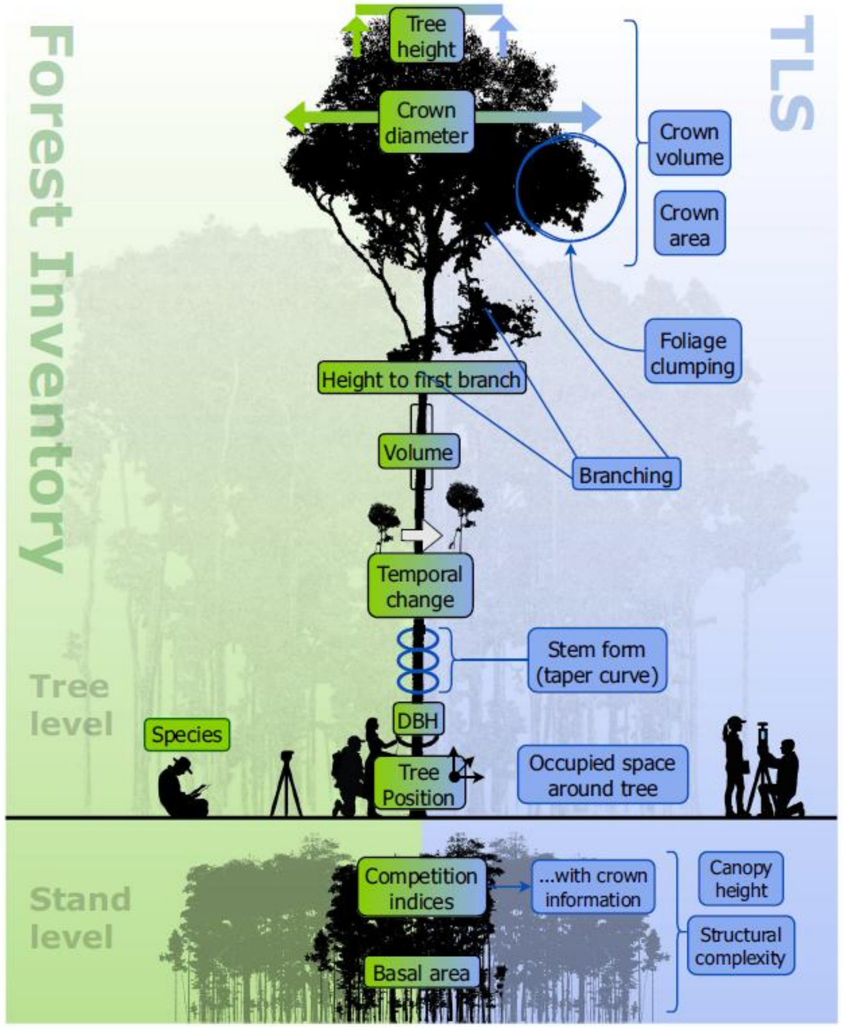


Pinus nigra ssp. salzmannii
Cambridge University Botanical Gardens

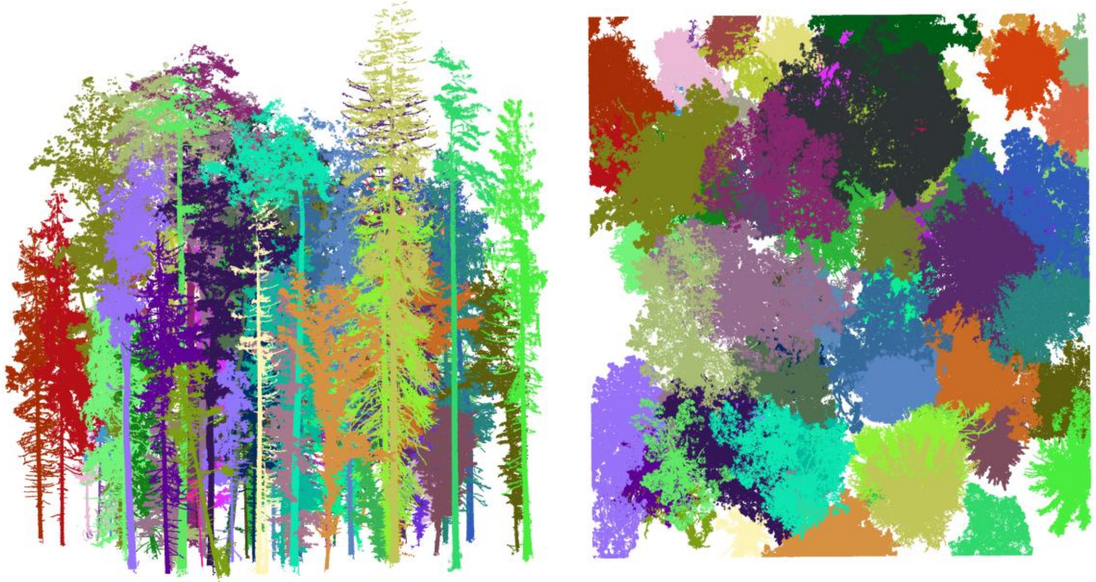
Proxy for:
Age
Carbon/biomass
Leaf area
Crown shape

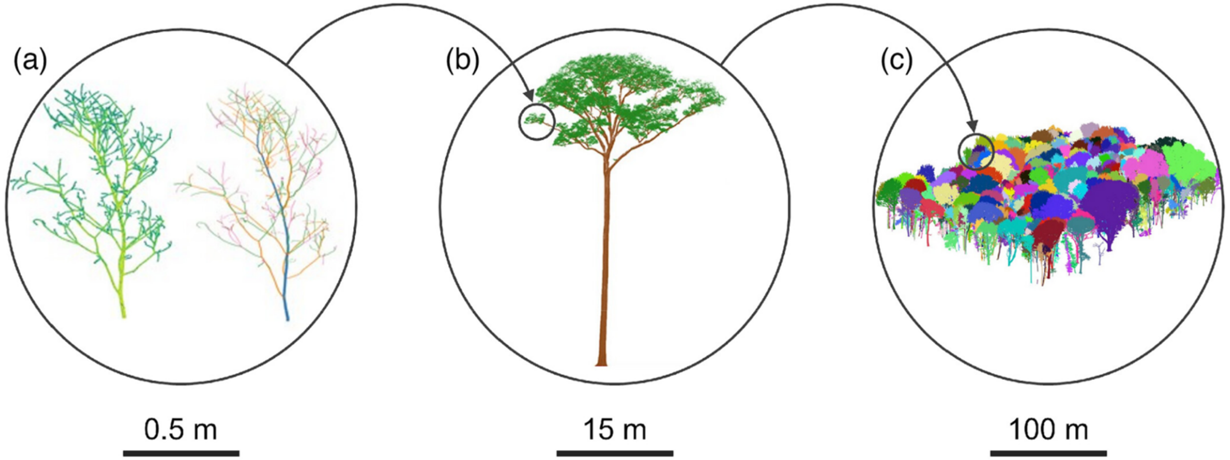
As a basis to **quanitify**:
Demography
Stand structure
Competitive interactions
Population dynamics

High resolution remote sensing offers rich direct measures (2D and 3D)



Integration of remote sensing and traditional methods (Maeda et al. 2026)





Lines et al. 2022

Processing raw data at scale to extract ecologically interesting properties is challenging

What is there?

How many trees are there?

Which species are they?

What is the quality of this habitat?

What size are the trees?

Is there regeneration?

How much X is there (biomass, nutrient, leaf area)?

What functions is this forest providing?

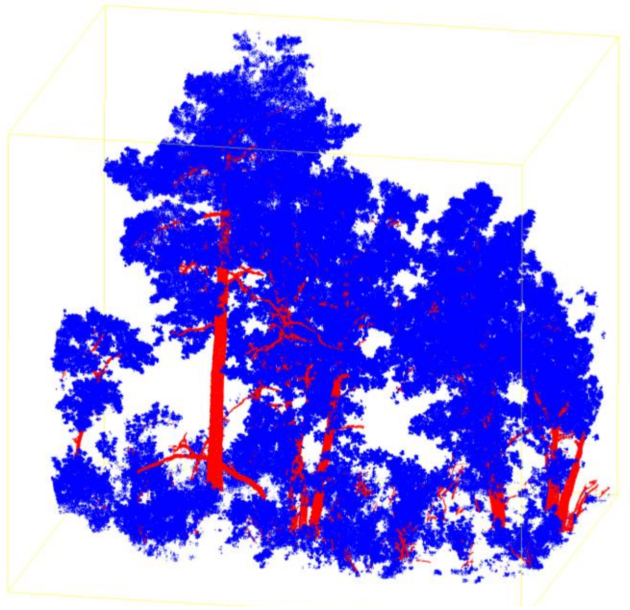
How resilient is this forest?

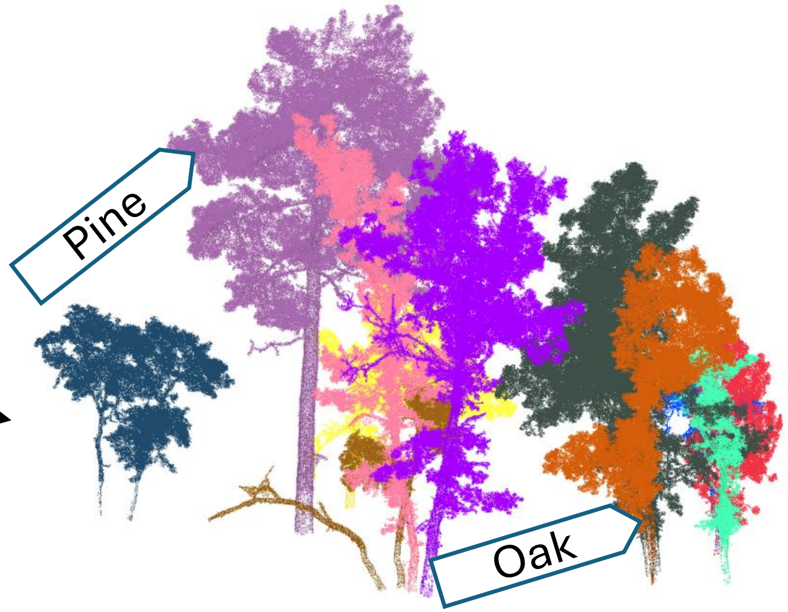
How **accurate** is this information?

Is it accurate **enough**?

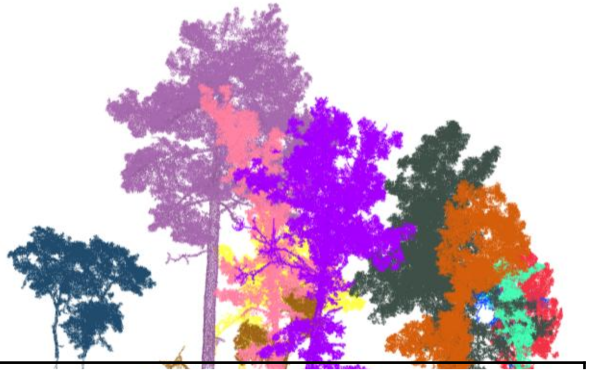
Deep learning as a solution for processing

- Semantic segmentation (leaves/wood/other)
- Instance segmentation (individuals)
- Species ID
- *Microhabitats, deadwood, lower vegetation…*





Have we made progress?



>15 studies on deep learning for 3D instance segmentation on TLS/MLS since 2021….

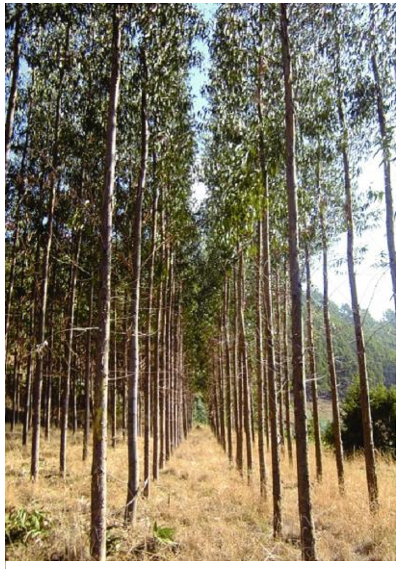
Approach		
PointNet		
YOLOv3 based		
PointNet++ based		
PointNet++		
PointGroup		
YOLOv5		
PointGroup		
Hybrid		
Hybrid		
RandLA-Net		
BFE-Net		
sparse UNet backbone		

time

The most recent major intercomparison study finds a **graph-based (non-DL) method** is current SOTA (Cherlet et al. 2026, *ISPRS JPRS*)

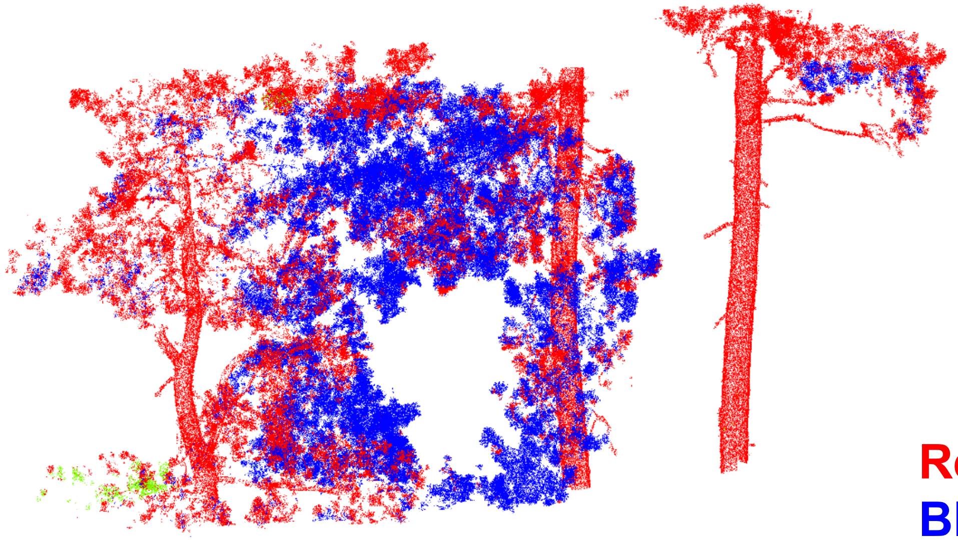

I HAVE THE BEST ONE!
MY SOLUTION IS THE BEST!
I HAVE THE BEST!
MINE WORKS EVERYWHERE!

