

ECOGRAPHY

Research article

Patterns and drivers of plant diversity across Australia

Karel Mokany, James K. McCarthy, Daniel S. Falster, Rachael V. Gallagher, Thomas D. Harwood, Robert Kooyman and Mark Westoby

K. Mokany (<https://orcid.org/0000-0003-4199-3697>) ✉ (Karel.Mokany@csiro.au) and T. D. Harwood, CSIRO, Canberra, Australia. – J. K. McCarthy (<https://orcid.org/0000-0003-3060-1678>), Manaaki Whenua – Landcare Research, Lincoln, New Zealand. – D. S. Falster (<https://orcid.org/0000-0002-9814-092X>), Evolution and Ecology Research Centre, Univ. of New South Wales, Sydney, NSW, Australia. – R. V. Gallagher (<https://orcid.org/0000-0002-4680-8115>), Hawkesbury Inst. for the Environment, Western Sydney Univ., Penrith, NSW, Australia. – RVG, R. Kooyman and M. Westoby (<https://orcid.org/0000-0001-7690-4530>), School of Natural Sciences, Macquarie Univ., Sydney, NSW, Australia. RK also at: Royal Botanic Garden, Sydney, NSW, Australia.

Ecography

2022: e06426

doi: 10.1111/ecog.06426

Subject Editor: Holger Kreft

Editor-in-Chief: Miguel Araújo

Accepted 26 August 2022



Biodiversity analyses across continental extents are important in providing comprehensive information on patterns and likely drivers of diversity. For vascular plants in Australia, community-level diversity analyses have been restricted by the lack of a consistent plot-based survey dataset across the continent. To overcome these challenges, we collated and harmonised plot-based vegetation survey data from the major data sources across Australia and used them as the basis for modelling species richness (α -diversity) and community compositional dissimilarity (β -diversity), standardised to 400 m², with the aim of mapping diversity patterns and identifying potential environmental drivers. The harmonised Australian vegetation plot (HAVPlot) dataset includes 219 552 plots, of which we used 115 083 to analyse plant diversity. Models of species richness and compositional dissimilarity both explained approximately one-third of the variation in plant diversity across Australia ($D^2 = 33.0\%$ and 32.7% , respectively). The strongest environmental predictors for both aspects of diversity were a combination of temperature and precipitation, with soil texture and topographic heterogeneity also important. The fine-resolution (≈ 90 m) spatial predictions of species richness and compositional dissimilarity identify areas expected to be of particular importance for plant diversity, including south-western Australia, rainforests of eastern Australia and the Australian Alps. Arid areas of central and western Australia are predicted to support assemblages that are less speciose or unique; however, these areas are most in need of additional survey data to fill the spatial, environmental and taxonomic gaps in the HAVPlot dataset. The harmonised data and model predictions presented here provide new insight into plant diversity patterns across Australia, enabling a wide variety of future research, such as exploring changes in species abundances, linking compositional patterns to functional traits or undertaking conservation assessments for selected components of the flora.

Keywords: alpha-diversity, beta-diversity, biodiversity, community, composition, dissimilarity, plot, richness, species, survey, uniqueness, vegetation



www.ecography.org

© 2022 The Authors. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Introduction

Understanding spatial biodiversity patterns and the processes that shape them is fundamental to improving our knowledge of the natural world (Cox et al. 2020). This knowledge is also vital to inform conservation policy, planning and management decisions aimed at retaining the earth's biodiversity in the face of increasing threats (Moilanen and Wilson 2009). We need to understand how diversity is distributed in order to ensure our conservation measures adequately represent biodiversity (Ferrier and Drielsma 2010).

While good knowledge of diversity patterns is important, for many taxonomic groups our understanding is still incomplete across much of the world. One such case is for plant diversity across Australia; with more than 21 000 native vascular plant species, and 93% of these being endemic (Chapman 2009, CHAH 2022), the Australian continent harbours unique plant assemblages inhabiting a broad range of ecosystems (i.e. alpine, temperate, tropical, arid, Mediterranean). Plant diversity in Australia faces increasing and interacting threats from human actions, such as ongoing land clearing, extensive grazing, altered fire and water regimes, climate change and invasive species (Broadhurst and Coates 2017). An improved understanding of spatial diversity patterns in the Australian flora and their likely drivers is an essential foundation on which to build conservation actions aimed at both halting declines and restoring ecosystems.

To date, research on Australian plant diversity patterns and processes has been based primarily on presence-only occurrence observations, such as from herbarium collections (Crisp et al. 2001, Goldie et al. 2010, Stevenson et al. 2012, González-Orozco et al. 2014, Gallagher 2016, Guerin et al. 2016, Thornhill et al. 2016). In this regard, Australia has good coverage, with > 16M occurrence observations for plant species recorded (Atlas of Living Australia 2021). However, there are challenges in using such data to analyse and understand patterns in plant community diversity, primarily due to the lack of information on species absences or abundances, and the inherent bias in the species that are documented or the places and times that species are recorded (Isaac and Pocock 2015). For example, separating 'true' patterns in species richness from observation effort is very difficult when using presence-only data (Gotelli and Colwell 2011, Neyens et al. 2019).

Survey data of species composition are strongly preferable for analyses of plant community diversity, providing a systematic record of the species that were both present and absent in a survey plot at the time surveyed (Bruehlheide et al. 2019, Večeřa et al. 2021). In Australia, eight individual state and territory government agencies are primarily responsible for collecting and collating plant community survey data. Harnessing such data for continental diversity analyses has been a substantial challenge, given these jurisdictions implement different community survey methods, and apply their own systems to document, store and provide access to those data. To date, Australia has had no national, standardised plant community survey dataset that combines available data from the state and territory agencies (Gellie et al. 2018). Plot-based

analyses of plant diversity patterns in Australia have, therefore, been restricted to particular regions (Austin et al. 1996, Hunter 2005, Guerin et al. 2013, Mokany et al. 2014, McCarthy et al. 2018) or have utilised a relatively small number of survey plots (Rice and Westoby 1983, Andermann et al. 2022). The TERN AusPlots initiative has surveyed ca 870 plots across Australia to date using standard methods; however, these data still have substantial spatial, environmental and taxonomic gaps (Guerin et al. 2021).

To advance our understanding of plant community diversity patterns across the Australian continent, we harmonised survey data from major sources across Australia into a single, consistent dataset. We report on the key attributes of this harmonised Australian vegetation plot (HAVPlot) dataset, and use it to assess spatial patterns in two key dimensions of diversity: species richness (α -diversity) and pairwise community compositional dissimilarity (β -diversity). Our consideration of compositional dissimilarity is particularly important, given its strong role in determining the total number of species (γ -diversity) at regional and continental extents (Mokany et al. 2011). We model both aspects of plant diversity as a function of environmental attributes selected from a broad range of spatial environment layers of potential importance, with the aim of elucidating likely continent-wide drivers of plant species richness and compositional turnover. Finally, we use these diversity models to generate fine-resolution (≈ 90 m) maps of predicted plant diversity patterns, identifying areas that are priorities for future surveys and for conserving Australia's plant diversity.

Material and methods

Study area

The region considered for the present analyses was the Australian continent and nearby islands. We did not consider islands that are part of the Australian territories but distant from the continent, such as Lord Howe Island, given the lack of consistent fine-resolution spatial environmental data that encompass these areas (e.g. soil data: Viscarra Rossel et al. 2015).

Plant community survey data

Given the absence of an existing harmonised plant community survey plot dataset for Australia, we obtained and collated data primarily from state/territory and federal agencies, which are custodians of the largest survey datasets. Although the Terrestrial Ecosystem Research Network (TERN) AEKOS data portal (TERN 2022) provides open access to plant community survey data for Australia, it currently provides limited harmonisation or method alignment, and represents only a subset of data held by state/territory agencies. Some major state agency datasets are not available via AEKOS (e.g. Victoria, Western Australia, Tasmania, Northern Territory) and others are missing more recent surveys. We therefore

obtained plant community survey data from the relevant state/territory custodians wherever possible, with AEKOS used as the source where that was not possible. In some cases we downloaded data directly from publicly accessible websites, while for others we required assistance and permission to obtain relevant data (Supporting information).

Given the different formats of the source data, we developed a simple, customised and structured data format to harmonise across sources, based broadly on the Veg-X schema (Wiser et al. 2011), with existing standards for data fields used wherever possible (Veg-X, Darwin Core). Source data were harmonised to the common format using a customised script in R (<www.r-project.org>). Taxonomic nomenclature was standardised to the Australian Plant Census (CHAH 2022), using code adapted from Falster et al. (2021), with only vascular plant species retained. In the present analyses, records were considered at the species level, with information on finer taxonomic resolution (e.g. sub-species) ignored.

