ECOGRAPHY

Research

When macroecological transitions are a fiction of sampling: comparing herbarium records to plot-based species inventory data

Greg R. Guerin, Alan N. Andersen, Maurizio Rossetto, Stephen van Leeuwen, Margaret Byrne, Ben Sparrow, Michelle Rodrigo and Andrew J. Lowe

G. R. Guerin (http://orcid.org/0000-0002-2104-6695) (greg.guerin@adelaide.edu.au), B. Sparrow and M. Rodrigo, Terrestrial Ecosystem Research Network, School of Biological Sciences, Univ. of Adelaide, North Terrace, Adelaide, SA, Australia. – A. N. Andersen, Research Inst. for the Environment and Livelihoods, Charles Darwin Univ., Darwin, NT, Australia. – M. Rossetto, National Herbarium of NSW, Royal Botanic Garden and Domain Trust, Sydney, NSW, Australia. – S. van Leeuwen and M. Byrne, Science and Conservation, Kensington, WA, Australia. – A. J. Lowe, Centre for Conservation Science and Technology, School of Biological Sciences, Univ. of Adelaide, North Terrace, Adelaide, SA, Australia.

Ecography
41: 1864–1875, 2018
doi: 10.1111/ecog.03607

Subject Editor: Miguel Nakamura
Editor-in-Chief:
Jens-Christian Svenning
Accepted 20 January 2018

Natural history collections are alternative data sources to plot-based species inventories for analysing macroecological species turnover. Herbarium records sample diversity well at regional level and are taxonomically validated. However, they are ad hoc from a sampling perspective, generating spatial and taxonomic biases. The implications of biased sampling on beta diversity (β) estimation, and use of herbarium data to identify macroecological transitions, remain unexplored. We tested sampling influences by comparing herbarium data with systematically collected inventory data from the Mount Lofty–Flinders Ranges region of Australia. We calculated β within moving windows across bioclimatic gradients using metrics varying in sensitivity to richness differences (pairwise/multi-site Sørensen β; Simpson β; Harrison et al. β_{-2}), and correlated β to species sampling and between herbarium and plot data. We tested whether generalised dissimilarity modelling (GDM) revealed the same compositional transitions in herbarium and plot data along environmental gradients. Sørensen, Simpson and multi-site Sørensen β had strong negative correlations with richness (indicating sampling bias) for herbarium data (Pearson's $r = -0.85, -0.80, -0.81$, respectively) but not plots ($r = -0.27, -0.28, -0.11$). Harrison et al. β_{-2} correlated poorly with richness (herbarium: $r = -0.16$; plots: $r = -0.14$) but herbarium and plot data were only weakly correlated ($r = 0.18$). All other metrics correlated poorly ($-0.03 < r < 0.16$) between datasets, suggesting biases. GDMs differed in variable importance but revealed similar transition zones for key gradients. We conclude that untransformed herbarium data are unsuitable for detecting macroecological transitions because turnover is linearly related to sampling intensity and correlates poorly with systematic surveys. Herbarium data should be used cautiously for β, even with methods insensitive to richness differences. However, herbarium data can robustly reproduce transition zones when modelled along environment gradients. We recommend this approach for detecting macroecological transitions using natural history data in the absence of plot data.

Keywords: beta diversity, macroecological transitions, natural history collections

NORDIC SOCIETY OIKOS

www.ecography.org

Introduction

Natural history collections have become a widely-used source of species occurrence data for macroecological analysis (Jácome et al. 2007, Engemann et al. 2015, 2016, Haque 2015, Hallgren et al. 2016, Robertson et al. 2016, Nualart et al. 2017), including the identification of climatic drivers of species distributions, diversity and composition (Wernberg et al. 2011, Fitzpatrick et al. 2013, Franklin et al. 2017) and phytogeographic regionalisation (Kooyman et al. 2013, González-Orozco et al. 2014, Gioia and Hopper 2017). Collections provide benefits, such as extensive spatial and temporal coverage, that cannot feasibly be replicated by new surveys, as well as verifiable species identity from physical samples and, ideally at least, a consistent and up-to-date taxonomy. Large databases containing these records are now routinely available online through open access data portals (Graham et al. 2004, Haque 2015, Hallgren et al. 2016).

