

Assessing Integrity Using Vegetation Structure and Composition

Megan J McNellie (✉ megan.mcnellie@environment.nsw.gov.au)

New South Wales Department of Industry: New South Wales Department of Planning Industry and Environment <https://orcid.org/0000-0002-8162-2551>

Josh Dorrrough

NSW Department of Planning Industry and Environment

Ian Oliver

NSW Department of Planning Industry and Environment

Jian DL Yen

Arthur Rylah Institute for Environmental Research

Simon Ferrier

CSIRO Land and Water

Graeme Newell

Arthur Rylah Institute for Environmental Research

Philip Gibbons

Australian National University

Research Article

Keywords: predictive modelling, site-based floristic records, spatially-explicit vegetation models, vegetation composition, vegetation indicators, vegetation structure, vegetation integrity

Posted Date: December 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1125611/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Context

The draft post-2020 Global Biodiversity Framework aims to achieve a 15% net gain in the area, connectivity and integrity of natural systems by 2050.

Objectives

First, we analyse the complexity (foliage cover) and composition (native species richness) of 6 plant functional groups relative to their empirically defined benchmark. Second, we extrapolate the spatial patterns in foliage cover and species richness to predict where different plant functional groups are above or below benchmark as spatially-explicit, continuous characteristics across the landscape.

Methods

We assess the integrity of vegetation relative to a numerical benchmark using the log of the response ratio (LRR) to reflect the proportional change in the response variable. We use ensembles of artificial neural networks to build spatially-explicit, continuous, landscape-scale models of cover and species richness to assess locations where functional groups meet or exceed benchmarks.

Results

Models of vegetation cover LRR performed well (R^2 0.79 – 0.88), whereas models of the vegetation richness LRR were more variable (R^2 0.57 – 0.80). Predicted patterns show that across the landscape (11.5 million ha), there is a larger area that meets or exceeds the cover benchmarks (approximately 112 000 ha or 1%), and an order of magnitude lower (approximately 10 000 ha or 0.1%) for richness benchmarks.

Conclusions

Spatially explicit maps of vegetation integrity can provide important information to complement assessments of area and connectivity. Our results highlight that net gains in the area, connectivity and integrity of ecosystems will require significant investment in restoration.

Introduction

The draft post-2020 Global Biodiversity Framework embarks on an ambitious global agenda to preserve and protect biodiversity (CBD 2020). The Framework aims for a “*net gain in the area, connectivity and integrity of natural systems of at least 5%*” by 2030 and that “*an increase of at least 15%*” by 2050. Furthermore, the Framework aims to ensure that “*at least 20% of degraded ecosystems are under restoration*”. Elements like area, connectivity and configuration are critical for conservation (Woodley et al. 2019) and relatively simple to assess, yet superficial (Barnes et al. 2018; Pressey et al. 2021) because

for some species, the quality, condition or integrity of patches within the landscape is more important than their geometry (Armstrong et al. 2021; Fleishman et al. 2002; Robles and Ciudad 2012). However, assessing and enhancing integrity is multifaceted. Often, historical reference states are regarded as the baseline from which we assess integrity. However, Rohwer and Marris (2021) state “*there is no objective, empirically measured property of integrity in ecosystems that can be shown to be high in 1850 and low in 2020*” (where the year 1850 is flagged as a pre-colonisation historical reference state in North America).

We address three shortfalls in assessing and subsequently enhancing integrity. First, we assess integrity using values such as structure (e.g., foliage cover) and composition (e.g., native species richness) (Rohwer and Marris 2021) so that enhancing integrity can be focused and targeted. Second, we quantify integrity relative to a measurable and falsifiable reference state. Third, we mapped areas that meet or exceed benchmarks as potential areas to secure a post-2020 net gain in integrity (Clements et al. 2019; De Vos and Cumming 2019; Maxwell et al. 2020; Pavlacky et al. 2021) and demonstrate these areas are both inside and outside the existing protected area network.

We treat vegetation as discrete plant functional groups defined by native species within 6 growth forms—tree, shrub, grass and grass-like (hereafter referred to as grass), forb, fern and remaining ‘others’ (Oliver et al. 2019). Our first objective was to analyse complexity (as summed foliage cover of each functional group) and composition (native species richness of each functional group) relative to their empirically defined benchmark. The second objective was to extrapolate the spatial patterns in foliage cover and species richness to predict where different plant functional groups are above or below benchmark across the landscape as spatially-explicit and continuous characteristics. These data-driven models show where and how much of the landscape can be regarded as being above or below benchmark, as opposed to inferential models that rely on an indirect interpretation of disturbance patterns such as land use (e.g., Newbold et al. 2016) or human footprint (e.g., Venter et al. 2016).

The status of different plant functional groups within ecosystems offers a multi-species specificity that we argue cannot be achieved with existing approaches to vegetation extent mapping, community type mapping or inferential habitat integrity or vegetation condition modelling. The latter attempt to simplify complex multi-attributes into a single index and in so doing can obscure important structural and compositional patterns within and among ecosystems across the landscape. Maps of vegetation cover and richness can also provide a direct measure of the intrinsic quality of landscapes, as opposed to extrinsic approaches proposed by Cook et al. (2019) which consider the shape, size, funding for management, pressures such as human footprint and adjacent land uses. When plant functional groups are considered relative to their benchmarks, these estimates of vegetation state provide a comprehensive and consistent approach to mapping current integrity, quality or condition. Spatial maps of the integrity, quality or condition of ecosystems will make an important contribution to assessing the protected area network to support global efforts to conserve and restore biodiversity post-2020.

Methods

Study area

The 11 519 052 hectare study area is located in north-eastern New South Wales (NSW), Australia. The area is dominated by privately owned land used for agriculture (52% land used for grazing and 23% used for cultivation, including irrigated cropping). Less than 10% of the study area is within the protected area network (757 217 ha) or travelling stock reserves (215 852 ha). Vegetation formations are diverse, ranging from Closed Rainforests in the eastern, elevated (1 510 m) regions to Arid Shrublands and Grasslands on the drier, western plains (Keith 2004).