Survey data refinement and scaling

For our analyses, we refined the plot dataset to those data most suited to assessing and modelling plant community composition. From the full set of collated and harmonised plant community survey data ($n=219\,552$ plots) (Fig. 1A, Supporting information) we retained only those plots that sampled all vascular plant species present aboveground at the time of survey (full floristic surveys), where the plots were contiguous areas and where the plot area was available ($n=158\,032$ plots) (Supporting information). Given the importance of plot area in influencing community diversity metrics, we further excluded plots for our analyses that were very small ($< 20\text{ m}^2$; $n=3354$) or very large ($> 20\,000\text{ m}^2$; $n=873$). We also retained only those plots where $\geq 70\%$ of the plant species present were native to Australia, and where $\geq 90\%$ of taxa present were identified to species level. These refinements led to a dataset comprising 115 083 plots that form the basis for our analyses and modelling (Fig. 1B, Supporting information). Almost one-half of the plant community survey plots were from the New South Wales BioNet collection ($n=58\,466$), with the Victorian Biodiversity Atlas also providing a large number of survey plots ($n=14\,888$) given the small area of the state (Supporting information). The reported coordinate uncertainty for the selected plots was generally low (median = 30 m, 95th percentile = 100 m).

To derive a comparable species richness metric across plots of different area, we scaled the surveyed native plant species richness of each plot (Supporting information) to the median plot area of 400 m^2 (Supporting information), assuming a simple species–area power relationship ($S=cA^z$) (Rosenzweig 1995) with standard scalar $z=0.25$ applied across all survey plots. The species–area power model has been shown to perform well at relatively fine spatial resolutions (Dengler et al. 2020) and a scalar of approximately $z=0.25$ for vascular plants at fine spatial resolutions is generally supported by a number of studies (Rosenzweig 1995, Crawley and Harral 2001, Drakare et al. 2006, Dengler et al. 2020, Dembicz et al.

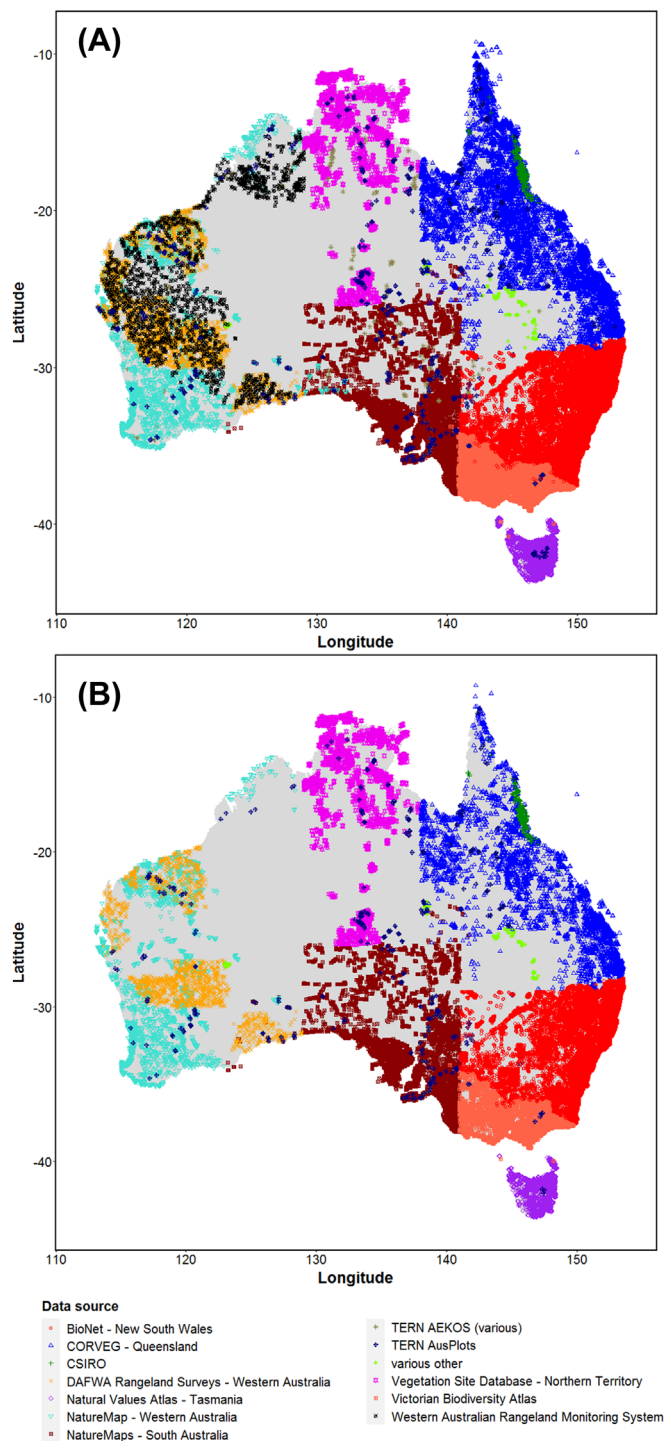


Figure 1. Locations and primary data sources for vegetation survey plots collated in the harmonised Australian vegetation plot (HAVPlot) dataset, including: (A) all 219 552 survey plots (Supporting information); (B) the 115 083 plant community survey plots used in the present analyses and modelling (Supporting information). The selected plots (B) are those from the full dataset that are full floristic surveys with known plot area, excluding very small ($< 20\text{ m}^2$) or large ($> 20\,000\text{ m}^2$) plots, excluding plots with $< 70\%$ of species recorded being native to Australia and excluding plots where $< 90\%$ of species were identified to species level.

2021). A sensitivity analysis indicated that the modelling results of our study are consistent when different species–area scalars are applied ($z=0.15$ – 0.30) (Supporting information).

Similarly we scaled the Sørensen's compositional dissimilarity between site-pairs to a plot area of 400 m², based on the area and species richness of the component plots, using the method described in Mokany et al. (2013). This approach uses the species–area power model to scale both the number of species in each plot in a pair as well as the number of species shared between the two plots, based on their area, then uses these scaled values to calculate their Sørensen's dissimilarity. For this we applied a scalar of $z_{\text{com}}=0.4$ in the power relationship for the number of species shared between a site-pair, as applied previously (Mokany et al. 2014). A sensitivity analysis indicated that the modelling results of our study are consistent when different scalars (z_{com}) are applied ($z_{\text{com}}=0.30$ – 0.45) (Supporting information).

Environmental predictors

Spatially complete and consistent environmental data for Australia were collated, derived and prepared, being aligned to the SRTM 3 second (≈ 90 m) digital elevation model (DEM) of Australia (Gallant et al. 2011). Topographically adjusted climate layers were generated using CSIRO's Terraforma package (Reside et al. 2013). This involved first generating monthly climate layers using ANUCLIM ver. 6.1 (Xu and Hutchinson 2010) based on the SRTM 3s DEM and underlying climate data for the period 1976–2005. The climate layers derived in ANUCLIM were then adjusted for the effects of topography (slope and direction) on local climate, based on fine-resolution solar radiation data, including mean monthly total shortwave (sloping surface), shortwave radiation ratio and net radiation (Gallant et al. 2014a, b, c). The climate variables generated using this approach were annual precipitation, precipitation seasonality (both solstice and equinox seasonality), annual evaporation (both potential and actual), annual water deficit (annual precipitation – potential evaporation), short-wave solar radiation (minimum monthly, maximum monthly), temperature (minimum monthly, maximum monthly, isothermality) and potential plant growth index.

Soil and landform layers at 3-s resolution were obtained from the TERN Soil and Landscape Grid of Australia (Gallant and Austin 2015, Grundy et al. 2015, Viscarra Rossel et al. 2015) including bulk density, available water-holding capacity, organic carbon content, nitrogen content, phosphorus content, pH and sand/silt/clay content, all aggregated to two depth ranges (0–30, 30–100 cm), plus an additional layer for total soil depth. These variables were supplemented with topographic descriptor variables derived from the 3-s DEM, including topographic wetness index (Gallant and Austin 2012b) and elevation focal range within a 300 m radius (Gallant and Austin 2012a), both of which may be useful in incorporating the effects of local topographic position on plant community composition. To account for the role that extended periods of inundation may have on plant

community composition, water observations from space (WOfS) data (Mueller et al. 2016) were used to derive a layer representing the average proportion of time each location was covered by freshwater (removing coastal and estuarine observations within 2 m elevation of the high tide mark). Environment data were extracted for the plant community plot locations, and correlations between environment variables were quantified to inform variable selection.

Species richness modelling

We generated a model of plant community species richness using generalised additive modelling (GAM) (Hastie and Tibshirani 1986), using the *mgcv* package in R (Worm et al. 2006, Wood 2016, <www.r-project.org>) with a Poisson link function and a maximum of 5 knots per predictor variable to avoid overfitting. Models were developed applying a cross-validation sample of 80% of randomly selected sites for model training, and the remaining 20% of sites were used for validation, with this random sampling procedure repeated 10 times. We derived a reduced set of candidate predictor variables from those described above by assessing the predictive power of each variable independently, then adding variables to the candidate set based on their individual explanatory power, ensuring no variables selected for further assessment were highly correlated (absolute Pearson's $R > 0.7$). From this initial candidate set of variables we applied a backward elimination variable selection approach. The performance of preliminary models to predict species richness for the testing sites was used to remove the least informative variable from the candidate set (assessed using deviance explained), stopping when a parsimonious set of statistically significant predictor variables remained.