Over-emphasis on standard model development means ‘successful’ tools are not useful

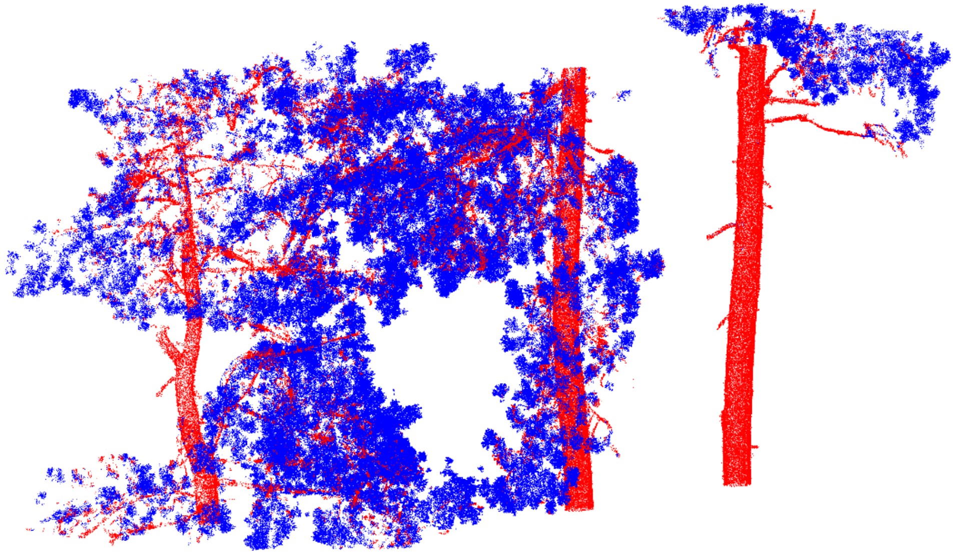


Original model trained in monoculture ecosystem, multi-sensor

Retrained on mixed natural forest, single sensor



Red – wood
Blue – leaf



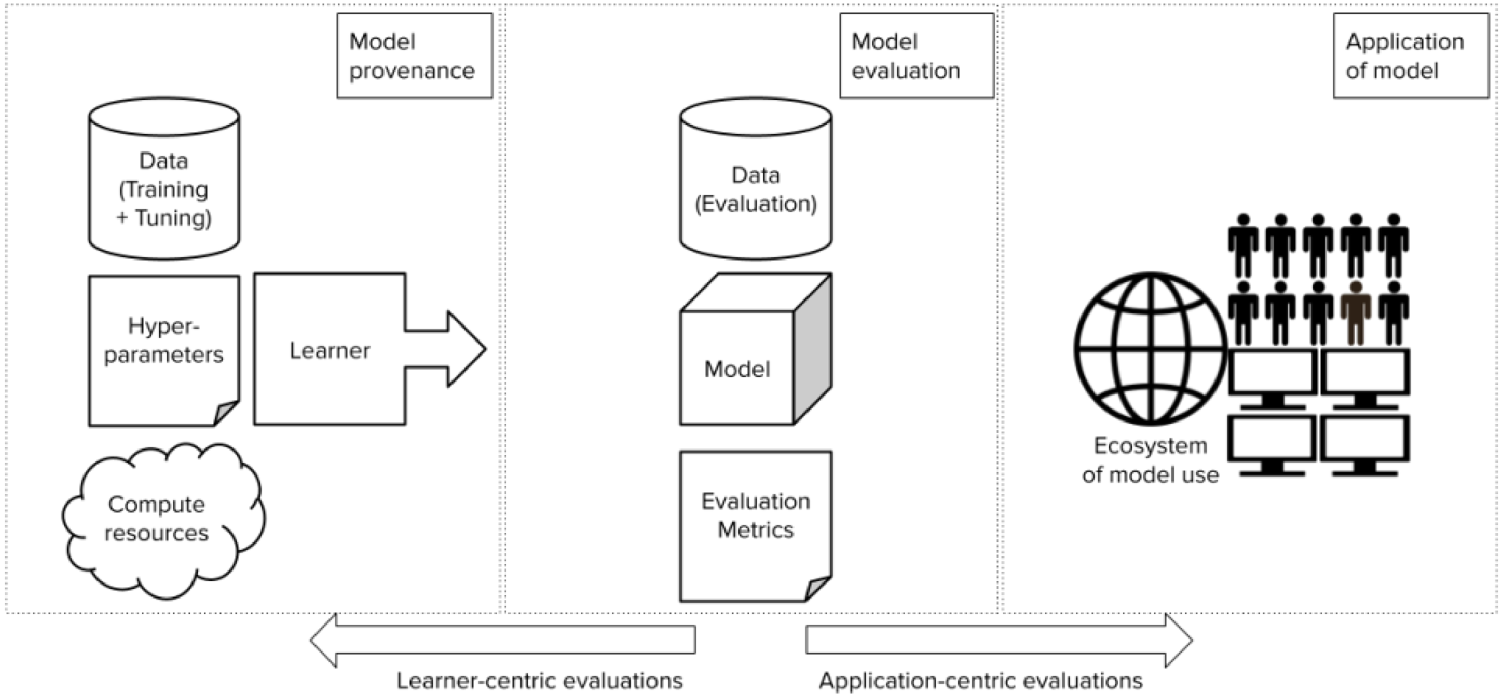
A focus on technical development

Agreed benchmarks

Model development

Model evaluation

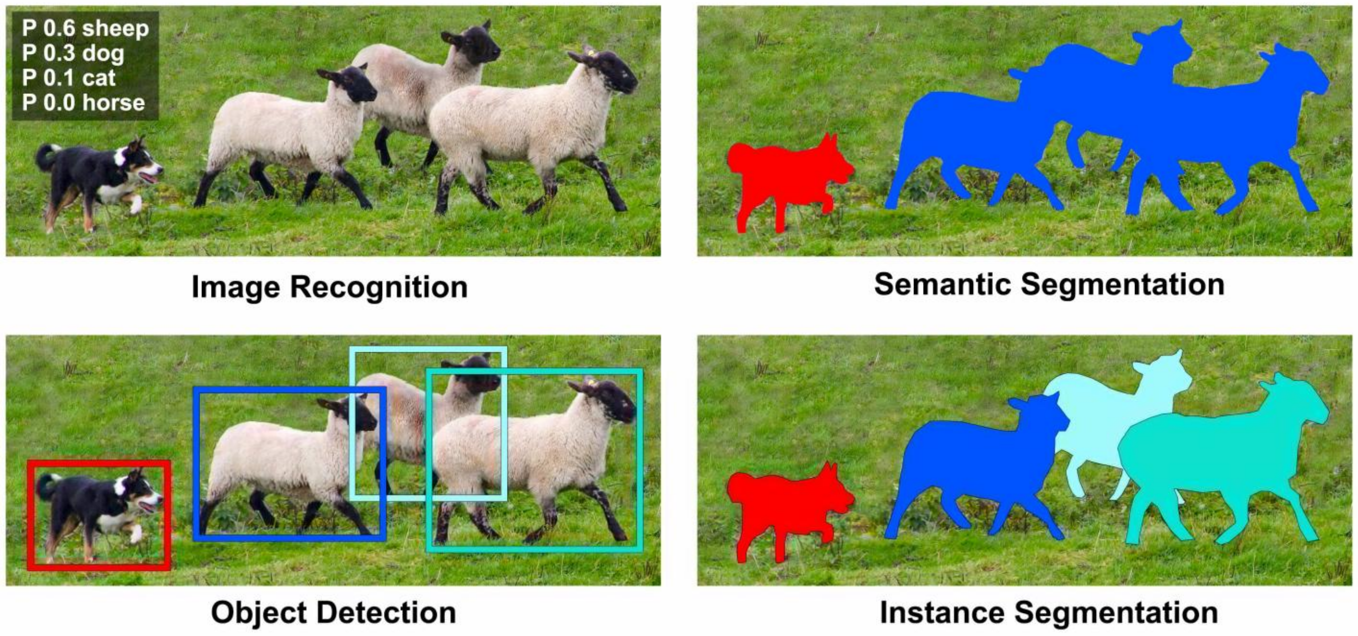
Task evaluation

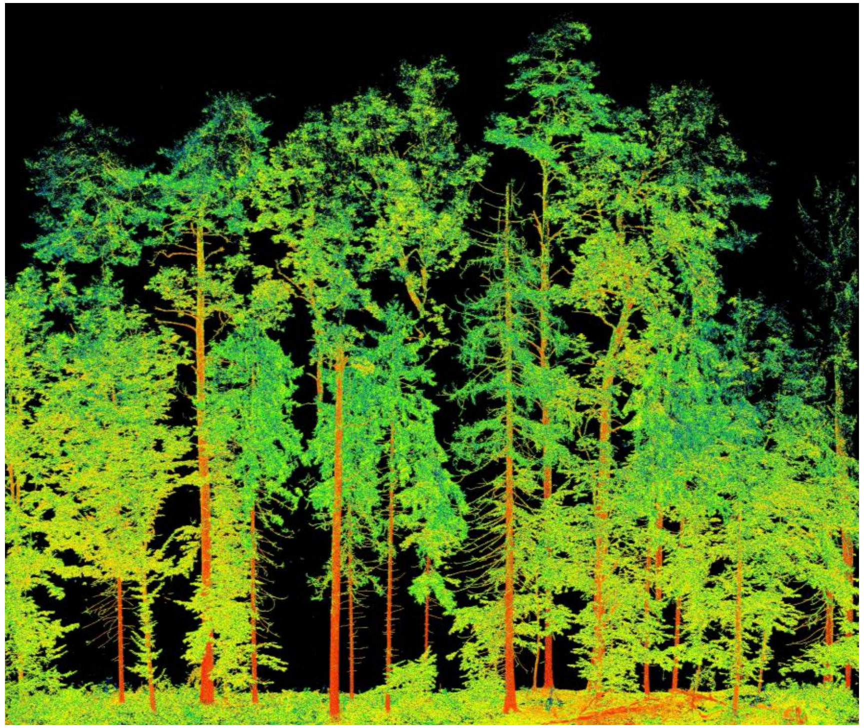


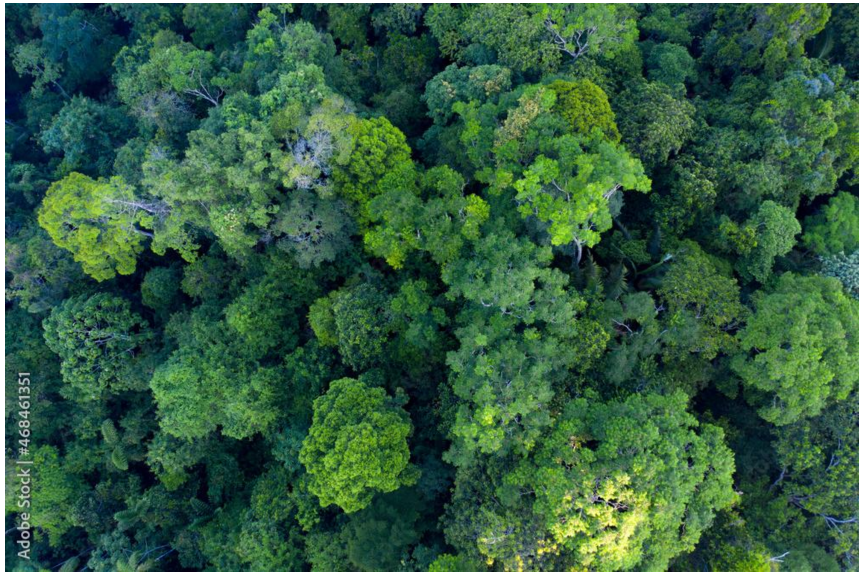
Hutchinson et al 2022

Benchmarking, evaluation, and truth

The history of machine learning is built on successes in *automating doable tasks*