Figure 1 outlines the 5

Scored site data – the response variables

We extend the dataset developed by McNellie et al. (2021) to include every native species allocated to one of 6 growth forms—tree, shrub, grass and grass-like (hereafter referred to as grass), forb, fern and remaining ‘others’ (Oliver et al. 2019). For these 6 different functional groups, we compare summed foliage cover and species richness to their empirical benchmarks (Yen et al. 2019) to describe the contemporary Best-on-Offer reference states (McNellie et al. 2020). In total, we used 12 response variables based on the summed cover and summed richness of native species in each of these 6 plant functional groups.

To score the structure and composition of each functional group, we compared the observed values of either summed cover or summed richness to a benchmark value (Yen et al. 2019). Benchmarks describe the contemporary reference state to which sites are compared. These benchmarks are the numerical context needed to define site-scaled biodiversity values for the contemporary Best-on-Offer reference state (McNellie et al. 2020). We used the log of the response ratio (LRR) to express the difference between observed and benchmark values: $LRR = \ln(\text{observed} / \text{benchmark})$ (Equation 1). The log of the response ratio reflects the proportional change in the response variable (Hedges et al. 1999) and is a useful measure to assess the strength or effect size of different processes among different taxa across different environmental gradients (Osenberg et al. 1997) where the expected magnitude of the difference is large. Negative LRR values indicate observations below the benchmark level, values equal to zero indicate observations at the benchmark level, and positive values indicate observations that exceed the benchmark level. Prior to calculating the LRR metric we added 0.01 to all observed values and benchmark values to avoid logarithms of zero values. We modelled summed cover and summed richness for each plant functional group separately. Background points (see Supp Material S3) were attributed with the lowest observed LRR value which was -9.747 for cover and -7.741 for richness functional groups. The dataset used to train the models contained 6 806 sites and an additional 2 462 background points. A subset of 3 021 floristic sites that matched the temporal period of the remotely sensed variables used to predict cover (2005 to 2012). We selected a random subset of 1 015 background points (Figure 1 – Step 1).

Environmental And Disturbance Variables – The Predictor Surfaces

Our choice of potential predictors was guided by ecologically informed expert-judgement. We prepared 29 potential surfaces that described patterns in macroclimate, soils and geology, topography, landscape modification and remote sensing (Figure 1 – Step 2 and Supplementary Material S1).

Prior to training the models, we tested for correlation by calculating Pearson's correlation coefficient and multicollinearity using the variance inflation factor (VIF) with a conservative threshold of 5 (O'Brien 2007) (Supplementary Material S2). This iterative approach to selecting the suite of predictor surfaces is described in McNellie et al. (2021).

The final suite of predictors reflects (i) macroclimatic environmental attributes that directly, and geological and topographic surfaces that indirectly influence the growth and morphology of vegetation (Austin 2002; Box 1981; Franklin 2009; Guisan and Zimmermann 2000; Pressey et al. 2000) and (ii) an *a priori* assessment of disturbance variables that modify and fragment vegetation (Table 1).

Table 1
 Predictor surfaces used to describe the environmental and disturbance characteristics of the study area

Environmental gradient	Predictor surface
Macroclimatic	isothermality
	rainfall seasonality ratio
	precipitation/evaporation
Soil and Geology	gravity
	% clay
	great soil group
Radiometric	thorium/potassium
Topographic	slope
	transformed aspect
	compound topographic index
	distance to drainage
	distance to wetland
Modification	land use
Landsat imagery	foliage projective cover
	normalised difference vegetation index
	bare ground

Modelling framework

Artificial neural network (ANN) models are advantageous for ecological applications where data do not meet parametric statistical assumptions and relationships between response and predictor data are complex, unknown or non-linear (Bishop 1995; Fielding 1999; McNellie et al. 2021). We chose multi-target ensembles of ANN (Statistica v10) (Statsoft Inc. 2013) because these are effective for resolving complex predictions and can handle redundant or co-linear predictor variables. ANN models can extrapolate beyond the range of site-based values for response variables which is more useful in predicting lower and higher ends of the data distribution (Heikkinen et al. 2012) and have been used to predict spatially-explicit characteristics of plant functional groups (McNellie et al. 2021).

Model Evaluation

The predictive performance of each ensemble model was evaluated by calculating the coefficient of determination (R^2) which we used to assess the strength of the relationships between predicted and observed values for the training, testing and hold-out subsets for each modelled condition attribute. It is important to note that model performance was judged by determining how well the model performed when applied to new data (the out-of-sample, validation or hold-out subset). Parity between the R^2 for the training and hold-out subsets indicates how well the model has been trained. We used the root mean squared deviation (RMSD) and the mean absolute error (MAE) to quantify the deviation from the 1:1 line. (see Supplementary S3). Both error estimates report errors on the same scale as the input data. We calculated Moran's Index from the model residuals to determine whether residuals were spatially autocorrelated (implemented in Spatial Statistics toolbox ArcGIS v10.4) (Figure 1 – Step 3).

Predicting condition across the whole landscape

We used ensembles of 25 ANN models to predict spatial patterns of vegetation condition across the entire landscape. Trained models were deployed to every grid cell in the prediction matrix to derive an estimate of vegetation condition in previously unobserved locations. These analyses produced 25 prediction models. The final predicted output for each grid cell was averaged to create a single ensemble model for each functional group.

We used spatial analysis to identify pixels at benchmark (LRR = 0) or above benchmark value (LRR > 0) for each functional group. We then calculated the coincident area where LRR equalled or exceeded benchmark value for all functional groups, with separate values for structure and composition. Finally, we calculated the area and locations where vegetation was at or above benchmark values for both structure and composition and we interrogate how much of this area is located within a protected area (Figure 1 – Steps 4 and 5).