Compositional dissimilarity modelling

We generated a model of pairwise community compositional dissimilarity for vascular plants using generalised dissimilarity modelling (GDM) (Ferrier et al. 2007), applying the *gdm* package in R (Fitzpatrick et al. 2021, <www.r-project.org>) using the default of three i-spline basis functions per predictor. Given the very large number of possible pairs of the survey plots (> 6 billion), we applied an integrated approach to site-pair sampling and variable selection, based on cross-validation. For each iteration of a cross-validation sample, 80% of sites were randomly selected to train the GDM, while the remaining 20% of sites were used to validate the predictions of the provisional models. For both sets of sites (training and testing) we applied the same methods to sample site pairs, with 300 000 site pairs generated for training the model (from the training plots) and 50 000 site pairs generated to test the model (from the testing plots). For each combination of model predictors assessed, this random sampling procedure was repeated 10 times.

Site-pair samples were generated for each of the test and training set of sites, based on a geographically weighted sampling scheme. This scheme is intentionally biased towards

selecting site pairs that are geographically closer together, hence likely to be more environmentally similar (across continental extents), while undertaking this sampling in an even manner across Australia. This approach ensures a good spatial and environmental coverage of sites in site pairs, from across the geographic and environmental space. For each sampling iteration, we applied a randomly situated net of sample nodes at 150 km distance, randomly sampling sites to combine into pairs using a Gaussian probability distribution around each node, and incorporated 10% of site pairs randomly selected over Australia (Mokany et al. 2018). We also included geographic distance as a potential predictor, calculated between each pair of locations using a projected coordinate system (Australian Albers, EPSG:3577).

As with the species richness model development, we derived a reduced set of candidate predictor variables based on individual explanatory power and predictor correlations. A final set of predictors was derived via backward elimination variable selection, stopping when a parsimonious set of statistically significant predictor variables remained in the final model, assessed via permutation test (Mokany et al. 2014, Fitzpatrick et al. 2021).

Spatial analyses

We spatially projected the models of plant community species richness and pair-wise compositional dissimilarity, using fine-resolution (3 s; ≈ 90 m) spatially complete layers of the predictor variables for the Australian land surface (Supporting information). For the compositional dissimilarity model, we visualised patterns in community dissimilarity by reducing the dimensionality of the model transformed predictor layers using principal components analysis (PCA), assigning a red, blue and green colour dimension to the first three PCA axes, then mapping this by combining the colours (Mokany et al. 2022c).

To identify areas of high potential importance for plant diversity, we combined the spatial projections of plant community species richness and pair-wise compositional dissimilarity to identify locations predicted to have more unique species assemblages containing a larger number of species. We quantified the compositional uniqueness of each location by calculating its average predicted dissimilarity to all locations in a random sample of 1% of the grid cells across Australia. We then normalised these compositional uniqueness values to a 0–1 range, normalised the predicted species richness values to a 0–1 range and calculated a ‘diversity importance’ score as the average of these two normalised values for each location across Australia.

To identify priority areas for collection and/or collation of additional survey plot data, we assessed the representativeness of the 115 083 survey plots used in our analyses (Fig. 1B), using the spatial projections of plant community compositional dissimilarity. Specifically, we calculated the mean predicted similarity of each location to the survey plot locations, divided by the mean predicted similarity of each location to a random sample of 1% of the grid cells across Australia

(Mokany et al. 2022c). Under this approach, locations with values > 1 are expected to be over-represented by the survey plot data while locations with values < 1 are expected to be under-represented by the survey plot data.

Results

Summary of plant community survey data

The plant community survey plots used for the present analyses and modelling ($n=115\,083$ plots) included 16 593 species (15 131 native species), with an average 28.9 ± 16.0 (SD) native species per plot (range = 1–248) (Fig. 1B, Supporting information). There were an average of 1.9 ± 3.1 (SD) non-native species per plot, which were excluded from the diversity analyses. The median frequency at which plant species were recorded in the survey plots was 21, with the most frequently recorded species being *Lomandra longifolia* ($n=22\,476$ survey plots), while 1987 species occurred in only a single survey plot. Of the pairs of survey plots selected for modelling community compositional dissimilarity ($n=3\,000\,000$), the mean Sørensen’s site-pair dissimilarity was 0.94 ± 0.08 (SD) (range = 0–1). The mean area of the plots analysed was $1097\text{ m}^2 \pm 2108$ (SD).

Species richness modelling

The final model of plant species richness contained nine predictor variables and explained 33.0% deviance in observed species richness (root mean square error = 13.9) (Table 1, Supporting information). The strongest predictors of plant species richness across Australia were temperature (maximum and minimum) and precipitation (Table 1, Fig. 2). The richness model also included predictors relevant to topography, soil texture, solar radiation and inundation frequency (Table 1, Fig. 2). There was no obvious bias in prediction error for scaled species richness in terms of the surveyed plot area (Supporting information). Spatial projection of the plant species richness model indicated the highest predicted species richness in the south-west, eastern ranges and the Kimberley region (north-west), while the lowest predicted richness was in arid central Australia (Fig. 3A).

Compositional dissimilarity modelling

The final model of compositional dissimilarity included nine predictor variables and explained 32.7% deviance in compositional dissimilarity (root mean square error = 0.07) (Table 1, Supporting information). The strongest predictors of plant community compositional dissimilarity across Australia were geographic distance, precipitation, potential plant growth index, maximum temperature and elevation focal range (Table 1, Fig. 4). Because of low to moderate correlation between predictor variables, the unique contribution of individual predictors to the multivariate model was relatively low (Table 1), with geographic distance having the greatest

Table 1. Variable contribution to the models of plant species richness and plant community compositional dissimilarity for Australia. For the species richness model, the importance of each variable in the multivariate model is indicated by the chi-squared statistic, while the importance of the predictor on its own is indicated by the deviance explained (D^2) in a single variable model. For the compositional dissimilarity model, the importance of each variable in the multivariate model is indicated by the loss of deviance explained when that variable is dropped, while the importance of the predictor on its own is indicated by the deviance explained in a single variable model. Empty cells indicate that the variable was not used in the model for that diversity metric.

Predictor variable	Species richness		Compositional dissimilarity	
	All variable model chi-squared	Single variable model D^2 (%)	All variable model loss of D^2 if dropped (%)	Single variable model D^2 (%)
Max. temperature warmest month ($^{\circ}\text{C}$)	53 629	9.3	0.6	16.5
Min. temperature coldest month ($^{\circ}\text{C}$)	15 491	1.1	0.5	7.6
Isothermality (%)	1391	1.7		
Mean annual precipitation (mm yr^{-1})	27 983	15.5	0.7	18.3
Max. monthly radiation (W m^{-2})	9028	6.6		
Potential plant growth index	3543	14.8	0.7	18.1
Elevation focal range, 300 m radius (m)	12 659	10.4	1.3	7.6
Soil silt content, 0–30 cm depth (%)	13 466	9.2		
Soil sand content, 0–30 cm depth (%)			0.5	5.2
Soil clay content, 0–30 cm depth (%)			0.3	5.7
Frequency of water coverage (%)	7538	2.0	0.4	0.9
Geographic distance (km)			2.0	19.2

All variables statistically significant at $p < 0.001$.

Multivariate species richness model deviance explained = 33.0%, $n = 115\,083$ plots, intercept = 3.25.

Multivariate compositional dissimilarity model deviance explained = 32.7%, $n = 300\,000$ site-pairs per cross-validation sample, intercept = 1.3.

irreplaceable explanatory power. The simplified spatial depiction of the compositional dissimilarity model (through the first three PCA axes of the transformed predictor variables) indicated areas of more similar plant communities within the south-east, within the monsoonal north, within the coastal south-west and within the eastern interior versus within the western interior (Fig. 3B), with notable differences between these regions. The first three PCA axes explained 55.3, 23.2 and 14.4% (92.9% total) of the variance in the transformed predictor variables.

Spatial analyses

Combining predicted spatial patterns of plant community species richness and compositional uniqueness to highlight areas of particular importance for plant diversity in Australia (Fig. 3C) shows similar patterns to the species richness projection (Fig. 3A), but with some nuance. For example, some areas have relatively high species richness but relatively lower diversity importance when incorporating compositional uniqueness, such as the north-west (Kimberley region) (Fig. 3C). In contrast, areas such as western Tasmania have moderate predicted species richness, but relatively higher diversity importance, due to high levels of predicted uniqueness (Fig. 3C). The survey plots utilised in the analyses were found to over-represent environments of south-eastern Australia, and under-represent central and western Australia (Fig. 3D). The Great Sandy Desert in north-western Australia was most poorly represented by the survey plots used in our analyses (Fig. 3D). This broadly reflected the coverage of the analysed plots in climate space (Supporting information), with good coverage in temperate environments but sparser coverage in the hottest and driest environments.

Discussion

Here we provide new insight into plant community diversity patterns across Australia. This analysis has identified important potential continent-wide environmental drivers of plant community species richness and compositional dissimilarity, with the response functions of plant diversity to these drivers quantified. Using these models, we have produced fine-resolution maps of predicted plant community diversity across Australia, which help to identify areas likely to be of particularly high importance for conservation planning and management across the continent. These advances in our understanding of Australian plant diversity are based on the harmonised plant community survey dataset we have compiled, providing a data-rich basis for our continental biodiversity analyses and broad potential to facilitate future research. The plot-based fine-resolution analyses that we present provide unique macroecological insights into plant diversity patterns at a continental extent (Beck et al. 2012, Biurrun et al. 2021, Večeřa et al. 2021).

