Matters arising: REPLYING TO A. Macintosh et al. (2024) Communications Earth & Environment:

https://doi.org/10.1038/s43247-024-01313-x

Title: National-scale datasets systematically underestimate vegetation recovery in Australian carbon farming projects

Authors:

Tim Moore[1], Andrew O'Reilly-Nugent[2,3], Kenneth Clarke[1,4], Cathleen Waters[5,6]

1. RegenCo, Rundle Mall, Adelaide, South Australia, 5000, Australia.
2. Climate Friendly, Level 24, 60 Margaret Street, SYDNEY New South Wales 2000
3. Adjunct Fellow, School of Biological, Earth and Environmental Sciences & Ecology Research Centre, University of New South Wales
4. Adjunct Senior Fellow, School of Agriculture, Food and Wine, University of Adelaide
5. GreenCollar, The Rocks, Sydney, NSW, 2000, Australia
6. Adjunct Fellow, Hawkesbury Institute for the Environment, Western Sydney University, Richmond, NSW, Australia

Corresponding author: tim.moore@regenco.earth

Word Limit: 1200; Word count: 1166 (not including Figures, Tables, References)

Figure limit: 2; No. figures: 2

Reference limit 15: No. references: 14

Competing interests' statement: the authors receive financial remuneration for administering carbon farming projects and monitoring their performance.

Limiting global warming below 2°C requires nature-based climate solutions which are expected to supply more than a third of cost-effective climate mitigation by 2030[1,2]. Regenerating native forests under the Australian Government's Australian Carbon Credit Unit (ACCU) program are delivering large-scale carbon storage across approximately 3.4 million hectares. Projects using the Human-induced regeneration (HIR) method[3] aim to restore native forests through improved land management, generating ACCUs that underpin both legislated emissions reduction and voluntary decarbonisation targets. Scientific rigour and transparency must underpin the integrity of the ACCU program. Constructive critiques of carbon crediting programs allow refinement over time, strengthening climate action[4]. However, flawed analyses can undermine investment decisions and diminish real outcomes when it impacts critical policy decisions[5.]

Macintosh et al. (2024[6]; hereafter Macintosh) contend that HIR activities are having limited influence on changes in woody vegetation cover in Australia. Macintosh analysed the National Forest and Sparse Woody Vegetation Dataset (NFSWVD[7]) and, elsewhere[8], Woody Cover Fraction (WCF[9]) to compare vegetation trends between credited HIR areas and adjacent comparison areas. Here, we show their assessment relies on two flawed assumptions:

1. that publicly available, national-scale datasets can accurately detect and quantify vegetation cover at the scale of individual projects, and
2. that adjacent comparison areas represent valid experimental controls.

We provide high quality reference data, collected on HIR projects as standard practice, as empirical evidence that these national-scale datasets systematically underestimate regeneration success on

HIR projects, and are therefore not fit for purpose as used by Macintosh. We also demonstrate that Macintosh' s experimental design undermines and does not support their stated conclusions.

1. NFSWVD is not fit-for-purpose as a standalone tool to examine HIR project performance

We assessed the accuracy of publicly available, national-scale vegetation datasets (NFSWVD and WCF; Supplementary Materials S1) using a reference dataset of high-resolution airborne lidar across representative HIR project sites in the Queensland, New South Wales and Western Australia rangelands (Supplementary Materials S2). Our analysis used randomised 75 x 75 m validation plots from five independent lidar surveys, covering approximately 59,000 ha and including credited and non-credited areas. Direct comparison against contemporary national-scale estimates demonstrated systematic underestimation of vegetation condition by the national data sets relative to the lidar reference data.

NFSWVD detected 40 % less sparse woody vegetation than reference lidar (Figure 2a), precisely the early stage regeneration that characterises successful HIR interventions. Furthermore, NFSWVD predicted sparse woody vegetation in different locations than the reference lidar data (Table 1). NFSWVD failed to detect four out of five (4/5) validation plots with predominantly sparse vegetation (omission) and only one in four (1/4) areas mapped by NFSWVD as sparsely wooded was classified correctly (commission). Commission errors in sparse woody vegetation predominantly occur by misclassifying forest as sparse woody vegetation, rather than by overestimating vegetation where there is none. More than half of areas classified as bare ground contained measurable tree cover above Australia's 2-meter forest height threshold (Figure 2a; Table 1). Finally, the overall accuracy is only 56.4 %, well below the regulated 85% accuracy threshold required for HIR project monitoring [10].