Results

Model assessment

For cover variables, R^2 for the hold-out subset ranged from 0.88 (integrity of forb cover) to 0.79 (integrity of fern cover) and R^2 for richness variables ranged from 0.80 (integrity of tree and forb richness) to 0.57 (integrity of fern and other richness) (Table 2). The RMSD for the models of cover condition ranged from 2.50 (fern cover) to 1.48 (forb cover); for the condition of functional group richness, RMSD ranged from 1.50 (condition of tree and grass richness) to 2.32 (condition of fern richness) (Table 2). MAE for cover condition models ranged from 0.81 (condition of tree cover) to 1.77 (condition of fern cover) and for richness condition attributes MAE ranged from 0.85 (tree richness) to 1.88 (condition of fern richness) (Table 2).

Sensitivity analysis for the ensembles of ANN models indicates land use, great soil group and foliage projective cover were the three most important predictors for informing the multi-target cover models; and land use, great soil group and % clay were the 3 most important predictors for informing the multi-target richness models (Supplementary Material S4). Estimates of Moran's Index suggested model residuals

were not spatially autocorrelated for the structure models (Supplementary Material S5). However, models of tree richness were dispersed and other richness were clustered.

Table 2
Coefficient of determination (R^2) for the ensemble of random samples for condition attributes for structure and composition.

Condition Attribute	R^2_{Train}	R^2_{Test}	$R^2_{\text{Hold-out}}$	RMSD	MAE
Tree cover	0.89	0.87	0.87	1.57	0.81
Shrub cover	0.85	0.82	0.83	1.82	1.11
Grass cover	0.89	0.87	0.87	1.51	0.93
Forb cover	0.90	0.88	0.88	1.48	0.88
Fern cover	0.82	0.79	0.79	2.50	1.77
Other cover	0.82	0.79	0.80	2.46	1.72
Tree richness	0.78	0.80	0.80	1.50	0.85
Shrub richness	0.77	0.79	0.79	1.55	0.92
Grass richness	0.79	0.80	0.81	1.50	0.85
Forb richness	0.77	0.79	0.80	1.58	0.92
Fern richness	0.54	0.57	0.57	2.32	1.88
Other richness	0.54	0.57	0.57	2.08	1.48
Numbers in bold highlight results for the hold-out subset.					

The root mean squared difference (RMSD) and mean absolute error (MAE) estimates show the mean deviation of predicted cover with respect to the observed cover and predicted richness values with respect to the observed richness values. Number of observations cover models, $n = 3\ 021$; richness models, $n = 9\ 268$.

Extrapolating Spatial Patterns Across The Landscape

Table 3 shows the area for each separate functional group where the LRR equalled or exceeded the benchmark value for structure and composition. Notably, the area where vegetation meets or exceeds benchmark is far greater for structure than for composition (Table 3; see Supplementary Material S6 for maps showing spatial patterns for each functional group). Figures 2 and 3 show a snapshot of different elements of mapped vegetation to demonstrate our results. Panel a shows extant vegetation communities (NSW Office of Environment and Heritage 2017). Panel b drills through the vegetation community to show predicted shrub richness (Figure 2b) or grass cover (Figure 3b) (McNellie et al. 2021). Panel c shows the results of the modelled prediction compared to the benchmark for each vegetation

community (maps of foliage cover and species richness for every functional see Supplementary Material S6). Despite vegetation mapping showing extant and uncleared vegetation communities (Panel a of Figure 2 and Figure 3), linked by continuous patches of individual functional groups (Panel b of Figure 2 and Figure 3), the relative integrity of these different functional groups does not always meet or exceed the benchmark (Panel c of Figure 2 and Figure 3). In addition, mapped results indicate locations outside the protected area network where the composition and structure of each growth form exceed benchmark value.

When all functional groups are interrogated together, our ensemble models predicted 111 691 ha (approximately 1% of the study area) where the structure of the 6 functional groups were continuously at or above benchmark value (in the same location). Of the six functional groups, fern cover is at or above benchmark value for approximately half of the study area. In contrast, these analyses show that the composition of different functional groups meets benchmark value in far fewer locations. Our ensemble models predicted that all 6 functional groups were at or above the composition benchmarks for only 10 371 ha (< 0.1% of the study area) although both grass richness and forb richness met benchmark values in approximately 10% of the study area. When we interrogated the areas where all 12 structure and composition attributes were predicted to be at or above benchmark value, 2 227 ha coincided and the largest contiguous area was <20 ha (Table 3).

Summarising patterns in the predicted Best-on-Offer reference state across tenure

Relating spatial patterns in integrity to tenure highlighted that protected area status is a coarse indicator for predicting where foliage cover or cover species richness might exceed benchmark (Table 4). Of the 2 227 ha where vegetation was at or exceeded benchmark values for both structure and composition, approximately 75% is within a protected area, and the remaining 25% occurs outside a protected area (Table 4), scattered across the study area.

Table 3

The extent of the predicted Best-on-Offer reference state for structure and composition for each functional group, with a summary of area (ha) where all functional groups are at or above benchmark at the same locations. See Supplementary Material S6 for spatially-explicit patterns for each functional group

Predicted functional group	Area (ha) at or above benchmark	% of study area	Largest contiguous area (ha)
Tree cover	654 801	5.69	30 770
Shrub cover	529 979	4.61	9 334
Grass cover	202 908	1.76	13 270
Forb cover	1 122 347	9.78	80 455
Fern cover	5 571 024	48.58	2 885 861
Other cover	3 679 350	32.06	950 379
Cover for all functional groups	111 691	0.97	1 635
Tree richness	201 122	1.75	8.279
Shrub richness	376 383	3.27	62 185
Grass richness	1 106 639	9.63	321 179
Forb richness	1 087 975	9.47	265 460
Fern richness	22 430	0.19	451
Other richness	197 894	1.72	49 877
Richness for all functional groups	10 371	0.09	448
Structure and composition	2 227	0.02	19

Table 4

Summarising patterns in the predicted Best-on-Offer (BOO) reference state within protected areas and travelling stock reserves, compared to other tenures (including private land).