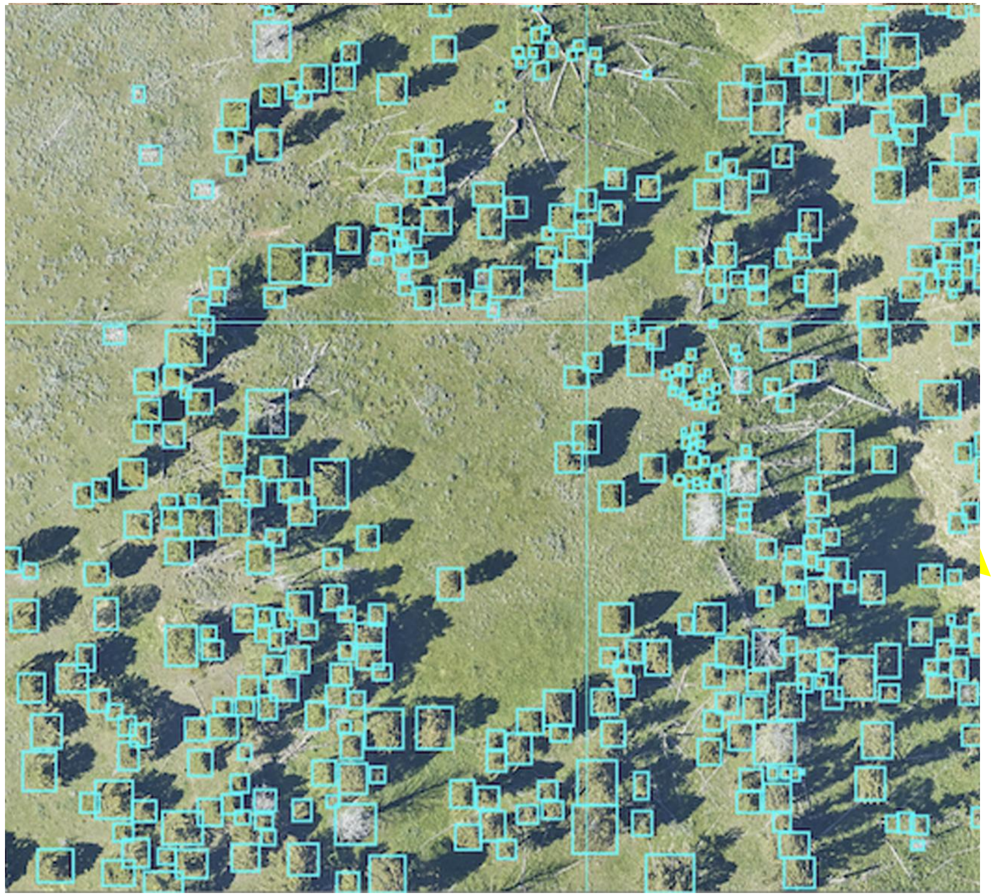


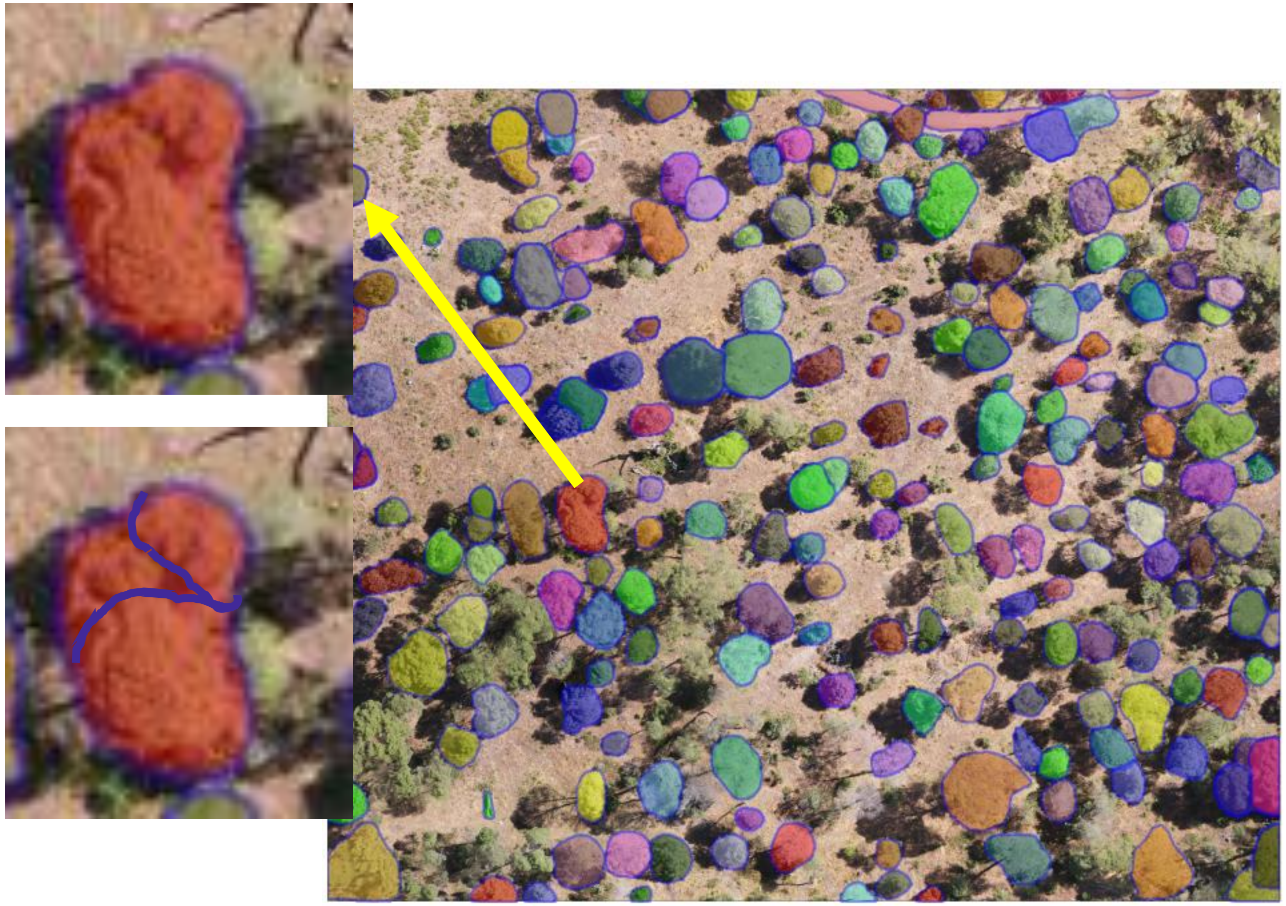


Did we skip a step?

The (**ground**) truth matters: individual tree segmentation in imagery

Crown segmentation tools *look* good

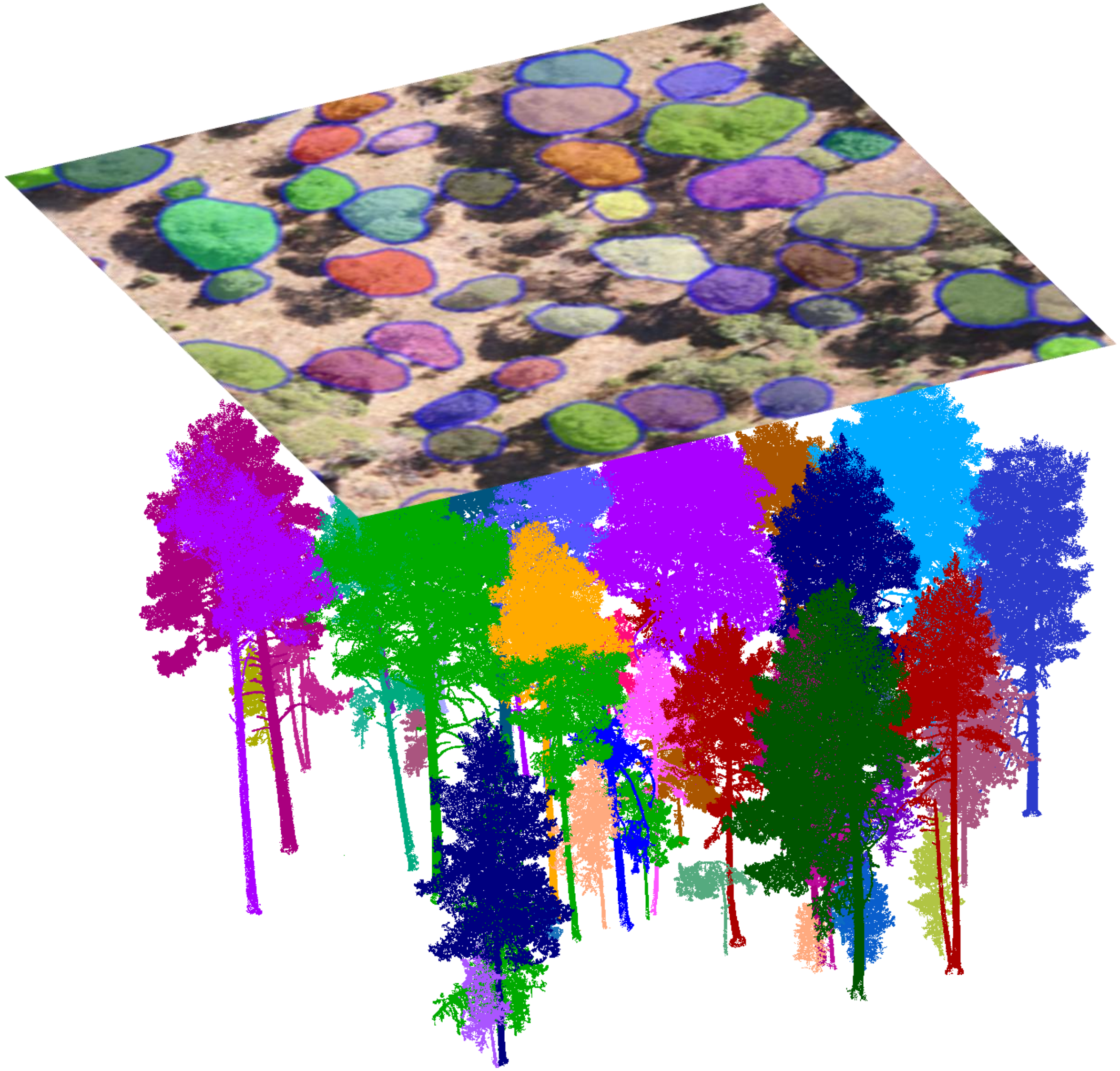


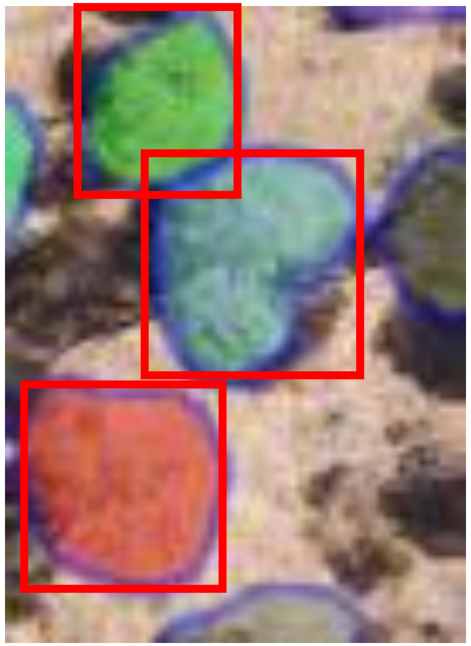


Ground truthing is **crucial** for believable and robust monitoring.

Weinstein et al. 2020

3D data offer the opportunity for high quality ground truth validation





Most 2D image tree segmentation methods are trained on manual delineation.

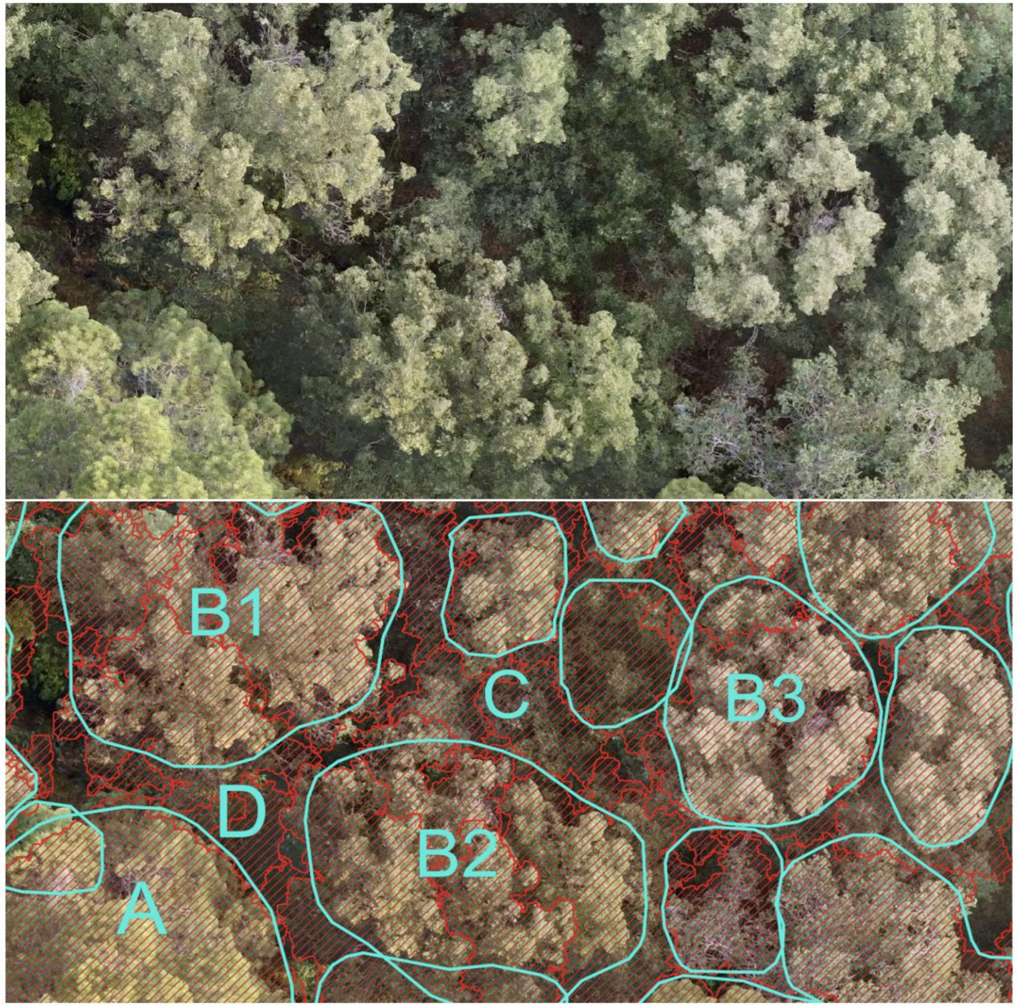
Co-located data from ground 3D data and aerial imagery allowed us to evaluate deep learning models.

We cut the 2D UAV imagery with our 3D individual trees to create high quality test data.

Compared two DL models' performance on TLS and manual labels, using bounding boxes and IoU 0.5, 0.75.