Potential environmental drivers of plant diversity in Australia

Identifying the relationships between plant diversity and potential environmental drivers is complicated by the varying degrees of correlation between environmental variables and the way these variables are combined in multivariate models, such as those we have developed. Previous plot-based analyses within Australia and around the world have consistently identified moisture and temperature as key drivers of vascular plant species richness and compositional turnover

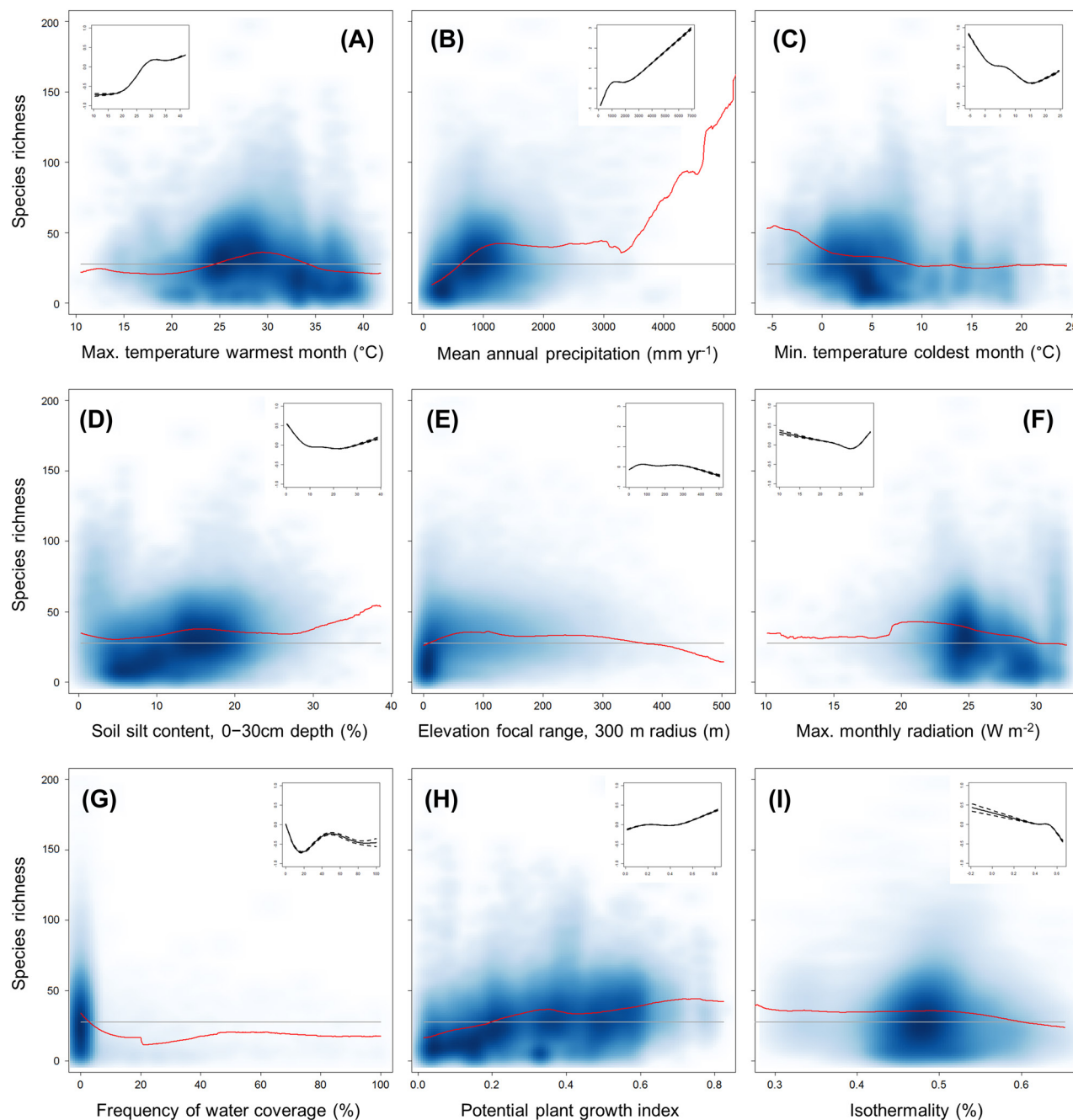


Figure 2. Plant species richness model response functions for the nine predictor variables selected. Species richness is scaled to the number of species per 400 m². Solid red lines indicate the predicted richness, when all other variables are held at their mean value across each section of the predictor gradient. Densities of observed values are shown in blue, with a horizontal line in grey marking the average species richness across all plots (27.5 species). The generalised additive modelling spline functions are shown in insets to each panel.

(Austin et al. 1996, Mokany et al. 2012, 2014, Keil and Chase 2019). Our analyses align with these previous studies, and with established theory on the evolution and distribution of the Australian flora (Hill 2017). Many previous studies of Australian plant diversity have been implemented at much coarser spatial resolutions (Crisp et al. 2001, González-Orozco et al. 2014, Thornhill et al. 2016), making direct comparison of identified key environmental drivers difficult,

given the influence of grain size on biodiversity patterns (Field et al. 2009, Keil and Chase 2019).

Annual precipitation was one of the most important predictors in our models (Table 1), with the response functions indicating that increases in precipitation at the low end of the gradient (100–1000 mm yr⁻¹) strongly promote increased species richness and turnover in community composition (Fig. 2, 4). Given the predominance of aridity across

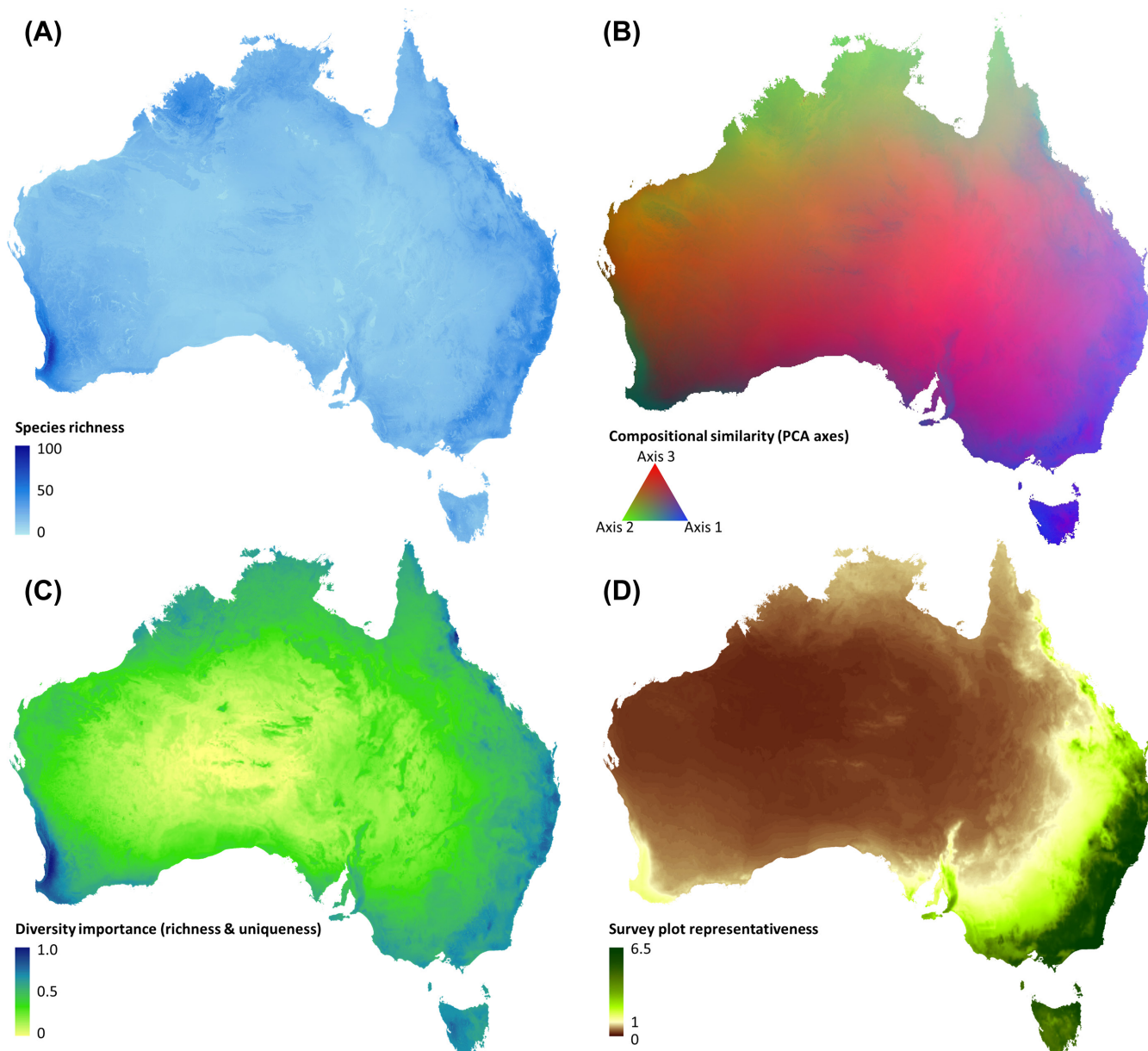


Figure 3. Spatial projections and analyses using the models of species richness and compositional dissimilarity for vascular plants across Australia. (A) Predicted species richness for a 400 m² area (median survey plot area). (B) Primary differences in predicted community composition based on a PCA of the generalised dissimilarity model (GDM) transformed environmental predictors, where more similar colours indicate more similar expected community composition. (C) An example of predicted diversity importance for each location, being a weighted average of the range normalised (0–1) predicted species richness and the range normalised (0–1) predicted compositional uniqueness (Supporting information). (D) Estimated representativeness of the 115 083 plots used in our analyses (Fig. 1B, Supporting information), being the mean predicted similarity of each location in Australia to the survey plot locations, relative to mean predicted similarity of each location to all Australia (values > 1 (green) are over represented while values < 1 (brown) are under-represented by the survey plots).