Predicted	Protected Area		Travelling Stock Reserve		Other	
	Area (ha) at or above benchmark	% of BOO	Area (ha) at or above benchmark	% of BOO	Area (ha) at or above benchmark	% of BOO
Coincident area with structure \geq benchmark (111 691 ha)	17 059	15	74	< 1	93 899	84
Coincident area with composition \geq benchmark (10 371 ha)	8 755	84	<1.5	<0.1	1 615	16
Coincident area with both structure and composition \geq benchmark (2 227 ha)	1 671	76	0	0	556	24

Discussion

Globally, conservation efforts need to be more targeted and more effective. Spatially explicit estimates of vegetation integrity are an essential tool to inform ecosystem conservation and restoration planning at a landscape scale. Yet ecological integrity is an opaque concept that is difficult to assess and enhance (Brown and Williams 2016; Rohwer and Marris 2021). To disentangle integrity, we demonstrate how tangible attributes of vegetation cover and richness can be assessed relative to their transparently defined, contemporary reference states (McNellie et al. 2020). Moreover, we deliver models that provide continuous gridded surfaces at fine scales (100 m) commensurate with the key ecological processes that shape landscape patterns. These models are focussed and data-driven, as opposed to inferential models that rely on disturbance-driven patterns (such as land use or human footprint) which aim to assess and enhance integrity based on notions of pristine, intact or pre-intensification ecosystems. However, our results show that land use was substantially more influential than other environmental or disturbance variables, suggesting that land use might be an appropriate proxy in some cases. Given our assessment of integrity can be sourced from existing data, and assessed against measurable and falsifiable benchmarks, we believe our approach could be applicable to terrestrial landscapes at a global scale.

Previous landscape-scaled assessments of integrity have used national conservation reserves (e.g., Harwood et al. 2016) or distance from human populations or infrastructure (e.g., Allan et al. 2019; Watson et al. 2018) as proxies to identify reference states from which to assess condition. However, protected areas are not always minimally disturbed and are often not representative of a majority of ecosystems (Joppa and Pfaff 2009). We found that approximately 85% of vegetation that meets or exceeds the cover benchmarks is outside the current protected area network. In contrast, 85% of the vegetation that meets or exceeds the richness benchmark is contained within the protected area network.

When using a contemporary reference frame, the intrinsic value of data-based models are useful measures of integrity, quality or condition. Here, we show that some areas with the greatest potential biodiversity value are not included in the protected area network (Archibald et al. 2020; Clements et al. 2019; De Vos and Cumming 2019). Targeting these areas may yield disproportionate benefits for conservation strategies because unprotected areas potentially face some of the greatest threats (Myers et al. 2000).

Inspection of the predicted patterns shows that, across the landscape, there is a larger area that meets or exceeds benchmark values for cover (approximately 112 000 ha) than for richness (approximately 10 000 ha). In addition, we identified areas where condition exceeds benchmark values outside the existing protected area network. Protected area status appears to be a limited indicator of vegetation integrity, especially when structural and compositional attributes are considered simultaneously. This may indicate that the composition of functional groups is more degraded and that efforts to restore landscapes will need to target species composition.

Here we have highlighted that when we use complex, multi-attribute information based on species' observations, it is difficult for all attributes to meet their benchmark simultaneously. This paints a picture of poor or incomplete integrity when assessing the contemporary landscape. This may contrast with common perceptions of integrity, especially in protected areas. However, people, including experts in vegetation assessment, use different cues to perceive integrity. When assessing vegetation communities experts may tend to be biased towards dominant woody species (Dorrough et al. 2021), rather than overall structure or composition (such as foliage cover and native species richness among functional groups).

One of the benefits of the models presented here is that individual attributes have not been combined into an index of structural or compositional integrity. When used as standalone attributes, each can be used to inform different aspects of landscape conservation. An additional benefit is that log response ratios can be summed. As a result, different functional groups can be summed to assess the integrity of combined attributes, such as non-woody (grasses, forbs, ferns and 'others') or woody (trees and shrubs) components of the landscape. Furthermore, the data are on a continuous scale and can be interrogated to find patterns in vegetation integrity across the landscape for operational and practical land management and conservation, such as active restoration. Modelled surfaces can identify where in the landscape most functional groups meet benchmark levels and, therefore, can inform targeted restoration. For example, our results indicate those areas where both shrub and tree functional groups meet or exceed benchmark levels, yet active restoration may be necessary to improve the structure and composition of forbs and grasses.

Given systematic conservation planning is an inherently spatial undertaking (Pressey et al. 2000), planning can be improved when the integrity of different functional groups is considered. The maps generated here incorporate the relationship of observed states relative to contemporary reference states, which is likely to improve systematic conservation planning by identifying locations that simultaneously

protect species from multiple functional groups. These maps show different degrees of change in integrity across the landscape and focus on a multi-species approach with an emphasis on structural and compositional characteristics of vegetation. These structural and compositional attributes can also serve as reliable input variables for species habitat modelling. For example, habitat models may require tailored functional groups relevant to species-specific ecological niches (Lawton et al. 2021; Robles and Ciudad 2012; van Schalkwyk et al. 2021).

Areas that could potentially be used to augment and extend the conservation network will inevitably be located on private lands. Refocusing the Global Biodiversity Framework post-2020 could concentrate new conservation efforts on creating protected areas that are tailored to meet future biodiversity targets. We have shown that some functional groups have large contiguous areas that facilitate well-connected landscapes within functional groups.

Declarations

Acknowledgements

We are grateful to T. Eyre, M. Lewis and R. Peet for their comments on an earlier version and we thank anonymous referees for all their valuable comments on the manuscript.

Funding or Competing Interests -The authors declare that they have no known competing financial interests, competing interests or personal relationships that could have appeared to influence the work reported in this paper.