Allen *et al.* (arXiv 2025 & in review)

Manual label ‘truth’ vs TLS ‘truth’ shows 2D aerial crown delineation more challenging than previously thought



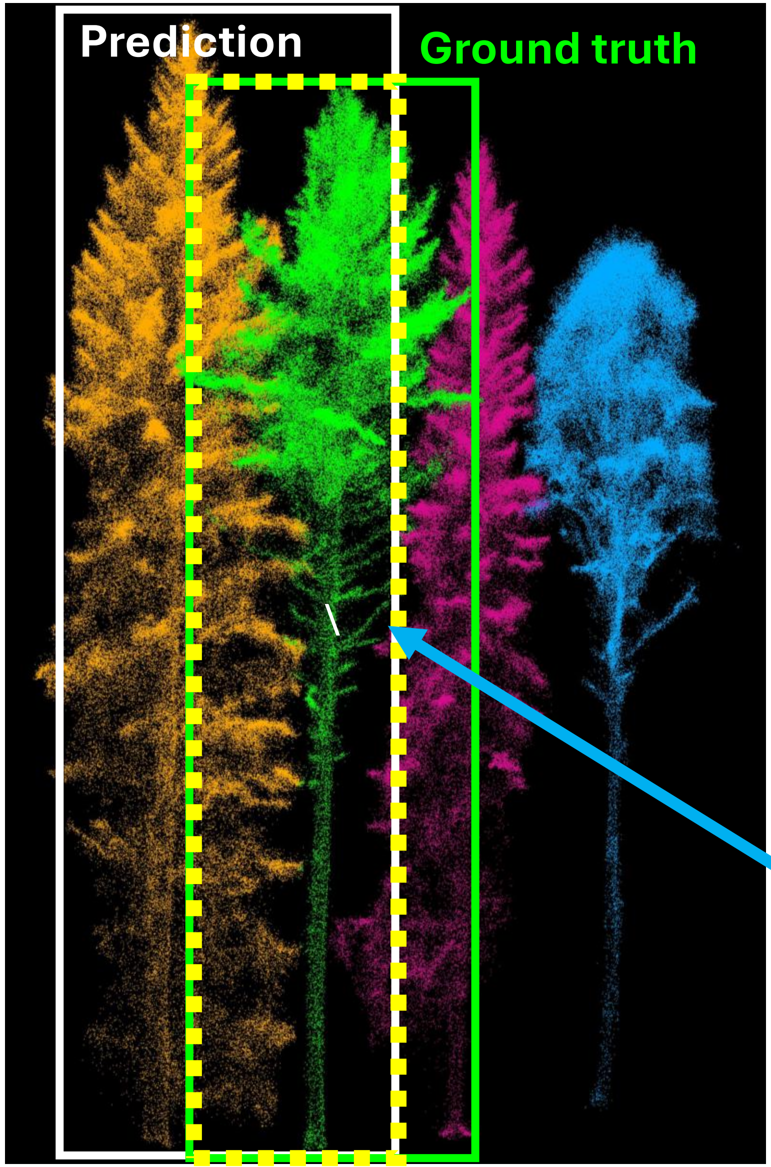
	Site	Mod	Tile Size	AP_{50}		AP_{75}		F1	
				All	Can	All	Can	All	Can
Manual drawn labels	Almorox (Manual)	DF							
		DT							
TLS labels, manually checked	Alto Tajo (TLS)	DF							
		DT							
	Joensuu (TLS)	DF							
		DT							

Typically, errors come from missing smaller trees, and mis-grouping multiple trees together.

Training on **manual** labels of aerial imagery seems to artificially inflate performance, likely due to biases towards easily identifiable canopy trees.

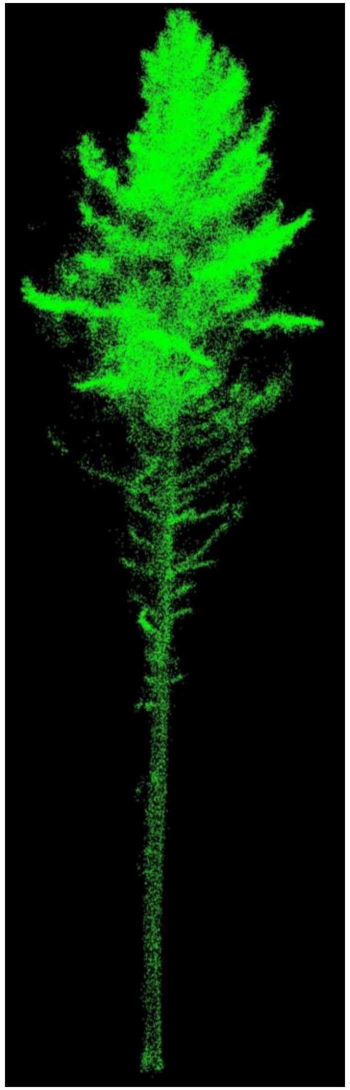
Allen *et al.* (arXiv 2025 & in review)

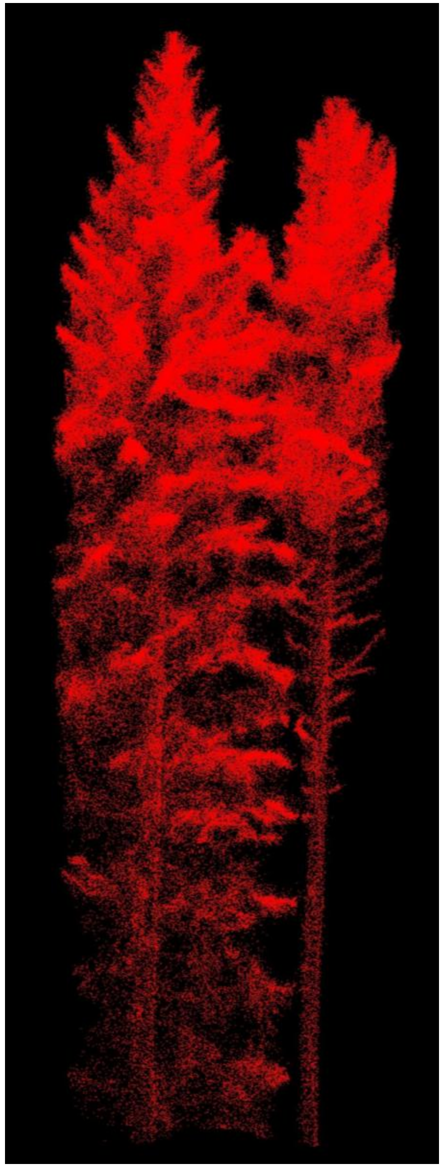
Defining success in instance segmentation: IoU



IoU > 0.5:

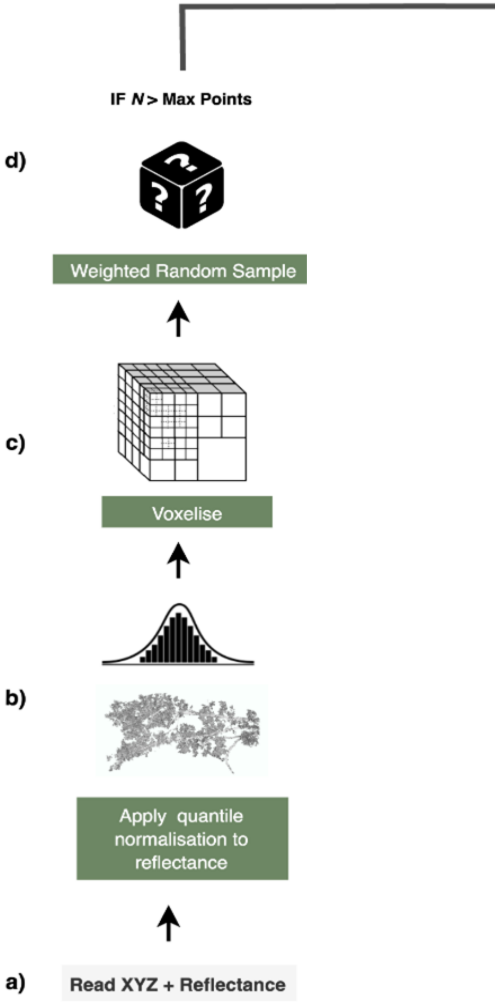
a “success” *(by ML standards)*

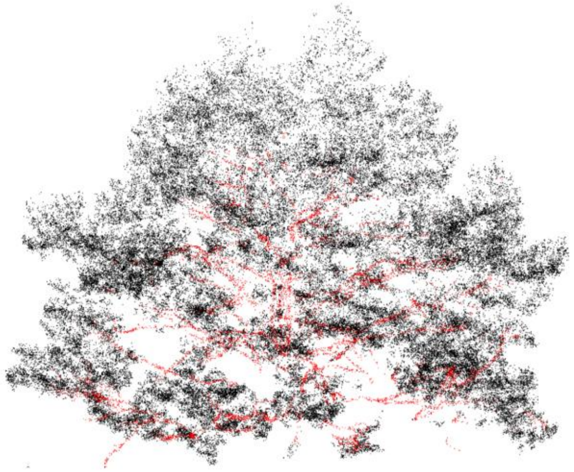




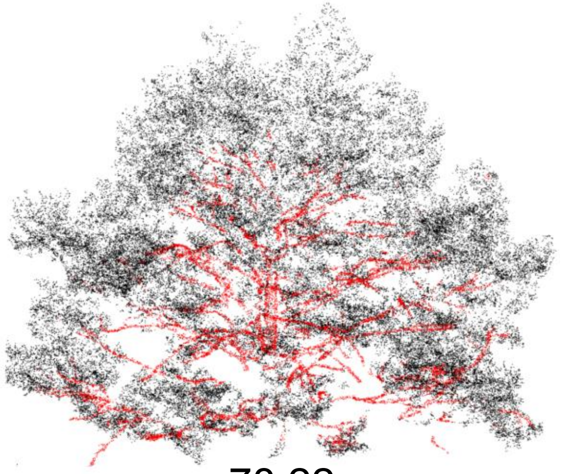
New ecological evaluation: leaf-wood segmentation

We tailored existing architecture *and* performance metrics to target hard to classify, rare, small features (wood).
Trained on three biome data (boreal/temperate/Mediterranean) and binary classification.





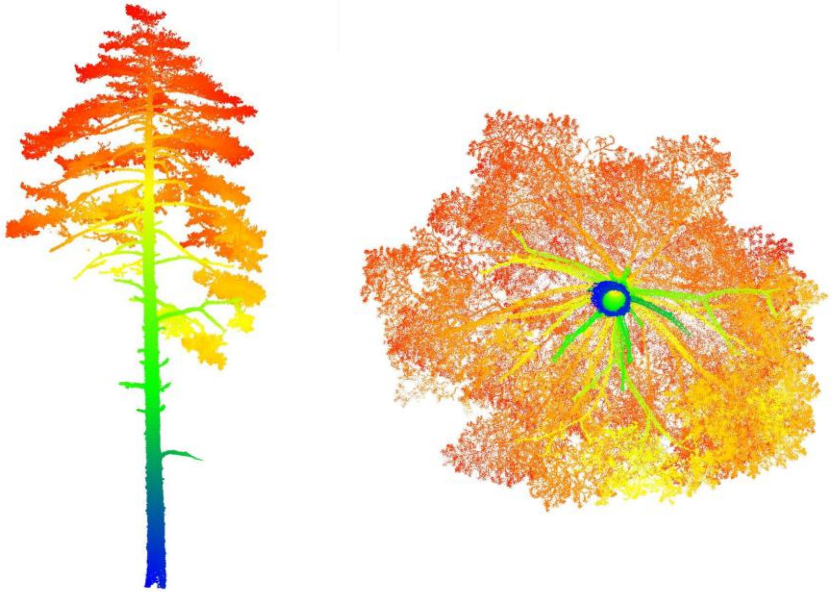
90:10



78:22

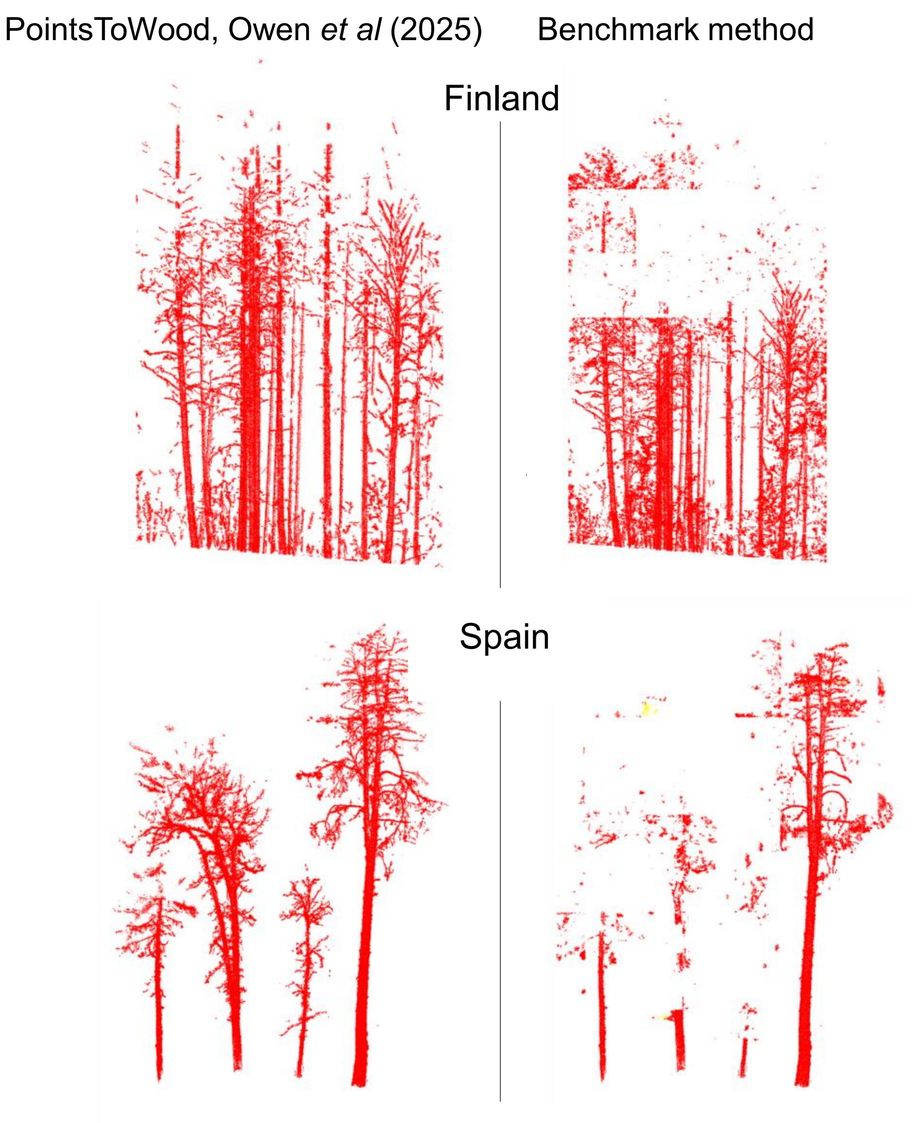
Downsampling to try to retain more of the rare class (wood).
Adding reflectance to the x-y-z information