Australia, with 67% of the land area having less than 500 mm yr⁻¹ mean annual precipitation, many plant species in low rainfall environments have likely adapted to relatively narrow bands of precipitation (Gallagher et al. 2019), contributing to these community-level diversity patterns. The response of plant diversity to temperature was also strong, though more complex. Maximum temperature was a stronger predictor of richness and turnover than minimum

temperature (Table 1), with richness being greater in plots with intermediate maximum temperature (~ 30°C) and low minimum temperature (< 5°C) (Fig. 1). The role of temperature in influencing plant community richness and turnover is highly dependent on the span of the temperature gradient considered and the interaction with other variables, such as precipitation (Pausas and Austin 2001, Sommer et al. 2010). While increasing temperature can promote plant growth

(Angilletta et al. 2003), in Australia the locations with the highest maximum temperatures also receive some of the lowest amounts of precipitation (Supporting information). In contrast, our data indicate that locations with low minimum temperatures have high plant species richness (Fig. 2), likely due to the reduction in plant size in colder environments (e.g. alpine) enabling a greater number of species to co-occur in a given area (Storch et al. 2018), though there is evidence contrary to this hypothesis for the Australian Alps (Mallen-Cooper and Pickering 2008).

Here we consider the interaction between energy (temperature, solar radiation) and moisture availability through the potential plant growth index (Xu and Hutchinson 2010), which was a strong predictor of both richness and compositional dissimilarity (Table 1). We observed a relatively constant increase in species richness with increasing plant growth index (Fig. 2), and a relatively constant rate of compositional turnover along this gradient (Fig. 4). This finding fits with widespread evidence from Australia and around the world for strong, generally positive relationships between plot-scale plant community diversity and plant growth (Specht and Specht 1993, Keil and Chase 2019).

Environmental heterogeneity is another key driver of plant species richness identified by previous plot-based studies (Costanza et al. 2011, Stein et al. 2014), and reflected in our analysis through elevation focal range (the range of elevations within a 300 m radius). Richness increased with elevation focal range values from 0 to 75 m, which covers 95.3% of locations in Australia, with evidence of lower species richness in the few areas where local topographic variability becomes very high (Fig. 2). Similarly, differences in elevation focal range explained differences in composition mostly along the lowest values (0–25 m) (Fig. 4), indicating that flat or near-flat locations were consistently different in their species composition from those with more topographic variability.

Substrate is also commonly observed to influence vegetation, though its role in influencing plant diversity is better understood at local–regional extents rather than continental extents (Pärtel 2002, Hulshof and Spasojevic 2020). Our analyses identified soil texture (soil clay, silt and sand content) having moderate importance in explaining changes in species richness and compositional dissimilarity across Australia (Table 1, Fig. 2, 4). Improved methods for incorporating region-specific influence of soil attributes in models of plant diversity across a continental extent would be valuable.

Geographic distance was a key predictor of plant compositional dissimilarity across Australia (Fig. 4A), as would be expected for a continental analysis. Much of the variation in compositional dissimilarity explained by geographic distance may be due to spatial autocorrelation in the environmental predictors (Table 1) (Mokany et al. 2022c), or by it acting as a surrogate for important environmental variables not included in our model. However, the spatial separation of locations undoubtedly has an important role in influencing compositional turnover across Australia. This is not only through biogeographic processes that cause divergent evolutionary

trajectories in different regions of Australia with similar environments, such as south-east and south-west Australia (Hopper and Gioia 2004), but also through stochastic community assembly processes playing out across large spatial extents (Hubbell 2001). More advanced consideration of geographic distance could further improve our predictions of compositional dissimilarity, such as using least-cost paths that consider environment barriers (Adriaensen et al. 2003).

Our modelling suggests that both plant species richness and compositional dissimilarity across Australia are related to a common set of environmental variables, with six out of the nine best predictor variables shared between the models of these two aspects of diversity (Table 1). This outcome is logical, given that the same inherent process influences both metrics – changes in the occurrences of species across environmental and spatial gradients (Whittaker 1960, Veech et al. 2002). In addition, the Sørensen dissimilarity metric we use incorporates turnover due to nestedness as well as replacement (Baselga 2010), so it is not independent of species richness.

It is also important to highlight that, while our analyses focus on continental drivers of plant diversity across Australia, we expect the environmental variables we identified will vary in importance across different regions of Australia. Owing to biogeographic idiosyncrasies and evolutionary history, we expect that other environmental variables will be strongly related to plant diversity when considered across smaller spatial extents (Pearson and Dawson 2003). In addition, the models we fitted have relatively modest explanatory power, explaining about one-third of the variation in species richness and compositional dissimilarity (Table 1). While we could have increased the explanatory power of our models by providing the splines for each predictor with more flexibility, we intentionally maintained simpler response functions in order to identify the main trends in diversity along each environmental gradient and to avoid overfitting. The relatively large amount of variation in plant diversity unaccounted for is likely to be at least partially due to different regional drivers across Australia, as well as to inherent stochasticity in species occurrences. The use of spatial environmental layers as predictors also likely reduced the explanatory power of our diversity models, given that each predictor layer has its own inherent error. The data we have prepared and analysed also enable more mechanistic exploration of environmental drivers of plant diversity in Australia, though this is limited at present by the lack of process-based diversity models applicable at the continental extent.

Spatial patterns of plant diversity across Australia

Projecting the models of plant species richness and compositional dissimilarity spatially across Australia generates patterns that broadly align with expectations and previous analyses (Andrew et al. 2021). Plant species richness is higher in the mesic temperate and tropical coastal ranges, while it is lowest in the arid interior (Fig. 3A). Particularly high richness in coastal south-western Australia aligns with existing