Contribution:

- conceived the ideas and designed methodology – ALL;
- proposed log response ratio (LRR) method – JD, JY;
- extracted the case study data and prepared LRR data – MJM;
- prepared all figures, tables and maps – MJM;
- led the writing of the manuscript – MJM;
- contributed critically to drafting and revising the manuscript and gave final approval for publication – ALL

Data Availability

Input data and maps will be made available via publicly archived datasets

References

1. Allan JR, Watson JEM, Di Marco M, O'Bryan CJ, Possingham HP, Atkinson SC, Venter O (2019) Hotspots of human impact on threatened terrestrial vertebrates. *PLoS Biology* 17(3):e3000158

<https://doi.org/10.1371/journal.pbio.3000158>

2. Archibald CL, Barnes MD, Tulloch AIT, Fitzsimons JA, Morrison TH, Mills M, Rhodes JR (2020) Differences among protected area governance types matter for conserving vegetation communities at-risk of loss and fragmentation. *Biol Conserv* 247:108533
<https://doi.org/10.1016/j.biocon.2020.108533>
3. Armstrong DP, Boulton RL, McArthur N, Govella S, Gorman N, Pike R, Richard Y (2021) Using experimental reintroductions to resolve the roles of habitat quality and metapopulation dynamics on patch occupancy in fragmented landscapes. *Conservation Biology* *In press* <https://doi.org/10.1111/cobi.13843>
4. Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling* 157(2-3):101-118 [https://doi.org/10.1016/s0304-3800\(02\)00205-3](https://doi.org/10.1016/s0304-3800(02)00205-3)
5. Barnes MD, Glew L, Wyborn C, Craigie ID (2018) Prevent perverse outcomes from global protected area policy. *Nature Ecology & Evolution* 2(5):759-762 <https://doi.org/10.1038/s41559-018-0501-y>
6. Bishop CM (1995) *Neural networks for pattern recognition*. Oxford University Press
7. Box EO (1981) Predicting physiognomic vegetation types with climate variables. *Vegetatio* 45(2):127-139 <https://doi.org/10.2307/20037033>
8. Brown ED, Williams BK (2016) Ecological integrity assessment as a metric of biodiversity: are we measuring what we say we are? *Biodiversity and Conservation* 25(6):1011-1035
<https://doi.org/10.1007/s10531-016-1111-0>
9. CBD (2020) Zero draft of the post-2020 global biodiversity framework. In: CDB (ed). Montreal, Canada,
10. Clements HS, Kerley GIH, Cumming GS, De Vos A, Cook CN (2019) Privately protected areas provide key opportunities for the regional persistence of large- and medium-sized mammals. *Journal of Applied Ecology* 56(3):537-546 <https://doi.org/10.1111/1365-2664.13300>
11. Cook CN, Valkan RS, McGeoch MA (2019) Beyond total area protected: A new set of metrics to measure progress in building a robust protected area estate. *Global Environmental Change* 58:101963 <https://doi.org/10.1016/j.gloenvcha.2019.101963>
12. De Vos A, Cumming GS (2019) The contribution of land tenure diversity to the spatial resilience of protected area networks. *People and Nature* 1(3):331-346 <https://doi.org/10.1002/pan3.29>
13. Dorrough J, Tozer M, Armstrong R, Summerell G, Scott ML (2021) Quantifying uncertainty in the identification of endangered ecological communities. *Conservation Science and Practice* 3(11):e537
<https://doi.org/10.1111/csp2.537>
14. Fielding A (1999) *Machine Learning Methods For Ecological Applications*. Springer US, Boston, MA
15. Fleishman E, Ray C, Sjögren-Gulve P, Boggs CL, Murphy DD (2002) Assessing the roles of patch quality, area, and isolation in predicting metapopulation dynamics. *Conservation Biology* 16(3):706-716 <https://doi.org/10.1046/j.1523-1739.2002.00539.x>

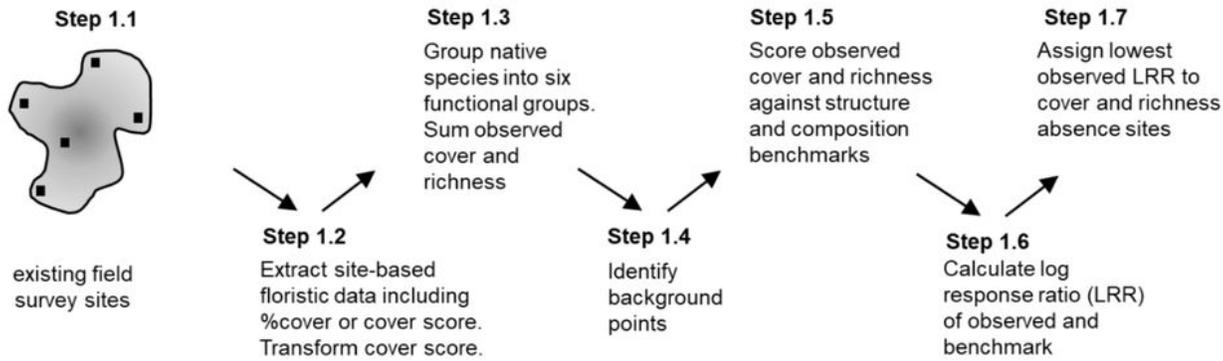
16. Franklin J (2009) Mapping Species Distributions: Spatial Inference and Prediction Cambridge University Press
17. Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* 135(2-3):147-186 [https://doi.org/10.1016/s0304-3800\(00\)00354-9](https://doi.org/10.1016/s0304-3800(00)00354-9)
18. Harwood TD, Donohue RJ, Williams KJ, Ferrier S, McVicar TR, Newell G, White M (2016) Habitat condition assessment system: a new way to assess the condition of natural habitats for terrestrial biodiversity across whole regions using remote sensing data. *Methods in Ecology and Evolution* 7(9):1050-1059 <https://doi.org/10.1111/2041-210X.12579>
19. Hedges LV, Gurevitch J, Curtis PS (1999) The meta-analysis of response ratios in experimental ecology. *Ecology* 80(4):1150-1156 [https://doi.org/10.1890/0012-9658\(1999\)080\[1150:Tmaorr\]2.0.Co;2](https://doi.org/10.1890/0012-9658(1999)080[1150:Tmaorr]2.0.Co;2)
20. Heikkinen RK, Marmion M, Luoto M (2012) Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography* 35(3):276-288 <https://doi.org/10.1111/j.1600-0587.2011.06999.x>
21. Joppa LN, Pfaff A (2009) High and far: biases in the location of protected areas. *PLoS ONE* 4(12):e8273 <https://doi.org/10.1371/journal.pone.0008273>
22. Keith D (2004) Ocean Shores to Desert Dunes: The Native Vegetation of New South Wales and the ACT. Department of Environment and Conservation NSW, Sydney
23. Lawton JA, Holland GJ, Bennett AF (2021) What determines the distribution of a threatened species, the brush-tailed phascogale *Phascogale tapoatafa* (Marsupialia: Dasyuridae), in a highly modified region? *Austral Ecology* 46(8):1404-1417 <https://doi.org/10.1111/aec.13094>
24. Maxwell SL, Cazalis V, Dudley N, Hoffmann M, Rodrigues ASL, Stolton S, Visconti P, Woodley S, Kingston N, Lewis E, Maron M, Strassburg BBN, Wenger A, Jonas HD, Venter O, Watson JEM (2020) Area-based conservation in the twenty-first century. *Nature* 586(7828):217-227 <https://doi.org/10.1038/s41586-020-2773-z>
25. McNellie MJ, Oliver I, Dorrrough J, Ferrier S, Newell G, Gibbons P (2020) Reference state and benchmark concepts for better biodiversity conservation in contemporary ecosystems. *Global Change Biology* 26(12):6702-6714 <https://doi.org/10.1111/gcb.15383>
26. McNellie MJ, Oliver I, Ferrier S, Newell G, Manion G, Griffioen P, White M, Koen T, Somerville M, Gibbons P (2021) Extending vegetation site data and ensemble models to predict patterns of foliage cover and species richness for plant functional groups. *Landscape Ecology* 36(5):1391-1407 <https://doi.org/10.1007/s10980-021-01221-x>
27. Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403(6772):853-858 <https://doi.org/10.1038/35002501>
28. Newbold T, Hudson LN, Arnell AP, Contu S, De Palma A, Ferrier S, Hill SLL, Hoskins AJ, Lysenko I, Phillips HRP, Burton VJ, Chng CWT, Emerson S, Gao D, Pask-Hale G, Hutton J, Jung M, Sanchez-Ortiz K, Simmons BI, Whitmee S, Zhang H, Scharlemann JPW, Purvis A (2016) Has land use pushed