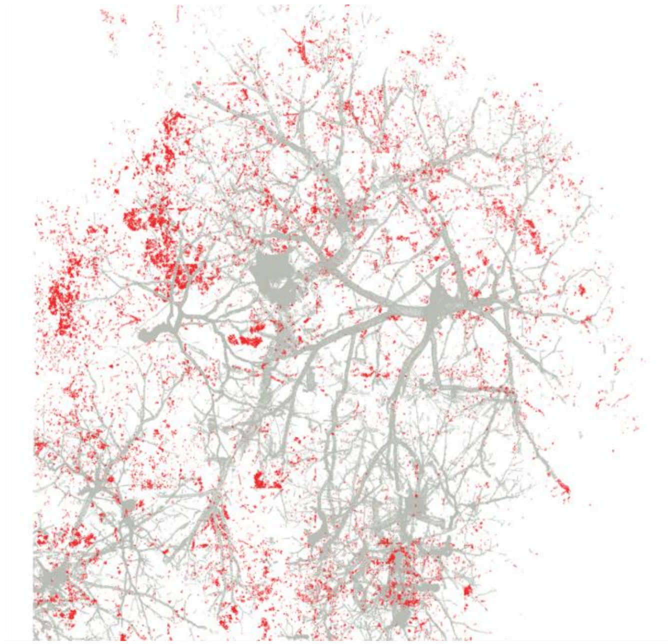
Evaluation in the hard to classify areas (twigs) by weighting accuracy by path length



Owen et al. (arXiv 2025 & in review)

Improved *ecologically relevant* performance vs benchmark (retrained)





Clear improvement on small element classification (twigs)

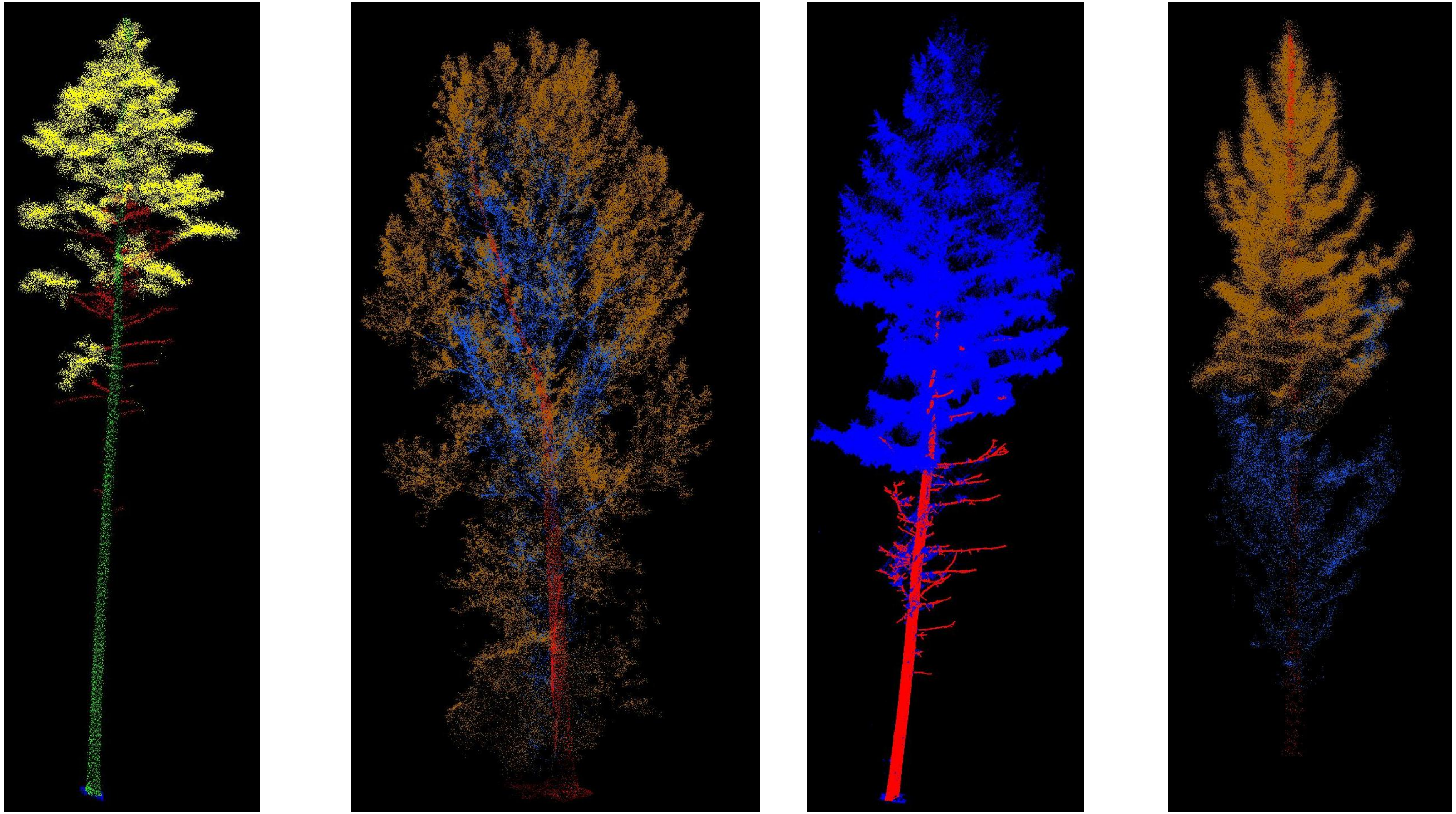
← **grey** shows our predictions; **red** shows false labels by benchmark method

Fine details matter for crown reconstruction, but represent a *tiny* fraction of the total points, making evaluation difficult.

Balanced accuracy **weighted by pathlength** highlights improvements in rare, messy, and hard to classify extremities that are hidden by standard metrics.

Owen et al. (arXiv 2025 & in review)

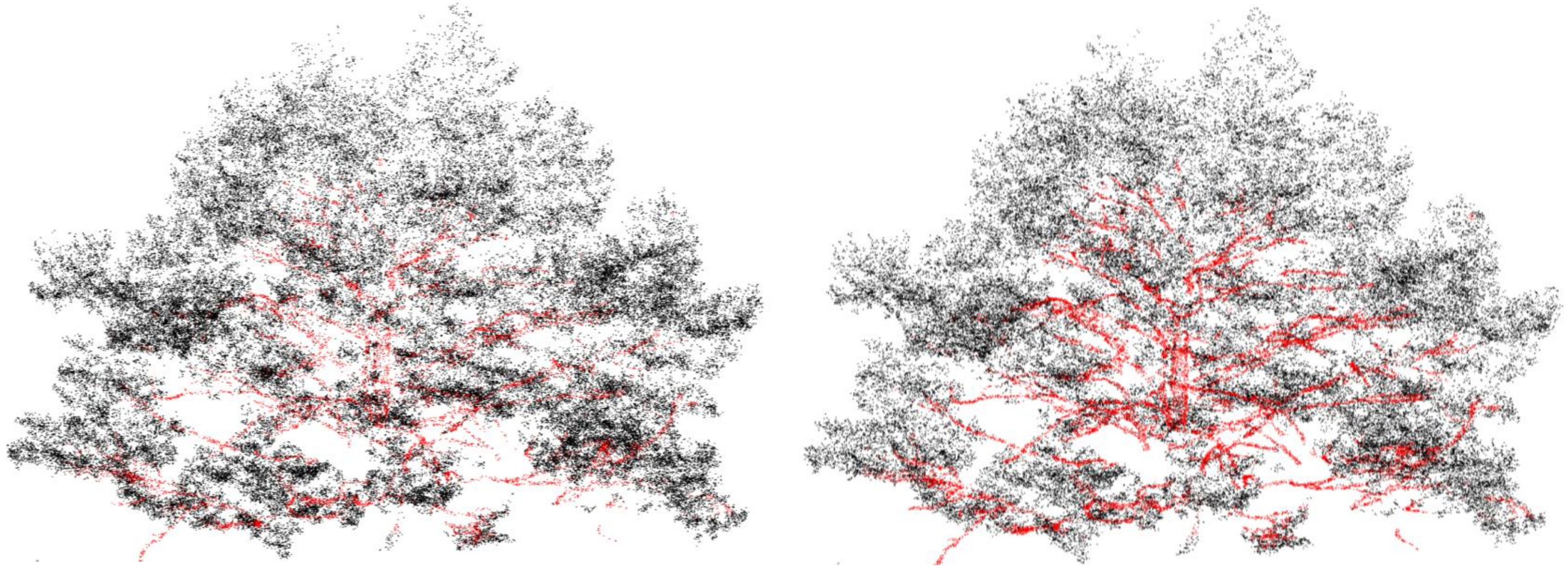
The state of benchmark datasets *(good labels are gold dust)*



Better models will do **worse** on poor ‘benchmark’ datasets

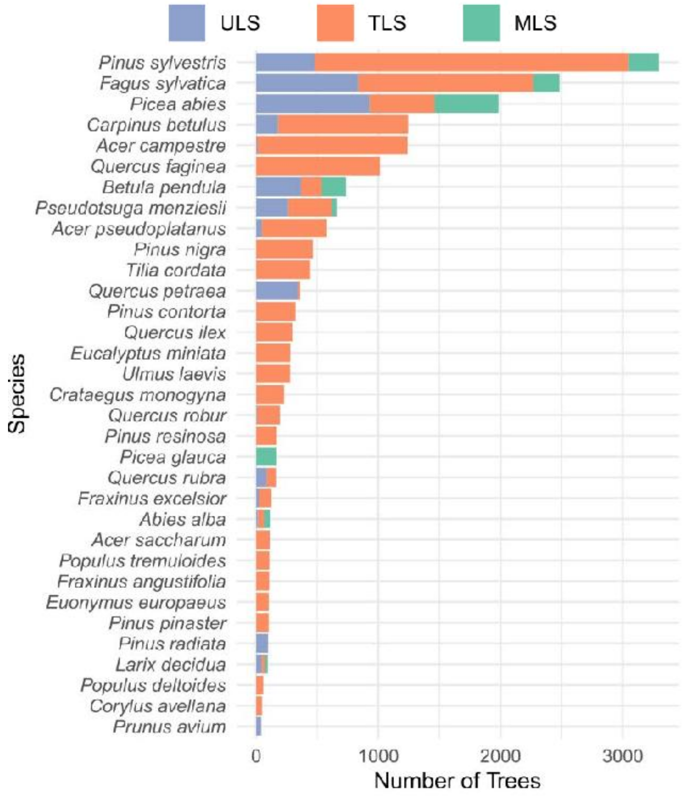
Ecological tasks need ecological metrics of success. But our data brings structural challenges: uneven classes are the norm

High density data cannot be ingested into networks. Standard downsampling loses key details



Owen *et al.* PointsToWood (arXiv & in review)

Rare species are undersampled but crucial for biodiversity monitoring.

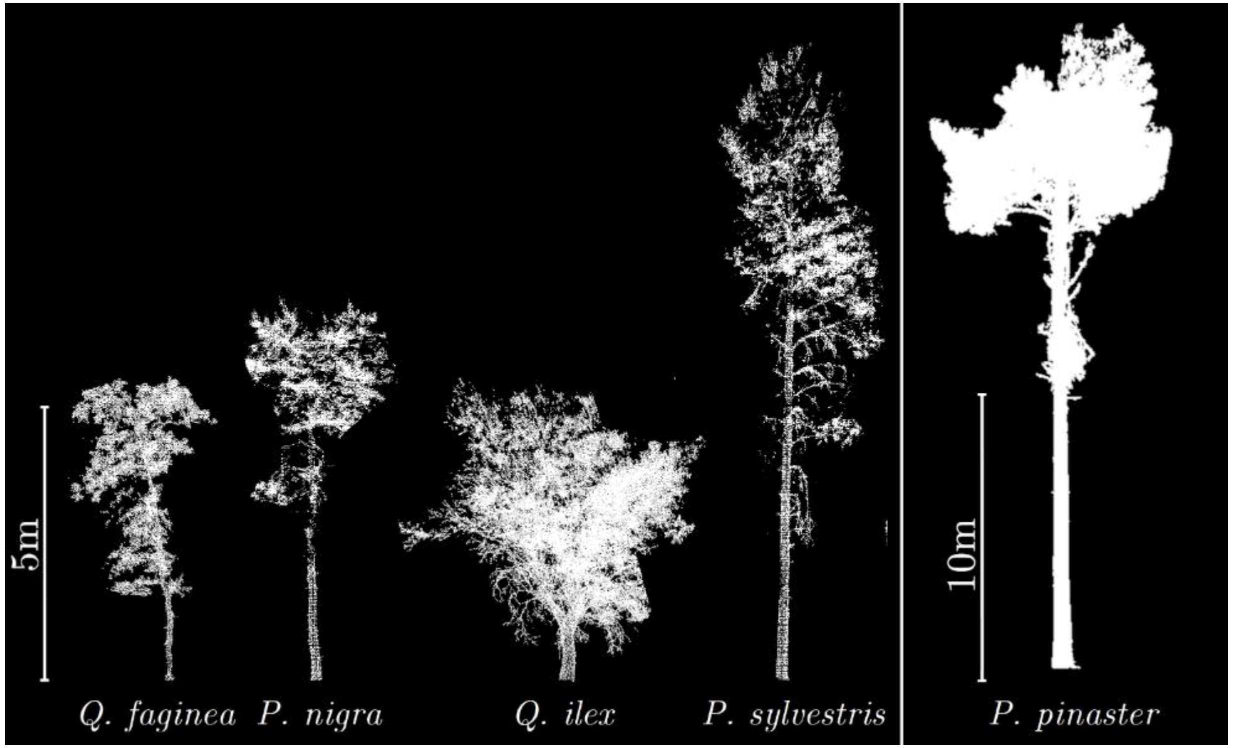


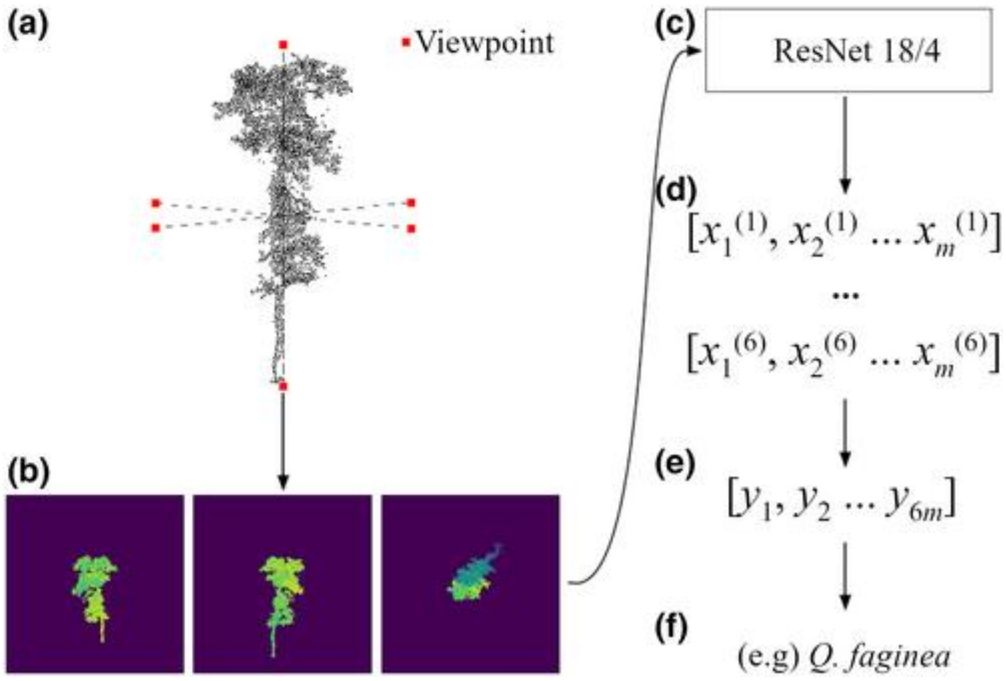
and its very time consuming to manually label millions of points

Puliti, Lines *et al.*
FOR-species20k *(MEE,* 2025)

What will help?