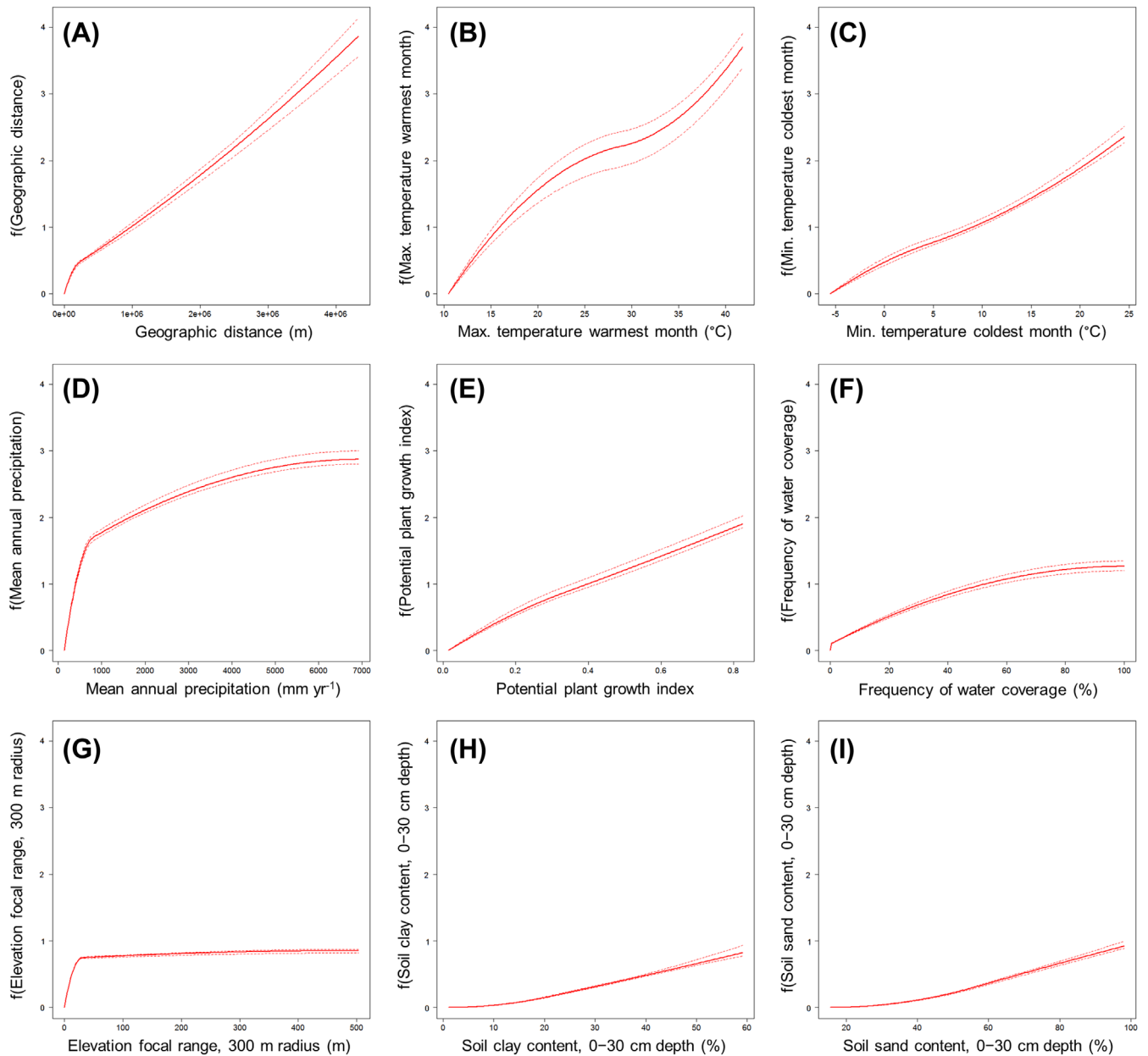


Figure 4. Fitted spline functions for each predictor variable included in the generalised dissimilarity model (GDM). For each panel, the variable's native scale is on the x-axis, and the GDM transformed values on the y-axis. The maximum height of the spline function (hence the maximum value of the transformed predictor: $f(x)$) indicates the overall importance of the predictor in explaining dissimilarities. Sections of the spline function with steeper slope indicate greater dissimilarity per unit change in the predictor variable along that section of the gradient. Dashed lines indicate extremes of predicted spline functions across 10 cross-validation sets.

evidence and the importance of the region as a known floristic diversity hotspot (Hopper and Gioia 2004). High species richness in the Kimberley region in north-western Australia is unexpected (Fig. 3A), and may at least partially be an artefact of the sparse data available for this region (Fig. 1), with most of the data that are available being from targeted surveys of small patches of rainforest (Kenneally et al. 1991) which sit within a much broader landscape dominated by tropical savanna vegetation.

Predictions from the model of compositional dissimilarity indicate many of the patterns expected for Australian plant

community turnover (Fig. 3B). These include broad alignment with the biomes and floristic regions recognised across Australia (González-Orozco et al. 2014, Ebach and Murphy 2020), including the eastern ranges, wet tropics, monsoonal tropics, south-west and the arid centre. We also visualised the model of compositional dissimilarity by calculating the expected compositional uniqueness of each location, being the average dissimilarity of each location to all other locations in Australia. This analysis again aligned with expectations, with more common environments such as the arid centre having lower predicted uniqueness than much more restricted and

rare environments, such as areas of rainforest, alpine areas and south-west Australia (Supporting information).

Here we have combined the spatial predictions of plant species richness and compositional uniqueness to map an index of 'diversity importance' across Australia. The locations predicted to support the most unique and speciose assemblages are in south-west Australia, areas of rainforest in coastal eastern Australia and the Australian Alps in the south-east (Fig. 3C). These areas identified in our analyses as being of high importance for plant diversity in Australia are likely to have more rare species of conservation interest than other locations, and they align with a large body of past research on the diversity of the Australian flora, especially the importance of relictual rainforest habitats (Webb et al. 1984, Kooyman et al. 2012).

The spatial predictions of plant diversity patterns across Australia we have derived and made available (Mokany et al. 2022a) could have a variety of conservation applications. The fine-resolution (≈ 90 m) maps of species richness (Fig. 3A) compositional uniqueness (Supporting information) and diversity importance (Fig. 3C) could be combined with spatial information on habitat condition (Harwood et al. 2016) and used to inform policy, planning and management of development or restoration activities (Ferrier and Drielsma 2010). These would be particularly valuable where consistent information is required at a national level or across state borders. The spatial prediction layers from the compositional dissimilarity model could also form the basis for a wide variety of further conservation assessments (Mokany et al. 2022c), such as assessing the representativeness of existing and proposed protected areas (Williams et al. 2016).

The harmonised plant community survey data

We spent considerable effort in collating and harmonising plant community survey data from across Australia to form the basis for our analyses. A key aspect of this data harmonisation process involved confirming important methodological aspects of each 'project' (set of linked plot surveys), particularly the dimensions of the area surveyed and the taxonomic scope of the survey (i.e. whether all or a subset of vascular plants were recorded). These important data are often not available from the source data files, though they are crucial when synthesizing data from multiple projects. Another important aspect of our data harmonisation process involved standardising and correcting species names to the Australian plant census (CHAH 2022), accounting for typos and use of different names for a species by different state agencies, and also changes in species names with ongoing taxonomic revision.

The harmonised continental plot dataset we have generated for Australia ('HAVPlot') should be a valuable resource for a wide range of future analyses, transforming quantitative understanding of the Australian flora (Mokany et al. 2022b). Potential applications include improved modelling of species distributions where presence-absence data are preferred over presence-only data, and quantifying gradients in vegetation functional attributes by combining the plot data

with newly available trait datasets (Falster et al. 2021). The HAVPlot dataset could also be combined with other continental plot datasets (Chytrý et al. 2016, Maitner et al. 2018) and incorporated into global plant community survey datasets (Bruehlheide et al. 2019), providing a more standardised and comprehensive representation of Australian plots in macroecological analyses. While some of the component datasets have license conditions that restrict us from sharing them, we have provided access to as much of the HAVPlot dataset that we are legally permitted to (205 084 plots) (Mokany et al. 2022b). Finally, we suggest that it would be beneficial for Australian state government agencies to move to more consistent approaches to providing open access to up-to-date ecological survey data.

The HAVPlot dataset we have produced implements a relatively simple relational data structure, harnessing existing standards wherever possible for naming and organising the data (Wiser et al. 2011). We note that the analyses presented here utilise only a portion of the total HAVPlot dataset (Fig. 1) and that alternative applications may use more or fewer of the complete set of plots available. It is also important to emphasise that the current HAVPlot dataset has substantial gaps in spatial, environmental and taxonomic coverage. Far fewer data are available in central, western and northern Australia compared to the more intensively surveyed south-east (Fig. 1, 3B, Supporting information). These areas and environments that are less well represented in the plot data, particularly the western deserts, are obvious priorities for future plant community survey projects to ensure a more balanced understanding of diversity patterns. Also noteworthy is the taxonomic representation of HAVPlot, with observations of 15 826 native species in the full dataset and 15 131 native species in the plots analysed here (Supporting information). While that is a large proportion of the 21 708 Australian native plant species currently documented (CHAH 2022), there remain thousands of species not represented in the HAVPlot dataset. Given that the species not surveyed are likely to be rare, it may take many more survey plots to achieve near-full taxonomic coverage.

There is also substantial scope to expand the current HAVPlot dataset by incorporating the many smaller plot datasets held by other agencies and universities. This could include the many plant community surveys undertaken as a part of environmental impact assessments (Samuel 2020), which are currently only represented in some of the state government datasets we have collated. Harmonising smaller datasets into HAVPlot could prioritise areas where existing coverage is poor, and could be implemented more efficiently through a collaborative open data publication approach. Refining existing information infrastructure (TERN 2022) to enable and promote user-contributed and harmonised plot datasets would also be beneficial.

Limitations and future priorities

Our study has harmonised community survey plot data to provide new insight into plant diversity drivers and patterns

across Australia. However, there are a number of important limitations with our study that could be addressed through future research. Most basically, while we have expended considerable time producing a harmonised plot dataset, it is likely that this contains some errors both from the source data as well as the harmonisation process. Further interrogation and use of the HAVPlot data for different purposes will help in identifying and correcting such errors. We also examine just two metrics of diversity; future studies that explore other metrics (e.g. evenness) would be valuable, and could address the harmonisation of the substantial amount of abundance data available in the HAVPlot dataset (Supporting information).

A key issue in using these data for diversity analyses is dealing with the different plot areas sampled, given the importance of sample area in influencing richness and compositional dissimilarity measures (Rosenzweig 1995, Mokany et al. 2013). While one approach is to incorporate plot area as a covariate in a multivariate model (Kreft and Jetz 2007), a method for implementing this within GDM has yet to be developed. To retain a standard approach to dealing with plot area for both species richness and compositional dissimilarity, we applied a relatively simple pre-modelling standardisation based on a species–area power model (Mokany et al. 2013) using a simple assumption of constant power-model scalars across all Australia. Future application of improved approaches to accounting for differences in plot area for analyses of Australian plant diversity would be beneficial, including better understanding and accounting for the likely environmental dependence in how diversity measures scale with area (Rosenzweig 1995).

Also of high value would be dedicated modelling and analysis into how environmental drivers of plant diversity vary between different regions within Australia. This could provide new insight into the biogeographic processes responsible for plant diversity patterns in Australia, as well as improved spatial predictions of diversity patterns, potentially harnessing geographically weighted modelling techniques (Fotheringham et al. 2002, Alves et al. 2018). Machine learning modelling approaches could also be worth exploring for improving the accuracy of diversity models for Australian plants, though these approaches may reduce interpretability and potentially increase extrapolation error in environments where plot coverage is poor (Elith et al. 2008, Andermann et al. 2022). It is also important to note that our exploration of potential environmental drivers of plant community diversity patterns harnessed spatial environmental layers that have their own bias, errors and uncertainty associated with them (Xu and Hutchinson 2010, Storlie et al. 2013). Hence, the finer spatial resolution of our models and maps compared to many other studies do not imply higher predictive accuracy.