Similarly, WCF failed to identify most vegetation below 20 % canopy cover, essentially missing early-stage regeneration (Figure 2b, c). Overall, WCF was moderately correlated with reference lidar (R^2 = 0.6), particularly at high canopy cover, however the magnitude of relative errors (126 % relRMSD) indicates substantial disagreement overall.

In summary, the publicly available datasets tended not to recognise regenerating woody with canopy cover < 20 % (NFSWVD and WCF), and commonly misrepresented forest at > 20 % cover as sparse woody vegetation (NFSWVD) when compared to lidar derived measurements. These findings invalidate MacIntosh's assumption that the national-scale datasets are appropriate for project level assessment of vegetation cover and change over time and support a previous ruling that national-scale data are not appropriate as standalone tools for the assessment of HIR projects[11].

ii. Issues with experimental design, analysis and presentation of data

Macintosh attempted to compare NFSWVD forest cover trends within credited areas of HIR projects and surrounding comparison areas extending up to three kilometres from the project perimeters. However, using fixed-width buffers resulted in dramatically mismatched comparison areas, ranging from 4% to 926% of the credited areas size (Supplementary Material 3).

Robust impact evaluation requires the careful section of control sites that match initial vegetation condition, soil properties, hydrological regimes, fire history, historical and current land management practices[12]. Macintosh's analysis, conducted within a region of high environmental variability[13], is flawed due to the assumption that geographic proximity adequately controls for these confounding factors. The use of poorly matched comparison areas (differing in size and character) violates basic principles of experimental design and introduces the likely risk of substantial bias. Comparison areas contain land unsuitable for forest growth, already heavily forested, and/or ineligible for undertaking

HIR activities, preventing analysis of a direction of bias. Macintosh obscured this variability by presenting their findings without confidence intervals or uncertainties (their Figure 3) but it is evident when the data are presented differently (Figure 2).

Finally, the statistical analysis treats year-to-year changes in NFSWVD in projects and adjacent areas as independent samples. Given the variable registration dates of the project cohort, the dataset is artificially inflated and heavily weighted toward early project stages, with very few observations of projects over ten years old. This unbalanced dataset likely underestimates the effect of HIR activities when vegetation change is expected to be gradual and non-linear over time. The absence of a detectable effect, regardless of measurement approach, reflects compounding methodological limitations rather than actual project performance.

A ’fit-for-purpose’ approach

Best practice in HIR project monitoring requires substantial, high-quality evidence to verify project performance[14]. Project proponents must demonstrate regeneration potential, including proof of young tree cohorts, implementation of management changes, and vegetation maps validated against independent reference data[10]. A high degree of confidence in HIR projects is essential for the ACCU program, given its role in Australia’s national mitigation strategy. Even with continued improvements in national products, project-level verification will remain essential given Australia's complex rangeland ecosystems.

While national-scale carbon monitoring presents significant challenges, advances in model-data fusion can improve assessment of vegetation change over large areas[15,]. These products remain important for tracking broad changes in landscapes and land use, and could be enhanced through integration with project-level data. We suggest that the data collected for HIR project monitoring be used as a key input to the development of Australia’s national scale forest monitoring system. We have already engaged in discussion with the Australian Government in this regard.

Effective climate action depends on transparency and scientific rigour. We have shown that Macintosh's analysis lacks the rigor needed for the reliable assessment of HIR project or overall program performance. Future policy development, including program review or new method development, should recognize the value and limitations of national monitoring systems and be cautious when considering if the information presented in Macintosh is adequate or useful. Systematic underestimation of regenerating vegetation could lead to incorrect conclusions about project effectiveness and discourage investment in nature-based climate solutions precisely when their scaling up is most crucial.

Acknowledgements

The authors thank Sam Shumack for their contributions to the analysis, and the efforts of three anonymous reviewers whose comments significantly improved the manuscript.

References

1. United Nations. COP26 Day 7: Sticking points and nature-based solutions (2022). https://www.un.org/en/climatechange/cop26-day-7-sticking-points-and-nature-based-solutions, 2022

2. B.W. Griscom, J. Adams, P.W. Ellis, R.A. Houghton, et. al. Natural climate solutions, Proc. Natl Acad. Sci. 11411645–50 (2017). https://doi.org/10.1073/pnas.1710465114.