Monitoring diversity: tree species identification with AI



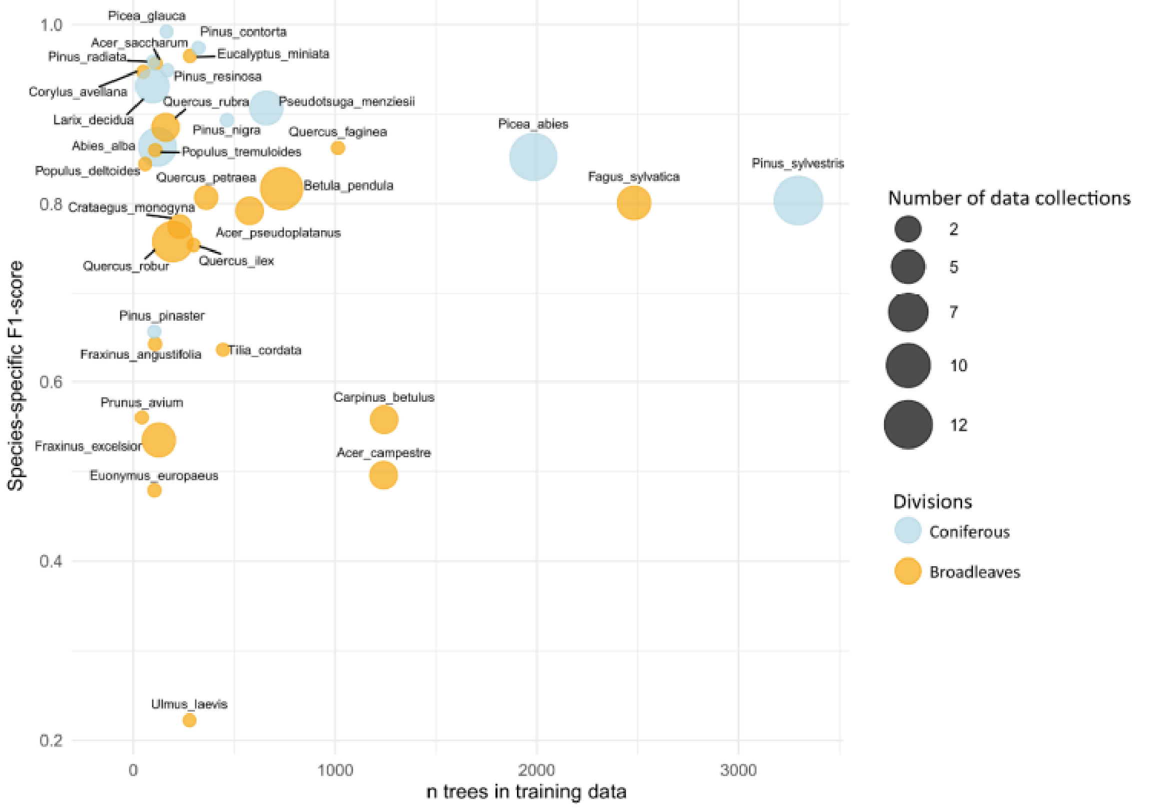


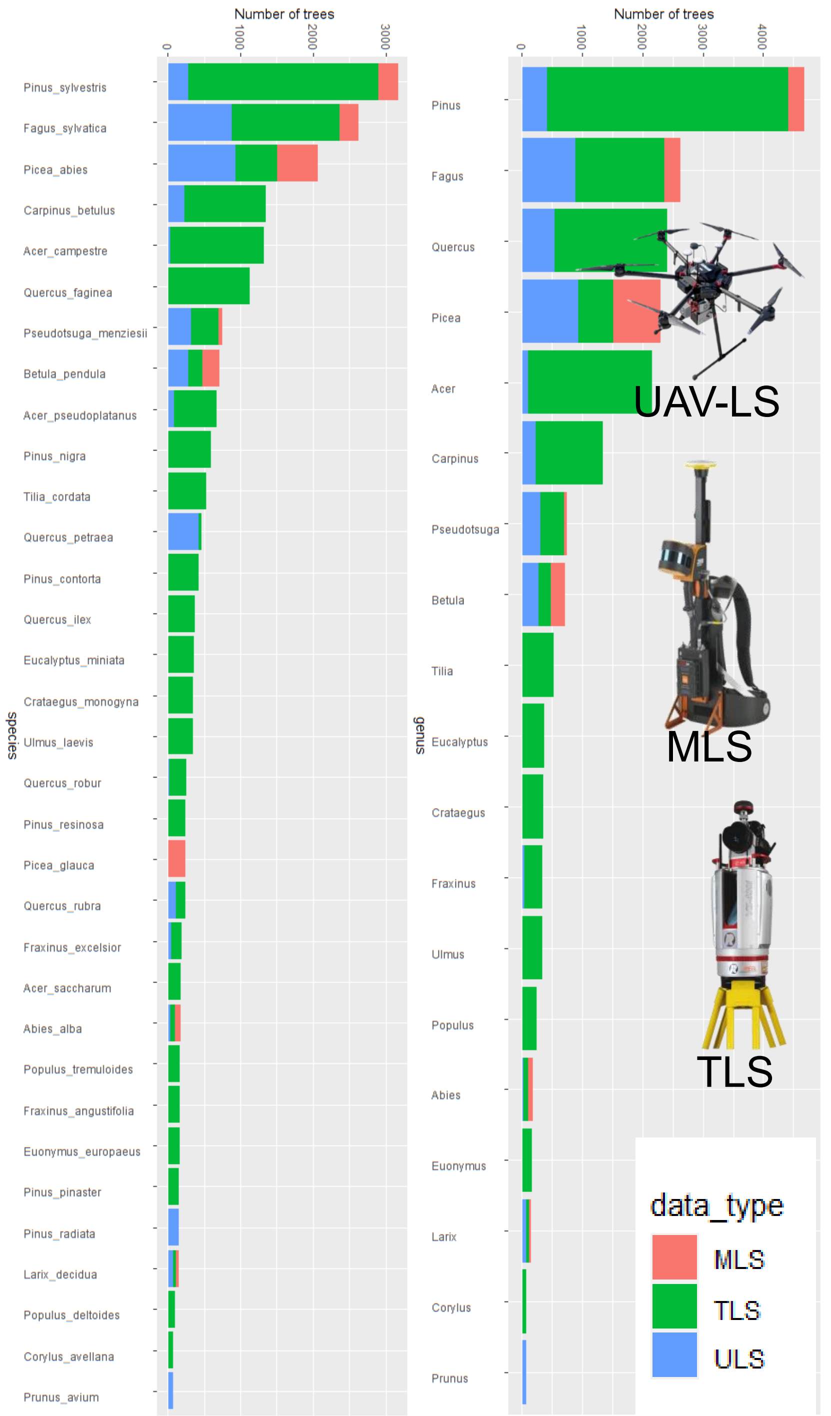
CNN for species classification, with camera projections and SimpleView

Allen *et al.* (2022)

Towards broad 3D point cloud tree species identification

More data probably helps, but there are likely other limits on performance

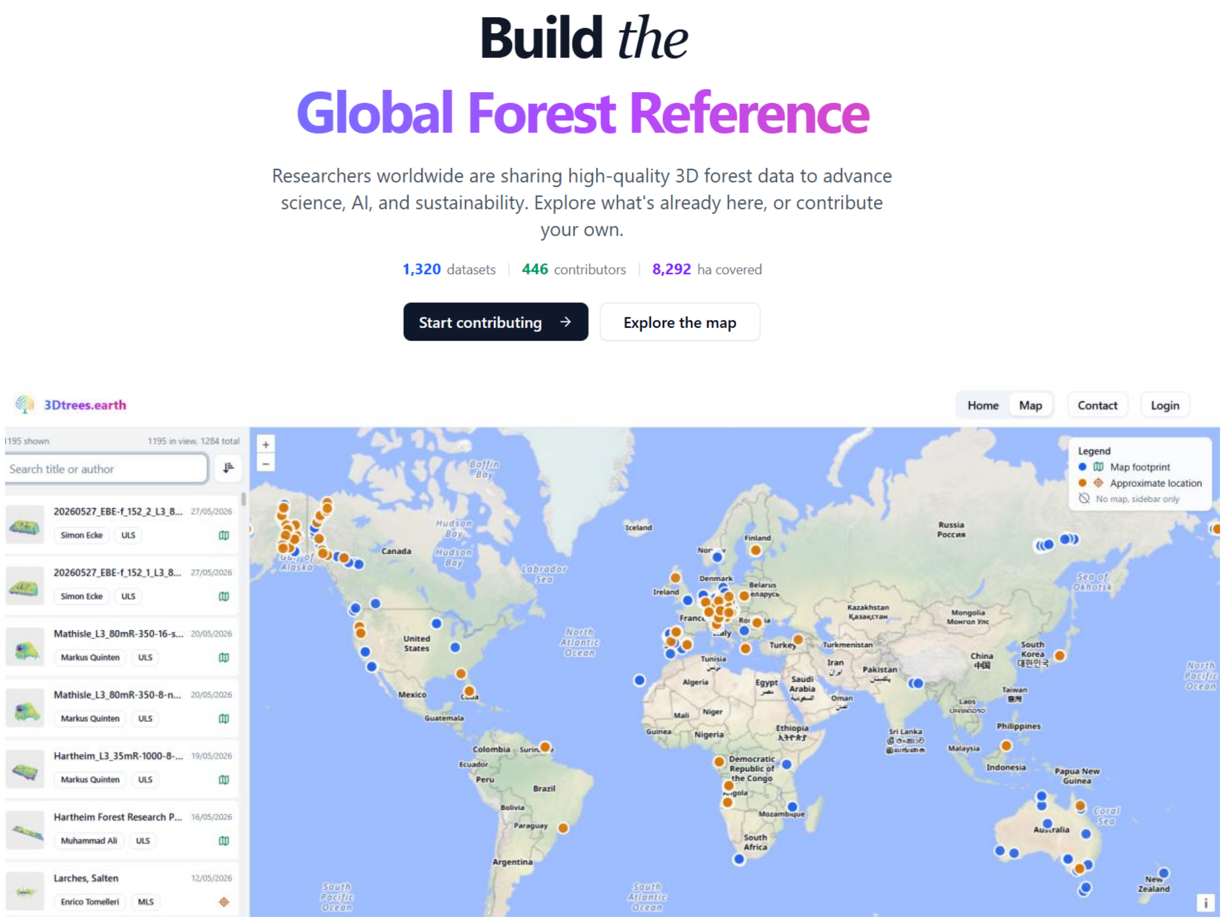




For-Species20k Puliti, Lines *et al.* (2025) *Methods in Ecology and Evolution*

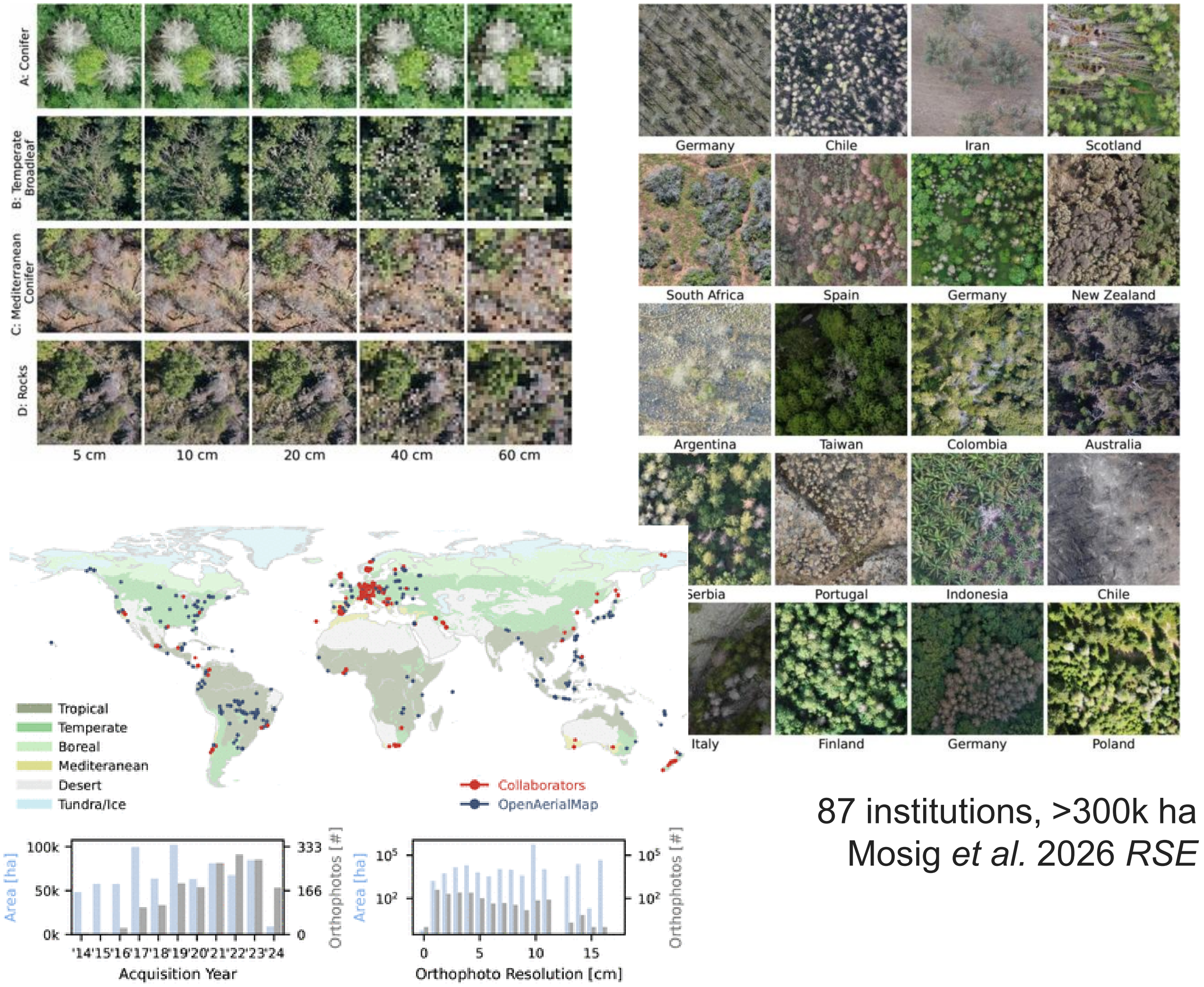
Community sourced benchmarking for data diversity

3dtrees.earth



Vajna-Jehle, Gerberding, Kattenborn et al. *in progress*

deadtrees.earth



87 institutions, >300k ha
Mosig *et al.* 2026 *RSE*

How do we get to the point of trusting outputs?