Conclusion

Here we have provided new insight into the patterns and likely drivers of plant diversity across Australia. While many

of our findings align with existing hypotheses and evidence, our assessment of both α - and β -diversity provides novel predictions that may stimulate new hypotheses and research, and provides improved information for biodiversity policy, planning and management. Importantly, our analyses have a firm quantitative basis in the large harmonised plot dataset we have collated and produced. This provides a new resource for future plant diversity research to improve, extend and utilise.

Acknowledgements – The data sources utilised are cited in full in Supporting information, and include: data supplied by Department of Environment and Natural Resources © Northern Territory of Australia; Natural Values Atlas (<www.naturalvaluesatlas.tas.gov.au>), 2022, © State of Tasmania; NSW BioNet Flora Survey Data Collection © State Government of NSW and Department of Planning, Industry and Environment 2013; Queensland CORVEG Database, ver. 8/3/2019 State of Queensland (Department of Environment and Science, <www.des.qld.gov.au/>); Victorian Biodiversity Atlas © State Government of Victoria, (accessed June 2017); NatureMap © State Government of Western Australia; NatureMaps © State Government of South Australia, Department for Environment and Water; TERN Ausplots, The Univ. of Adelaide (<www.adelaide.edu.au>), Adelaide, South Australia – supported by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS); Desert Ecology Research Group Plots © 2015–2018 Rights owned by the Univ. of Sydney; AusCover © 2011–2013 The Univ. of Queensland (Joint Remote Sensing Research Program).

Funding – This research was supported by a Julius Career Award to KM by CSIRO.

Author contributions

Karel Mokany: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Funding acquisition (lead); Investigation (lead); Methodology (lead); Project administration (lead); Resources (lead); Software (lead); Supervision (lead); Validation (lead); Visualization (lead); Writing – original draft (lead); Writing – review and editing (lead). **James K. McCarthy:** Conceptualization (supporting); Data curation (supporting); Formal analysis (supporting); Investigation (supporting); Methodology (supporting); Project administration (supporting); Software (supporting); Visualization (supporting); Writing – review and editing (supporting). **Daniel S. Falster:** Conceptualization (supporting); Data curation (supporting); Formal analysis (supporting); Funding acquisition (equal); Investigation (supporting); Methodology (supporting); Project administration (supporting); Software (supporting); Validation (supporting); Writing – review and editing (supporting). **Rachael V. Gallagher:** Conceptualization (supporting); Investigation (supporting); Methodology (supporting); Writing – review and editing (supporting). **Thomas D. Harwood:** Data curation (supporting); Methodology (supporting); Software (supporting); Writing – review and editing (supporting). **Robert Kooyman:** Conceptualization (supporting); Data curation (supporting); Investigation (supporting); Methodology (supporting); Writing – review and editing (supporting). **Mark Westoby:** Conceptualization

(supporting); Investigation (supporting); Methodology (supporting); Project administration (supporting); Writing – review and editing (supporting).

Transparent peer review

The peer review history for this article is available at <<https://publons.com/publon/10.1111/ecog.06426>>.

Data availability statement

Data are available from the CSIRO Data Access Portal: <<https://doi.org/10.25919/mk24-1792>> (Mokany et al. 2022a) and <<https://doi.org/10.25919/5cex-4s70>> (Mokany et al. 2022b).

Supporting information

The Supporting information associated with this article is available with the online version.

References

- Adriaensen, F. et al. 2003. The application of 'least-cost' modelling as a functional landscape model. – *Landsc. Urban Plan.* 64: 233–247.
- Alves, D. M. C. C. et al. 2018. Geographic variation in the relationship between large-scale environmental determinants and bat species richness. – *Basic Appl. Ecol.* 27: 1–8.
- Andermann, T. et al. 2022. Estimating alpha, beta and gamma diversity through deep learning. – *Front. Plant Sci.* 13: 839407.
- Andrew, S. C. et al. 2021. Functional diversity of the Australian flora: strong links to species richness and climate. – *J. Veg. Sci.* 32: e13018.
- Angilletta, M. J. et al. 2003. Tradeoffs and the evolution of thermal reaction norms. – *Trends Ecol. Evol.* 18: 234–240.
- Atlas of Living Australia 2021. Atlas of living Australia. – <www.ala.org.au/>.
- Austin, M. P. et al. 1996. Patterns of tree species richness in relation to environment in southeastern New South Wales, Australia. – *Aust. J. Ecol.* 21: 154–164.
- Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. – *Global Ecol. Biogeogr.* 19: 134–143.
- Beck, J. et al. 2012. What's on the horizon for macroecology? – *Ecography* 35: 673–683.
- Biurrun, I. et al. 2021. Benchmarking plant diversity of Palaearctic grasslands and other open habitats. – *J. Veg. Sci.* 32: e13050.
- Broadhurst, L. and Coates, D. 2017. Plant conservation in Australia: current directions and future challenges. – *Plant Divers.* 39: 348–356.
- Bruehlheide, H. et al. 2019. sPlot – a new tool for global vegetation analyses. – *J. Veg. Sci.* 30: 161–186.
- CHAH 2022. Australian Plant Census, Centre of Australian National Biodiversity Research. – Council of Heads of Australasian Herbaria (CHAH), <<https://id.biodiversity.org.au/tree/51354547>>.
- Chapman, A. 2009. Numbers of living species in Australia and the world. – Australian Government.
- Chytrý, M. et al. 2016. European vegetation archive (EVA): an integrated database of European vegetation plots. – *Appl. Veg. Sci.* 19: 173–180.
- Costanza, J. K. et al. 2011. Multi-scale environmental heterogeneity as a predictor of plant species richness. – *Landsc. Ecol.* 26: 851–864.
- Cox, C. B. et al. 2020. Biogeography: an ecological and evolutionary approach. – Wiley.
- Crawley, M. J. and Harral, J. E. 2001. Scale dependence in plant biodiversity. – *Science* 291: 864–868.
- Crisp, M. D. et al. 2001. Endemism in the Australian flora. – *J. Biogeogr.* 28: 183–198.
- Dembicz, I. et al. 2021. Fine-grain beta diversity of Palaearctic grassland vegetation. – *J. Veg. Sci.* 32: e13045.
- Dengler, J. et al. 2020. Species–area relationships in continuous vegetation: evidence from Palaearctic grasslands. – *J. Biogeogr.* 47: 72–86.
- Drakare, S. et al. 2006. The imprint of the geographical, evolutionary and ecological context on species–area relationships. – Blackwell Science, pp. 215–227.
- Ebach, M. C. and Murphy, D. J. 2020. Carving up Australia's arid zone: a review of the bioregionalisation of the Eremaean and Eyrean biogeographic regions. – *Aust. J. Bot.* 68: 229–244.
- Elith, J. et al. 2008. A working guide to boosted regression trees. – *J. Anim. Ecol.* 77: 802–813.
- Falster, D. et al. 2021. AusTraits, a curated plant trait database for the Australian flora. – *Sci. Data* 8: 254.
- Ferrier, S. and Drielsma, M. 2010. Synthesis of pattern and process in biodiversity conservation assessment: a flexible whole-landscape modelling framework. – *Divers. Distrib.* 16: 386–402.
- Ferrier, S. et al. 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. – *Divers. Distrib.* 13: 252–264.
- Field, R. et al. 2009. Spatial species-richness gradients across scales: a meta-analysis. – *J. Biogeogr.* 36: 132–147.
- Fitzpatrick, M. C. et al. 2021. gdm: generalized dissimilarity modeling. – R package ver. 1.5.0-1, <<https://CRAN.R-project.org/package=gdm>>.
- Fotheringham, A. S. et al. 2002. Geographically weighted regression: the analysis of spatially varying relationships. – Wiley.
- Gallagher, R. V. 2016. Correlates of range size variation in the Australian seed-plant flora. – *J. Biogeogr.* 43: 1287–1298.
- Gallagher, R. V. et al. 2019. Safety margins and adaptive capacity of vegetation to climate change. – *Sci. Rep.* 9: 8241.
- Gallant, J. and Austin, J. M. 2012a. Relief – elevation range over 300 m derived from 1" SRTM DEM-S. ver. 2. – CSIRO Data Collection.
- Gallant, J. and Austin, J. M. 2012b. Topographic wetness index derived from 1" SRTM DEM-H. ver. 2. – CSIRO Data Collection.
- Gallant, J. C. and Austin, J. M. 2015. Derivation of terrain covariates for digital soil mapping in Australia. – *Soil Res.* 53: 895–906.
- Gallant, J. et al. 2011. SRTM-derived 3 second digital elevation models ver. 1.0. – Geoscience Australia, Commonwealth of Australia.
- Gallant, J. et al. 2014a. Mean monthly net radiation modelled using the 1" DEM-S – 3" mosaic. ver. 2. – CSIRO Data Collection.
- Gallant, J. et al. 2014b. Mean monthly total shortwave radiation on a sloping surface modelled using the 1" DEM-S – 3" mosaic. ver. 2. – CSIRO Data Collection.
- Gallant, J. et al. 2014c. Mean monthly shortwave radiation ratio modelled using the 1" DEM-S – 3" mosaic. ver. 2. – CSIRO Data Collection.