3. Carbon Credits (Carbon Farming Initiative) (Human-Induced Regeneration of a Permanent Even-Aged Native Forest—1.1) Methodology Determination (2013). https://www.legislation.gov.au/F2013L01189/latest/text

4. Chubb, I., Bennett, A., Gorring, A., Hatfield-Dodds, S., (2022). Independent Review of ACCUs, Department of Climate Change, Energy, the Environment and Water, Canberra, December. https://www.dcceew.gov.au/sites/default/files/documents/independent-review-accu-exec-summary.pdf

5. Marcos Martinez, R., Baerenklau, K. Controlling for misclassified land use data: A post-classification latent multinomial logit approach, Remote Sensing of Environment, 170, (2015) https://doi.org/10.1016/j.rse.2015.09.025

6. Macintosh, A., Butler, D., Larraondo, P. et al. Australian human-induced native forest regeneration carbon offset projects have limited impact on changes in woody vegetation cover and carbon removals. Comm Earth Environ 5, 149 (2024). https://doi.org/10.1038/s43247-024-01313-x

7. Department of Climate Change, Energy, the Environment and Water. National Forest and Sparse Woody Vegetation Data (Version 7.0 - 2022 Release) (Commonwealth of Australia, Canberra, 2023). https://researchdata.edu.au/national-forest-sparse-2022-release/2590362

8. Macintosh A et al. Non- compliance and under-performance in Australian human-induced regeneration projects (2024) . The Rangeland Journal 46, RJ24024. https://doi:10.1071/RJ24024

9. Liao, Z., Van Dijk, A., Binbin He, B., et. al. Woody vegetation cover, height and biomass at 25-m resolution across Australia derived from multiple site, airborne and satellite observations. Int J Appl Earth Obs Geoinf 93 (2020). https://doi.org/10.1016/j.jag.2020.102209

10. Clean Energy Regulator. Guidelines on stratification, evidence and records or projects under the Human Induced Regeneration of a Permanent Even Aged Native Forest and Native Forest from Managed Regrowth methods (2019). https://cer.gov.au/node/3783

11. Emissions Reduction Assurance Committee. *Review of the Human-Induced Regeneration and Native Forest from Managed Regrowth methods*. (Commonwealth of Australia, Canberra, 2018).

12. Adams, VA., Barnes M., Pressey R. Shortfalls in Conservation Evidence: Moving from Ecological Effects of Interventions to Policy Evaluation. One Earth, 1,1 (2019). https://doi.org/10.1016/j.oneear.2019.08.017

13. Tongway, D.J. and Ludwig, J.A. Vegetation and soil patterning in semi-arid mulga lands of eastern Australia. Australian Journal of Ecology 15 (1990). https://doi.org/10.1111/j.1442-9993.1990.tb01017.x

14. Brack, C. L. Gateway Regeneration Checks for Human Induced Regeneration projects https://cer.gov.au/document_page/independent-review-gateway-checks-december-2024

15. McNicol, I.M., Keane, A., Burgess, N.D. et al. Protected areas reduce deforestation and degradation and enhance woody growth across African woodlands. Commun Earth Environ 4, 392 (2023). https://doi.org/10.1038/s43247-023-01053-4

Tables

Table 1. Confusion matrix comparing classification results between NFSWVD (rows) and reference lidar dataset (columns) for 500 validation points, with associated omission and commission errors and overall accuracy (OA).

		Reference lidar				
		Non-woody	Sparse woody	Forest	Total	Commission
NFSWVD	Non-woody	131	97	42	270	51.5%
	Sparse woody	4	22	55	81	72.8%
	Forest	6	14	129	149	13.4%
	Total	141	133	226	500	
	Omission	7.1%	83.5%	42.9%		OA: 56.4%