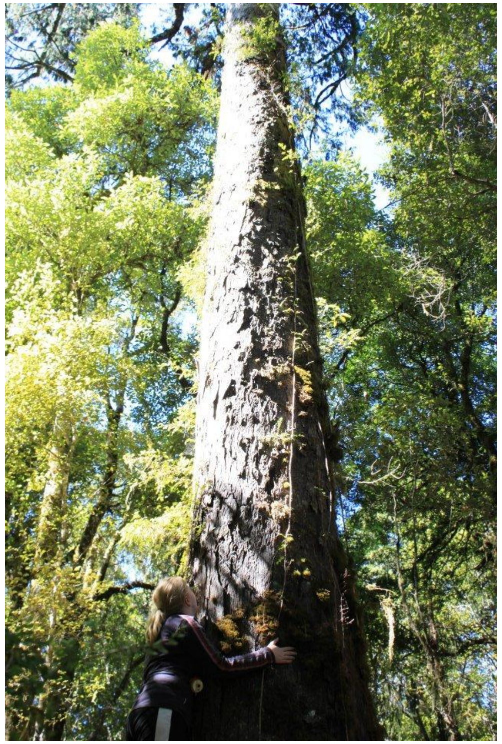
Forests and sensors are diverse and data is hard to label: progress is only possible by working together as a community.

However, not all labelled data is good enough for use and we should be strict on what gets used for benchmarking.

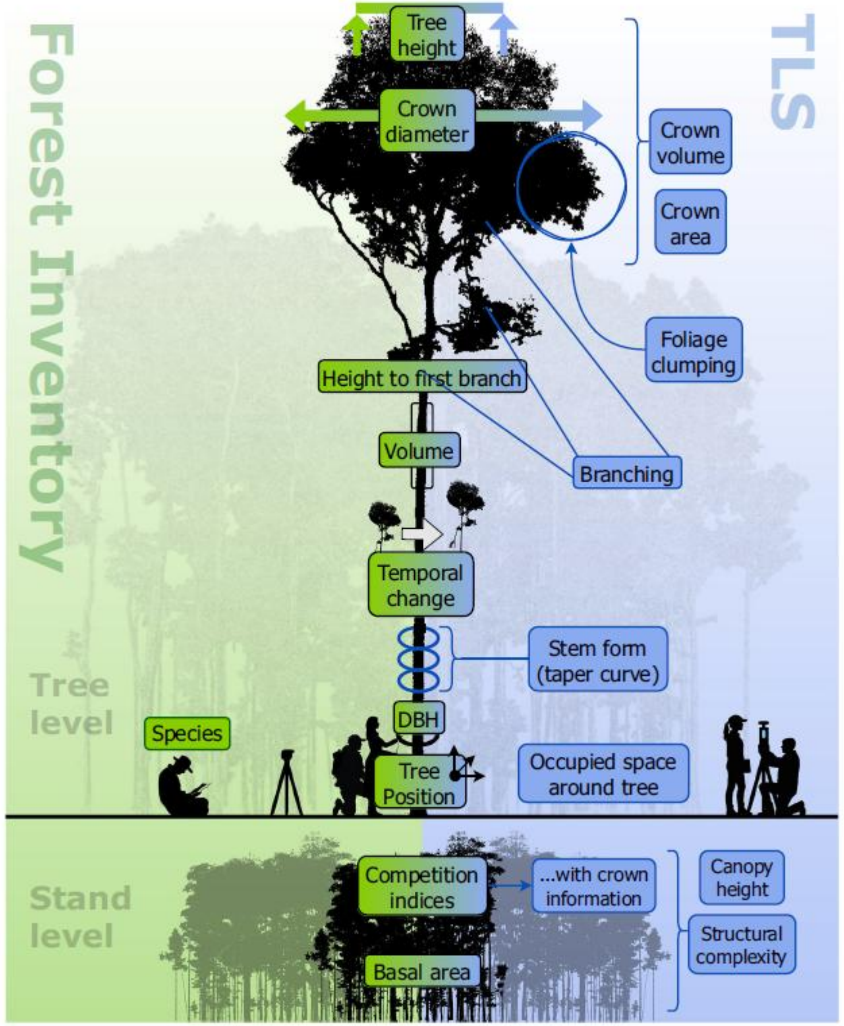
There is enough open data to demonstrate:

- Real usefulness of tools for *tasks (not just F1 scores)*, **and**
- Improvements over other methods

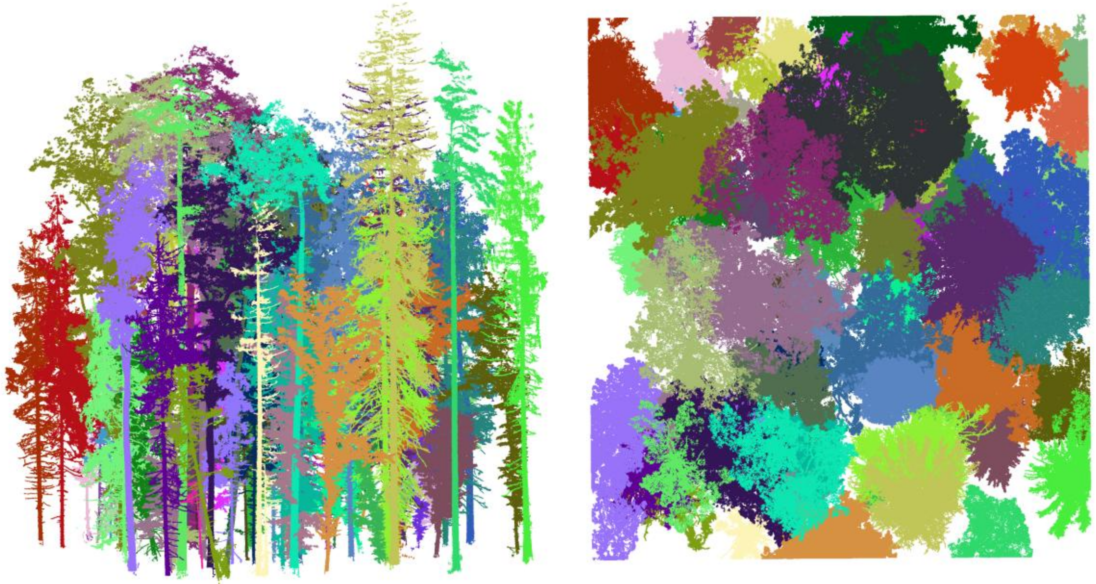
And demonstrating these should be required by us a scientists, reviewers and editors.

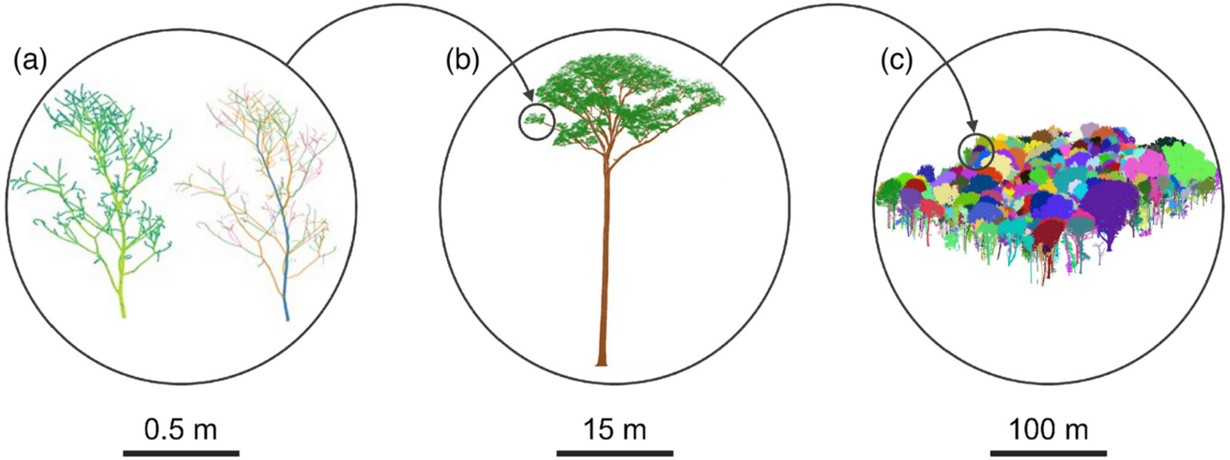


Are high resolution data worth it?

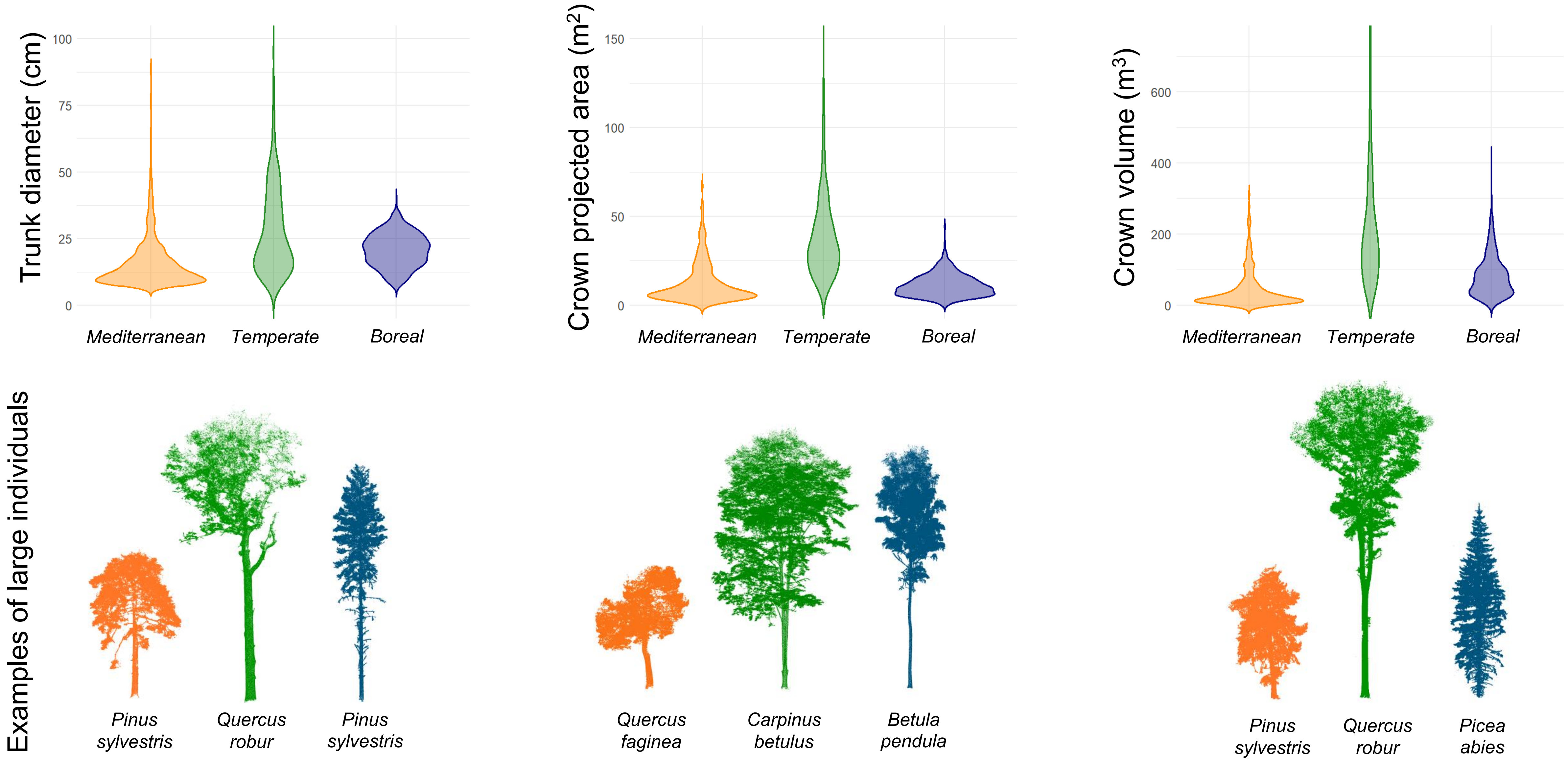


Integration of remote sensing and traditional methods (Maeda et al. 2026)