- terrestrial biodiversity beyond the planetary boundary? A global assessment. *Science* 353(6296):288
<https://doi.org/10.1126/science.aaf2201>
29. NSW Office of Environment and Heritage (2017) The NSW State Vegetation Type Map: Methodology for a regional-scale map of NSW plant community types. In: NSW Office of Environment and Heritage (ed). Sydney, Australia.,
 30. O'Brien RM (2007) A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity* 41(5):673-690 <https://doi.org/10.1007/s11135-006-9018-6>
 31. Oliver I, McNellie MJ, Steenbeeke G, Copeland L, Porteners MF, Wall J (2019) Expert allocation of primary growth form to the NSW flora underpins the Biodiversity Assessment Method. *Australasian Journal of Environmental Management* 26(2):124-136
<https://doi.org/10.1080/14486563.2019.1595186>
 32. Osenberg Craig W, Sarnelle O, Cooper Scott D (1997) Effect size in ecological experiments: the application of biological models in meta-analysis. *The American Naturalist* 150(6):798-812
<https://doi.org/10.1086/286095>
 33. Pavlacky DC, Jr., Hagen CA, Bartuszevige AM, Iovanna R, George TL, Naugle DE (2021) Scaling up private land conservation to meet recovery goals for grassland birds. *Conservation Biology* 35(5):1564-1574 <https://doi.org/10.1111/cobi.13731>
 34. Pressey RL, Hager TC, Ryan KM, Schwarz J, Wall S, Ferrier S, Creaser PM (2000) Using abiotic data for conservation assessments over extensive regions: quantitative methods applied across New South Wales, Australia. *Biol Conserv* 96(1):55-82 [https://doi.org/10.1016/s0006-3207\(00\)00050-1](https://doi.org/10.1016/s0006-3207(00)00050-1)
 35. Pressey RL, Visconti P, McKinnon MC, Gurney GG, Barnes MD, Glew L, Maron M (2021) The mismeasure of conservation. *Trends in Ecology & Evolution* 36(9):808-821
<https://doi.org/10.1016/j.tree.2021.06.008>
 36. Robles H, Ciudad C (2012) Influence of habitat quality, population size, patch size, and connectivity on patch-occupancy dynamics of the middle spotted woodpecker. *Conservation Biology* 26(2):284-293 <https://doi.org/10.1111/j.1523-1739.2011.01816.x>
 37. Rohwer Y, Marris E (2021) Ecosystem integrity is neither real nor valuable. *Conservation Science and Practice* 3(4):e411 <https://doi.org/10.1111/csp2.411>
 38. Statsoft Inc. (2013) *Electronic Statistics Textbook*. <http://www.statsoft.com/textbook/>, Tulsa, Oklahoma,
 39. van Schalkwyk J, Gaigher R, Pryke JS, Samways MJ (2021) Within-corridor heterogeneity is more important than corridor design for maintaining butterfly functional and taxonomic diversity. *Journal of Applied Ecology* *In press* <https://doi.org/10.1111/1365-2664.14006>
 40. Venter O, Sanderson EW, Magrath A, Allan JR, Behr J, Jones KR, Possingham HP, Laurance WF, Wood P, Fekete BM, Levy MA, Watson JEM (2016) Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nature Communications* 7(1):12558
<https://doi.org/10.1038/ncomms12558>