Figures

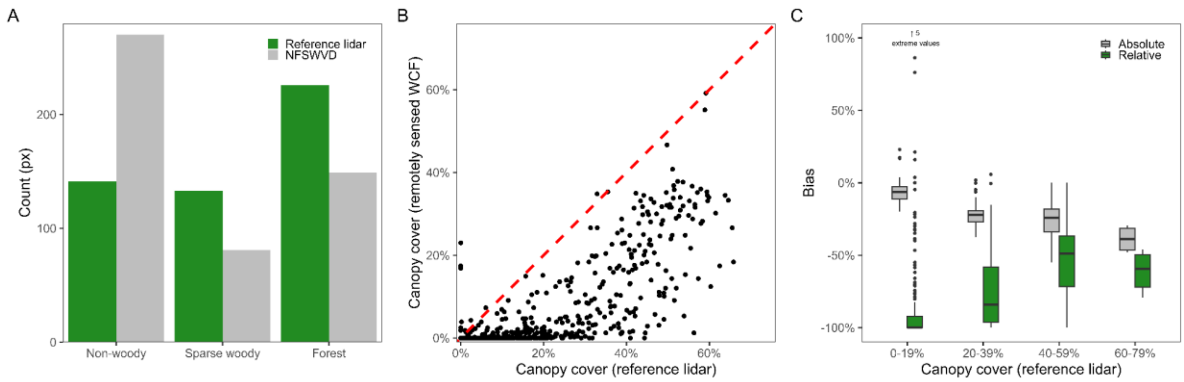

Figure 1. Canopy cover characterization and accuracy assessment. (A) Comparison of class frequencies for NFSWVD against reference lidar. (B) Comparison of WCF canopy cover against reference lidar. The dashed 1:1 line indicates perfect agreement. (C) Absolute and relative bias of WCF by canopy cover.

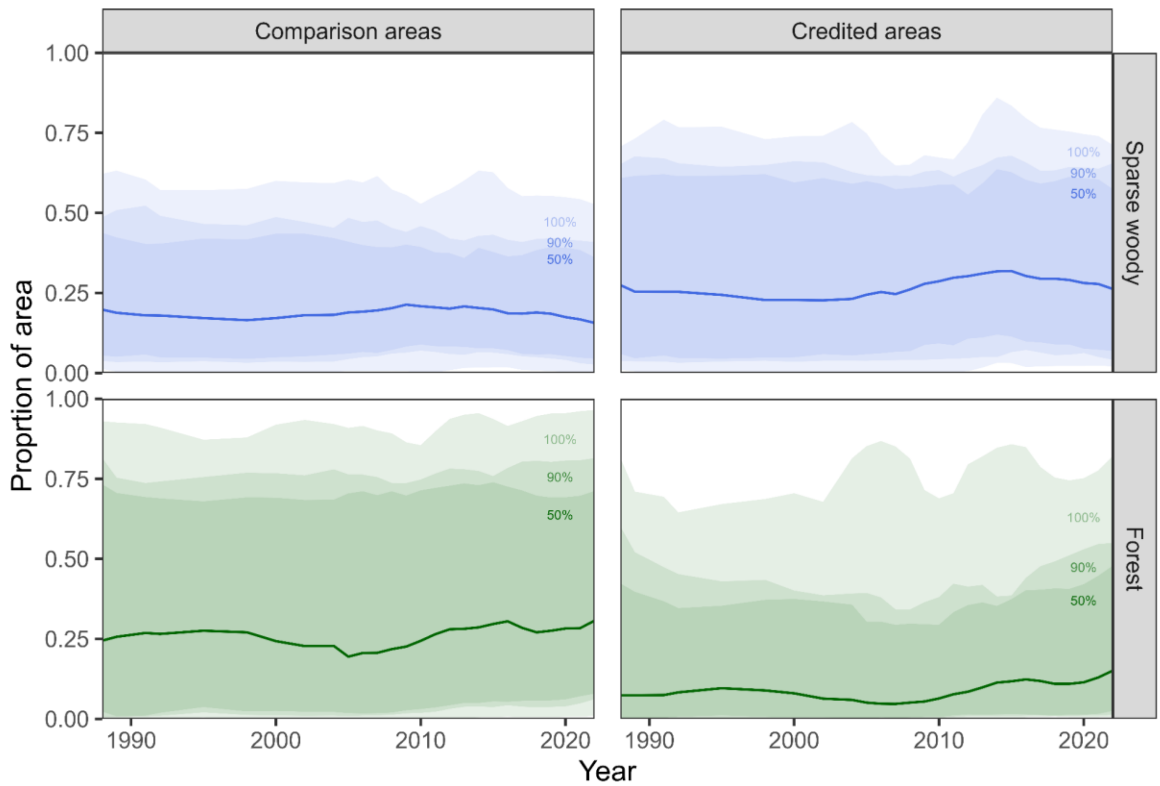

Figure 2. Median trends (solid lines) and quantiles (50%, 90% and 100%) of NFSWVD forest and sparse woody vegetation proportions within credited and comparison. Reproduced with data from Macintosh et al (2024).