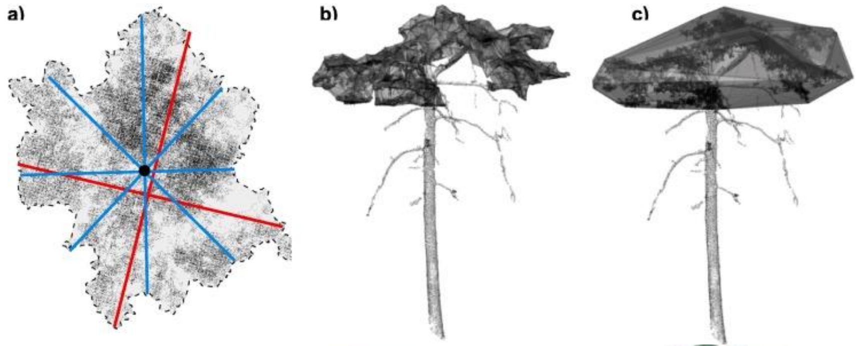


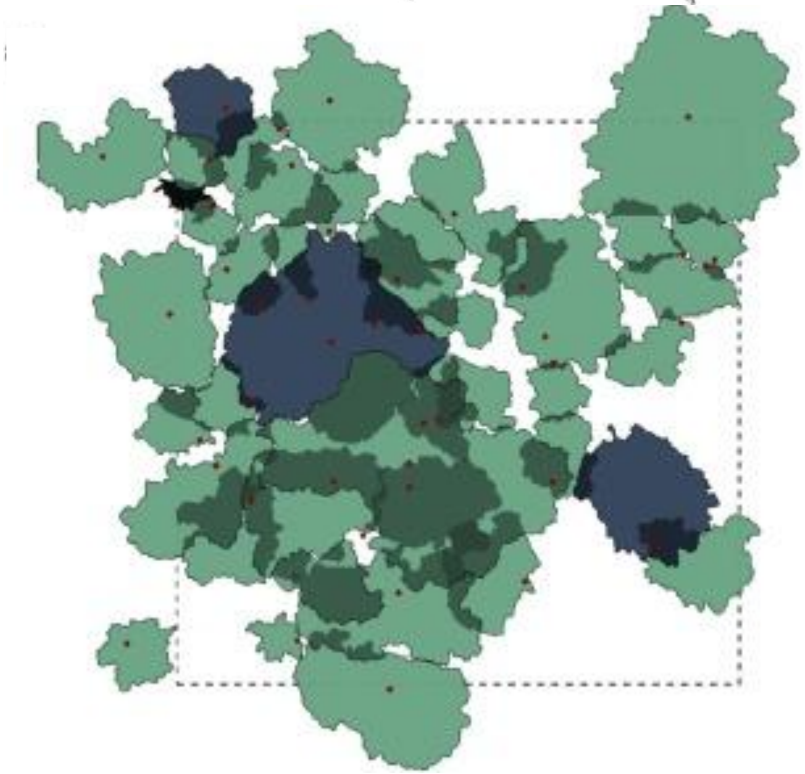


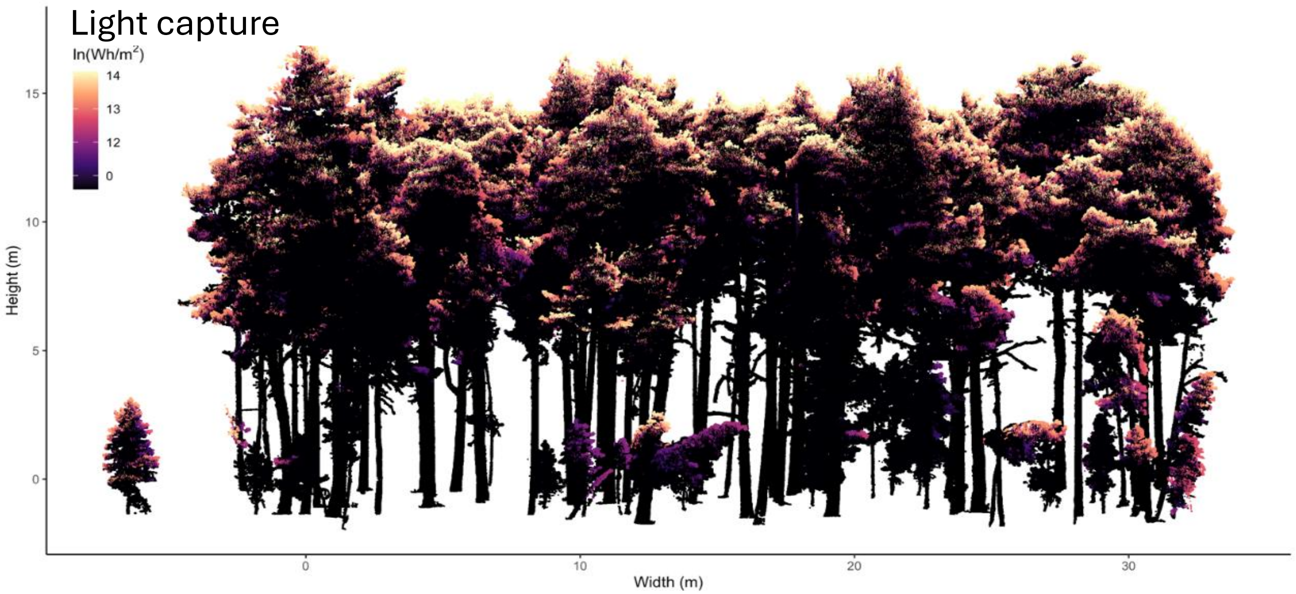
Exploring the diversity of tree and forest structures



Study structural-functional relationships directly



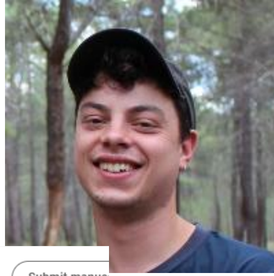




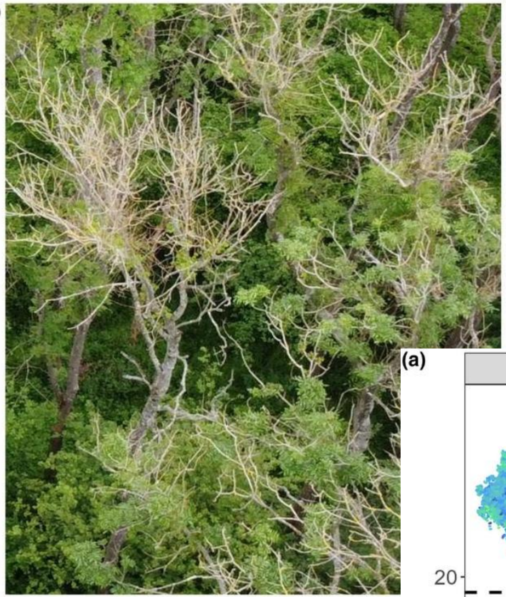
- Improved allometries for standard field measurements
- Aboveground biomass and carbon
- Leaf area and biomass allocation
- Competition
- Microhabitat and microclimate
- Realistic radiative transfer

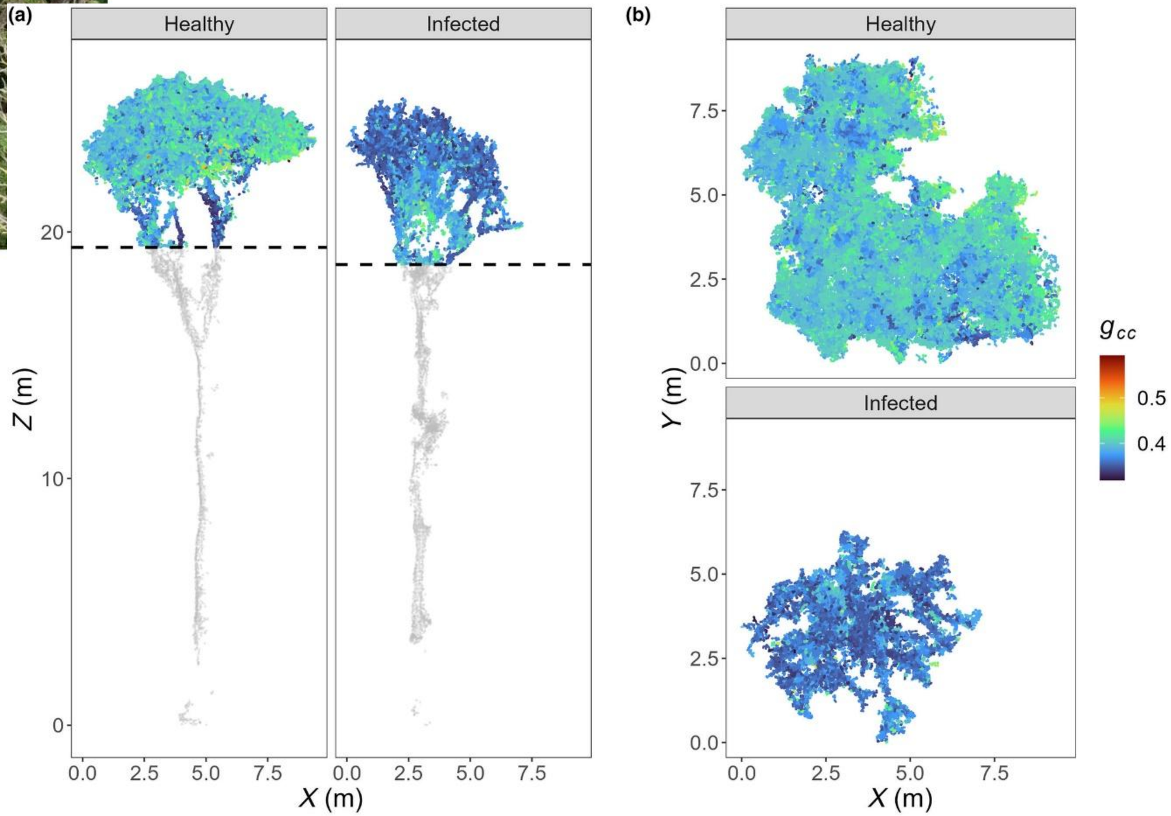
Owen et al. (2021),
Owen and Lines (2024)

Spectral-structural-temporal insights into disease



(a)



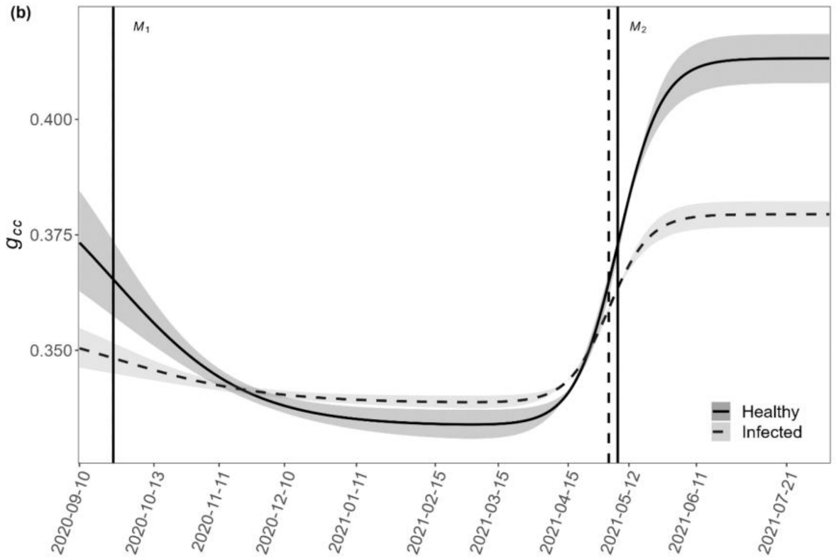


3D path length analysis shows characteristic greening of ash dieback infected trees


Science
Current Issue First release papers Archive About Submit manus
RESEARCH ARTICLE ADAPTATION

Rapid polygenic adaptation in a wild population of ash trees under a novel fungal epidemic

CAREY L. METHERINGHAM, WILLIAM J. PLUMB, WILLIAM R. M. FLYNN, JONATHAN J. STOCKS, LAURA J. KELLY, MIGUEL NEMESIO GORRIZ, STUART W. D. GRIEVE, JUSTIN MOAT, EMILY R. LINES, [...], AND RICHARD A. NICHOLS +1 authors Authors Info & Affiliations



Green-up occurred at a higher rate in trees with higher genomic estimated breeding values & less damage from disease

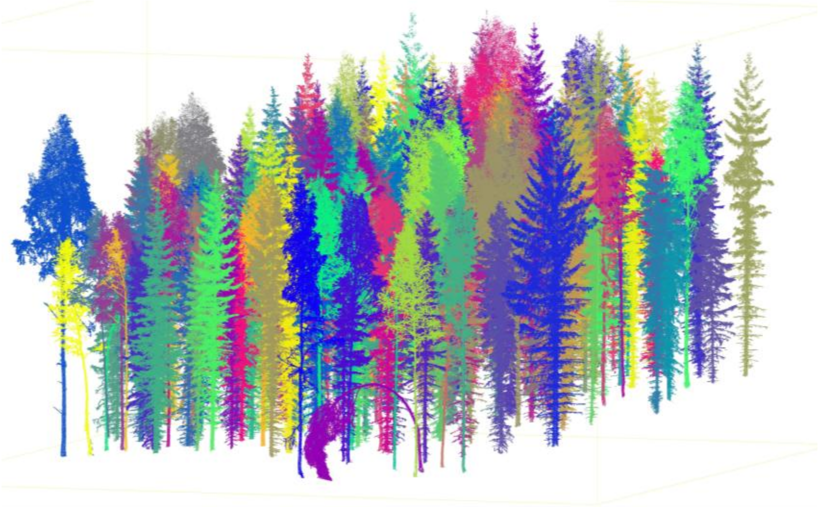
Flynn *et al.* (2024); Metheringham *et al.* (2025)

High resolution two and three-dimensional data offer a route towards joint carbon, biodiversity and ecosystem functioning monitoring (direct and proxy) from single data sources.

To do this at scale, we *probably* need deep learning (not yet proven).

With a little more thought we will reach trustworthy implementation.





Research Group

Anne Baranger

Tom Ratsakatika

Harry Leung

Desmond Teh

Ellie Kent

Thank you

linesresearchgroup.com erl27@cam.ac.uk

Group alumni
Matt Allen
Amandine Debus
Will Flynn
Milto Miltiadou
Harry Owen

(some of)
Our collaborators on this work
Stuart Grieve
Stefano Puliti
Paloma Ruiz-Benito
Phil Wilkes

WET WOODLANDS RESEARCH

UKRI Future Leaders Fellowships

UKRI Natural Environment Research Council

UNIVERSITY OF CAMBRIDGE

Centre for Landscape Regeneration

ai@cam

COST Action CA20118
3DForEcoTech
Three-dimensional forest ecosystem monitoring & better understanding by terrestrial-based technologies