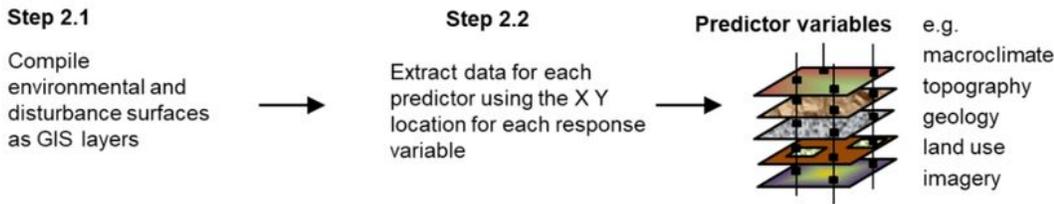
41. Watson JEM, Evans T, Venter O, Williams B, Tulloch A, Stewart C, Thompson I, Ray JC, Murray K, Salazar A, McAlpine C, Potapov P, Walston J, Robinson JG, Painter M, Wilkie D, Filardi C, Laurance WF, Houghton RA, Maxwell S, Grantham H, Samper C, Wang S, Laestadius L, Runting RK, Silva-Chávez GA, Ervin J, Lindenmayer D (2018) The exceptional value of intact forest ecosystems. *Nature Ecology & Evolution* 2(4):599-610 <https://doi.org/10.1038/s41559-018-0490-x>
42. Woodley S, Locke H, Laffoley D, MacKinnon K, Sandwith T, Smart J (2019) A review of evidence for area-based conservation targets for the post-2020 Global Biodiversity Framework. *PARKS: The International Journal of Protected Areas and Conservation* 25(2):31-46
<https://doi.org/10.2305/IUCN.CH.2019.PARKS-25-2en>
43. Yen JDL, Dorrough J, Oliver I, Somerville M, McNellie MJ, Watson CJ, Vesk PA (2019) Modeling biodiversity benchmarks in variable environments. *Ecological Applications* 29(7):1-16
<https://doi.org/10.1002/eap.1970>

Figures

STEP 1. Score sites relative to cover and richness benchmarks – the response variables



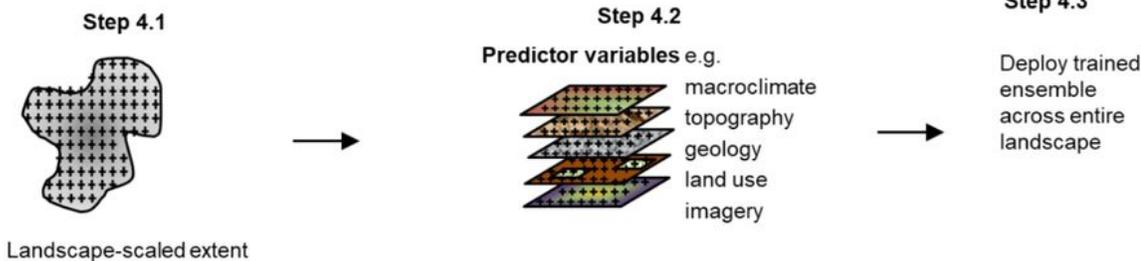
STEP 2. Assimilate environmental and disturbance gradients – the predictor surfaces



STEP 3. Train multi-target artificial neural network models



STEP 4. Predict across the landscape



STEP 5. Represent model predictions and residual error as raster maps

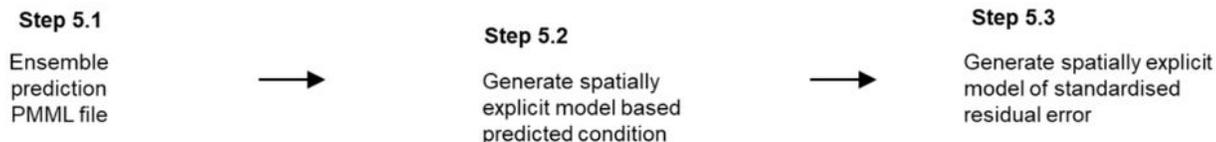


Figure 1

Outline of the five components of this study: (i) classifying site-based floristic data into plant functional groups and scoring structure and composition relative to Best-on-Offer benchmarks; (ii) sourcing spatial layers that describe environmental and disturbance gradients; (iii) training multi-model artificial neural networks (ANN) models; (iv) using the trained ANN models to predict structure and composition of plant functional groups across the whole landscape; and (v) rendering the average results from ensembles of

predicted models into a spatially-explicit map for each attribute, accompanied by the standardised residual error. ANN - Artificial neural network; LRR – log of the response ratio; PMML - Predictive Model Markup Language