- Gellie, N. J. H. et al. 2018. Overview of plot-based vegetation classification approaches within Australia. – *Phytocoenologia* 48: 251–272.
- Goldie, X. et al. 2010. Evolutionary speed limited by water in arid Australia. – *Proc. R. Soc. B* 277: 2645–2653.
- González-Orozco, C. E. et al. 2014. Quantifying phytogeographical regions of Australia using geospatial turnover in species composition. – *PLoS One* 9: e92558.
- Gotelli, N. J. and Colwell, R. K. 2011. Estimating species richness. – In: Magurran, A. E. and McGill, B. J. (eds), *Biological diversity: frontiers in measurement and assessment*. Oxford Univ. Press, pp. 39–54.
- Grundey, M. et al. 2015. Soil and landscape grid of Australia. – *Soil Res.* 53: 835–844.
- Guerin, G. R. et al. 2013. Spatial modelling of species turnover identifies climate ecotones, climate change tipping points and vulnerable taxonomic groups. – *Ecography* 36: 1086–1096.
- Guerin, G. R. et al. 2016. Identifying centres of plant biodiversity in South Australia. – *PLoS One* 11: e0144779.
- Guerin, G. R. et al. 2021. Using generalised dissimilarity modelling and targeted field surveys to gap-fill an ecosystem surveillance network. – *J. Appl. Ecol.* 58: 766–776.
- Harwood, T. D. et al. 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 Ecol. Evol.* 7: 1050–1059.
- Hastie, T. and Tibshirani, R. 1986. Generalized additive models. – *Stat. Sci.* 1: 297–310.
- Hill, R. S. 2017. *History of the Australian vegetation: cretaceous to recent*. – Cambridge Univ. Press.
- Hopper, S. D. and Gioia, P. 2004. The Southwest Australian Floristic Region: evolution and conservation of a global hot spot of biodiversity. – *Annu. Rev. Ecol. Syst.* 35: 623–650.
- Hubbell, S. 2001. *The unified neutral theory of biodiversity and biogeography*. – Princeton Univ. Press.
- Hulshof, C. M. and Spasojevic, M. J. 2020. The edaphic control of plant diversity. – *Global Ecol. Biogeogr.* 29: 1634–1650.
- Hunter, J. T. 2005. Geographic variation in plant species richness patterns within temperate eucalypt woodlands of eastern Australia. – *Ecography* 28: 505–514.
- Isaac, N. J. B. and Pocock, M. J. O. 2015. Bias and information in biological records. – *Biol. J. Linn. Soc.* 115: 522–531.
- Keil, P. and Chase, J. M. 2019. Global patterns and drivers of tree diversity integrated across a continuum of spatial grains. – *Nat. Ecol. Evol.* 3: 390–399.
- Kenneally, K. F. et al. 1991. Floristics and phytogeography of Kimberley rainforests, Western Australia. – In: McKenzie, N. L. et al. (eds), *Kimberley rainforests of Australia*. Surrey Beatty, pp. 93–131.
- Kooyman, R. et al. 2012. Australian tropical and subtropical rain forest community assembly: phylogeny, functional biogeography and environmental gradients. – *Biotropica* 44: 668–679.
- Kreft, H. and Jetz, W. 2007. Global patterns and determinants of vascular plant diversity. – *Proc. Natl Acad. Sci. USA* 104: 5925–5930.
- Maitner, B. S. et al. 2018. The BIEN R package: a tool to access the Botanical Information and Ecology Network (BIEN) database. – *Methods Ecol. Evol.* 9: 373–379.
- Mallen-Cooper, J. and Pickering, C. M. 2008. Linear declines in exotic and native plant species richness along an increasing altitudinal gradient in the Snowy Mountains, Australia. – *Austral Ecol.* 33: 684–690.
- McCarthy, J. K. et al. 2018. Predicting community rank-abundance distributions under current and future climates. – *Ecography* 41: 1572–1582.
- Moilanen, A. and Wilson, K. A. 2009. *Spatial conservation prioritization: quantitative methods and computational tools*. – Oxford Univ. Press.
- Mokany, K. et al. 2011. Combining A- and B-diversity models to fill gaps in our knowledge of biodiversity. – *Ecol. Lett.* 14: 1043–1051.
- Mokany, K. et al. 2012. Dynamic macroecology and the future for biodiversity. – *Global Change Biol.* 18: 3149–3159.
- Mokany, K. et al. 2013. Scaling pairwise β -diversity and α -diversity with area. – *J. Biogeogr.* 40: 2299–2309.
- Mokany, K. et al. 2014. Identifying priority areas for conservation and management in diverse tropical forests. – *PLoS One* 9: e89084.
- Mokany, K. et al. 2018. Enhancing landscape data: capacity building for GDM analyses to support biodiversity assessment. – CSIRO.
- Mokany, K. et al. 2022a. Plant diversity spatial layers for Australia, ver. 1. – CSIRO Data Collection, <<https://doi.org/10.25919/mk24-1792>>.
- Mokany, K. et al. 2022b. Harmonised Australian vegetation plot dataset (HAVPlot), ver. 1. – CSIRO Data Collection, <<https://doi.org/10.25919/5cex-4s70>>.
- Mokany, K. et al. 2022c. A working guide to harnessing generalized dissimilarity modelling for biodiversity analysis and conservation assessment. – *Global Ecol. Biogeogr.* 31: 802–821.
- Mueller, N. et al. 2016. Water observations from space: mapping surface water from 25 years of Landsat imagery across Australia. – *Remote Sens. Environ.* 174: 341–352.
- Neyens, T. et al. 2019. Mapping species richness using opportunistic samples: a case study on ground-floor bryophyte species richness in the Belgian province of Limburg. – *Sci. Rep.* 9: 19122.
- Pärtel, M. 2002. Local plant diversity patterns and evolutionary history at the regional scale. – *Ecology* 83: 2361–2366.
- Pausas, J. G. and Austin, M. P. 2001. Patterns of plant species richness in relation to different environments: an appraisal. – *J. Veg. Sci.* 12: 153–166.
- Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – *Global Ecol. Biogeogr.* 12: 361–371.
- Reside, A. E. et al. 2013. Climate change refugia for terrestrial biodiversity: defining areas that promote species persistence and ecosystem resilience in the face of global climate change. – National Climate Change Adaptation Research Facility, Gold Coast.
- Rice, B. and Westoby, M. 1983. Plant species richness at the 0.1 hectare scale in Australian vegetation compared to other continents. – *Vegetatio* 52: 129–140.
- Rosenzweig, M. L. 1995. *Species diversity in space and time*. – Cambridge Univ. Press.
- Samuel, G. 2020. Independent review of the EPBC act – final report. – Dept of Agriculture Water and the Environment.
- Sommer, J. H. et al. 2010. Projected impacts of climate change on regional capacities for global plant species richness. – *Proc. R. Soc. B* 277: 2271–2280.
- Specht, A. and Specht, R. L. 1993. Species richness and canopy productivity of Australian plant communities. – *Biodivers. Conserv.* 2: 152–167.

- Stein, A. et al. 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. – *Ecol. Lett.* 17: 866–880.
- Stevenson, L. A. et al. 2012. Species richness and endemism of Australian bryophytes. – *J. Bryol.* 34: 101–107.
- Storch, D. et al. 2018. The more-individuals hypothesis revisited: the role of community abundance in species richness regulation and the productivity–diversity relationship. – *Ecol. Lett.* 21: 920–937.
- Storlie, C. J. et al. 2013. Improved spatial estimates of climate predict patchier species distributions. – *Divers. Distrib.* 19: 1106–1113.
- TERN 2022. TERN AEKOS. – <<http://aekos.org.au>>.
- Thornhill, A. H. et al. 2016. Continental-scale spatial phylogenetics of Australian angiosperms provides insights into ecology, evolution and conservation. – *J. Biogeogr.* 43: 2085–2098.
- Večeřa, M. et al. 2021. Mapping species richness of plant families in European vegetation. – *J. Veg. Sci.* 32: e13035.
- Veech, J. A. et al. 2002. The additive partitioning of species diversity: recent revival of an old idea. – *Oikos* 99: 3–9.
- Viscarra Rossel, R. A. et al. 2015. The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project. – *Soil Res.* 53: 845–864.
- Webb, L. J. et al. 1984. A floristic framework of Australian rainforests. – *Aust. J. Ecol.* 9: 169–198.
- Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. – *Ecol. Monogr.* 30: 280–338.
- Williams, K. J. et al. 2016. Assessing the ecological representativeness of Australia's terrestrial National Reserve System: a community-level modelling approach. – CSIRO.
- Wiser, S. K. et al. 2011. Veg-X – an exchange standard for plot-based vegetation data. – *J. Veg. Sci.* 22: 598–609.
- Wood, S. N. 2017. Generalized additive models: an introduction with R (2nd ed.). – Chapman and Hall.
- Worm, B. et al. 2006. Impacts of biodiversity loss on ocean ecosystem services. – *Science* 314: 787–790.
- Xu, T. and Hutchinson, M. 2010. ANUclim ver. 6.1 user guide. – Fenner School of Environment and Society, The Australian National Univ.