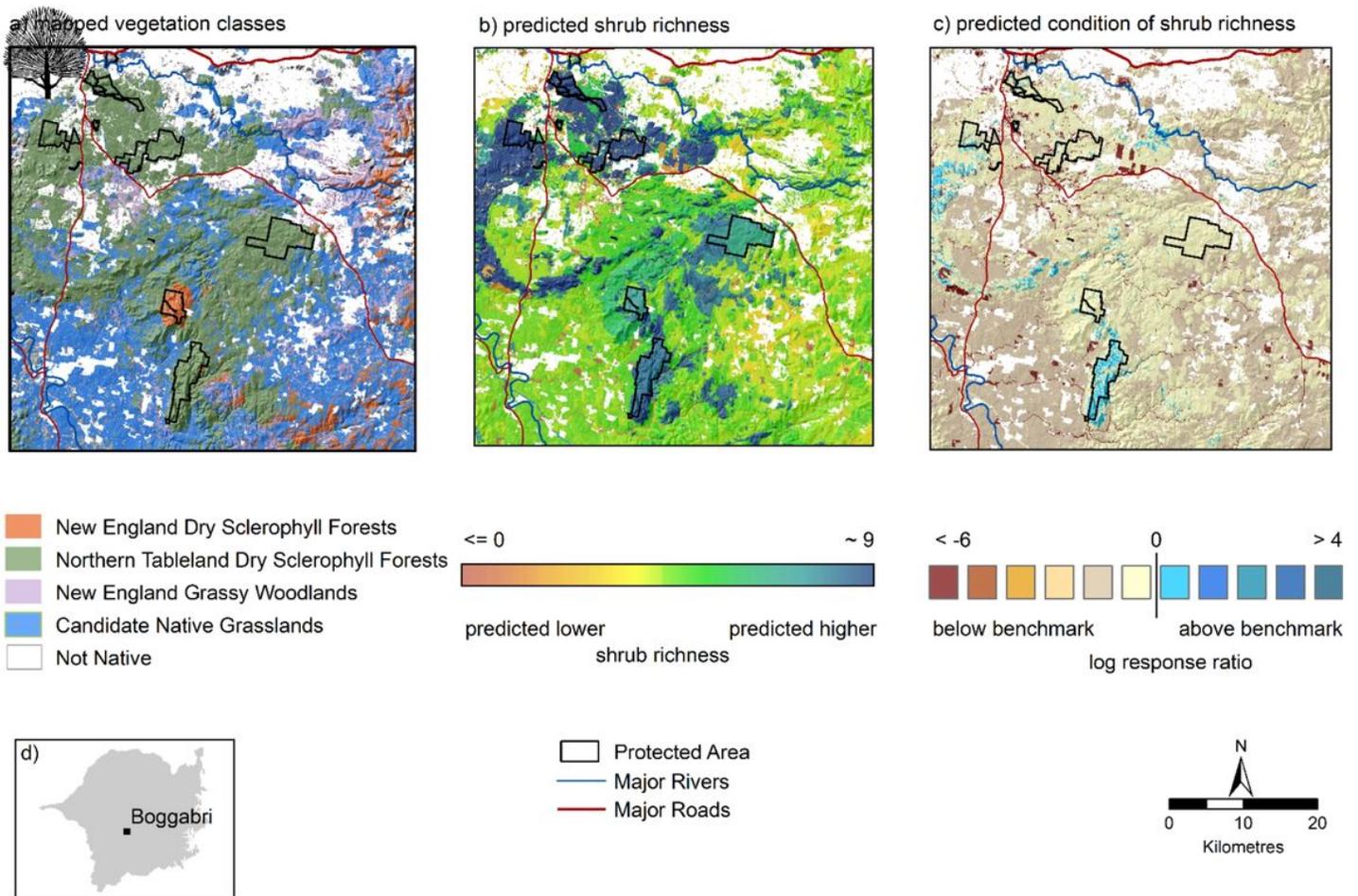


Figure 2

a) vegetation mapping showing extant vegetation communities; b) a fine-scale and continuous model of shrub richness (from McNellie et al. 2021); c) the integrity of shrub richness relative to benchmark; d) location diagram within the study area

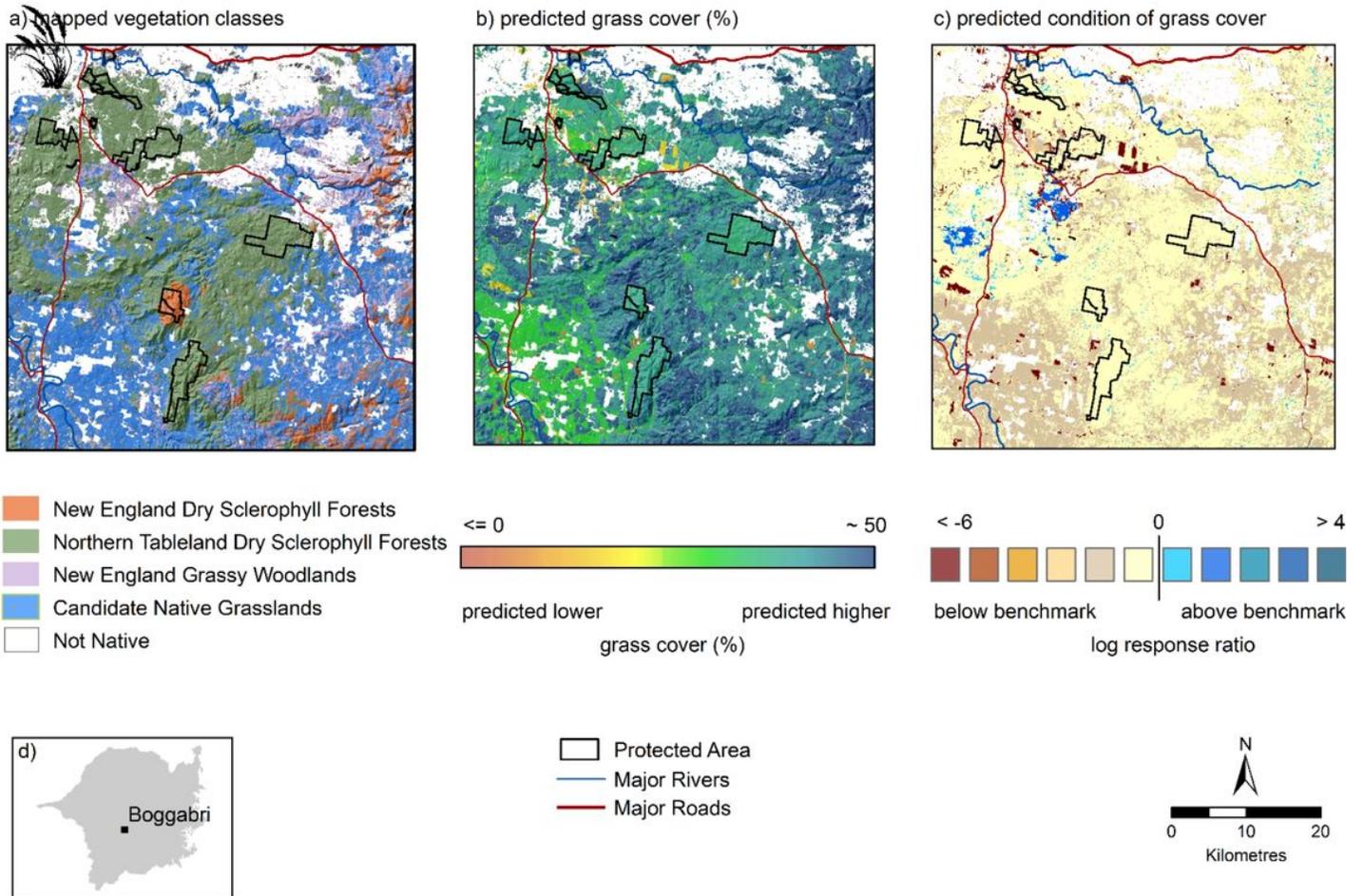


Figure 3

a) vegetation mapping showing extant and intact vegetation communities; b) a fine-scale and continuous model of predicted grass cover (%) (from McNellie et al. 2021); c) the integrity grass cover relative to benchmark; d) location diagram within the study area

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [McNellieUploadSupp20211129.docx](#)