However, natural history collections tend to be spatially and taxonomically biased due to uneven sampling intensity and scope (Engemann et al. 2015, Haque 2015). This has raised concerns as to the validity and appropriate use of such data for inferring biogeographic patterns (Hortal et al. 2007). In particular, the reliability of collections data for assessing patterns of beta diversity (β) under some circumstances is open to question. Such assessments are critical for gaining a robust understanding of climatically driven ecological transition zones (Fitzpatrick et al. 2013), as a basis for forecasting ecological impacts of climate change (Engemann et al. 2015), and for planning appropriate climate-sensitive adaptation and restoration strategies (Harris et al. 2006).

A recurring limitation of using specimen data for regional planning is the treatment of areas that have received little or no sampling. Model-based approaches have been proposed that predict biodiversity patterns based on environmental relationships developed from better-sampled grid cells (Albuquerque and Beier 2016). However, it has been demonstrated that when species are under-sampled, β estimates can become artificially high (Harrison et al. 1992, Cardoso et al. 2009, Guerin et al. 2016). Extrapolation from well-sampled grid cells can be particularly unreliable if sampling intensity is itself correlated with environmental features. In particular, remote and sparsely populated areas often represent unique environmental domains, yet are typically under-sampled.

There have been several attempts to deal with perceived biases in natural history collection data for macroecological analysis (Rowe 2005, Maldonado et al. 2015, Robertson et al. 2016), including accounting for sampling bias in observed species richness and turnover (Mokany et al. 2012, Engemann et al. 2015, Burley et al. 2016, Gioia and Hopper 2017) and spatial gaps in sampling (Albuquerque and Beier 2016). Many studies use common β metrics, such as Bray–Curtis dissimilarity, that have good ecological properties but are sensitive to sampling differences, as default parameters (DeWalt et al. 2005, Gavin 2009) with little qualification. Others deliberately use metrics that should be less sensitive to sampling because they ignore richness differences (Burley et al. 2016, Gioia and Hopper 2017), but do not always assess whether biases are alleviated, or account for the influence of absolute observed richness rather than richness difference (González-Orozco et al. 2014). These approaches all have limitations (Robertson and Barker 2006, Hortal et al. 2007), but there have been few attempts to test the reliability of species turnover assessments from collections or atlas-type data by comparison with data from systematic species inventory plots or relevés. This is a critical gap given general acknowledgement that systematically sampled datasets are less susceptible to the biases of more opportunistic or selective voucher sampling (Franklin et al. 2017).

Here we investigate these issues by addressing the following questions: 1) are estimates of raw plant species turnover (according to various metrics) influenced by sampling intensity? 2) Are turnover estimates derived from herbarium data congruent with those from systematic, plot-based sampling? 3) Are differences in mapped transitions between raw herbarium and plot-based data alleviated by modelling species turnover along environmental gradients?

The motivation for the study is to understand the limitations of using natural history records (specifically, herbarium data) in macroecology and to assess which methods are more robust. We addressed the questions for a study region traversing a major bioclimatic gradient in South Australia, making empirical comparisons between species records from herbarium data (140 624 records) and a set of systematically surveyed plant species composition plots from the Biological Survey of South Australia, which were stratified to sample habitat diversity (125 085 records). We examined species turnover among map grid cells in relation to sampling intensity and tested whether any discrepancies between the two data sources are alleviated by: a) choice of β metric; and b) extracting compositional variance with respect to a set of key gradients in a model setting.

Methods

Study area

The study area was defined by three contiguous bioregions within South Australia: Kanmantoo, Flinders Lofty Block and Murray Darling Depression (Thackway and Cresswell 1995) that are ecologically diverse and well-sampled by herbarium collections and vegetation survey plots (Fig. 1). The study area is topographically diverse, covers strong gradients in climate and soil type, and is a centre of plant biodiversity for South Australia (Guerin et al. 2016). It is traversed by a large-scale monitoring transect (TREND) established to investigate spatial and temporal transitions in ecological communities, particularly with respect to climatic changes (Guerin et al. 2014, Caddy-Retalic et al. 2017).

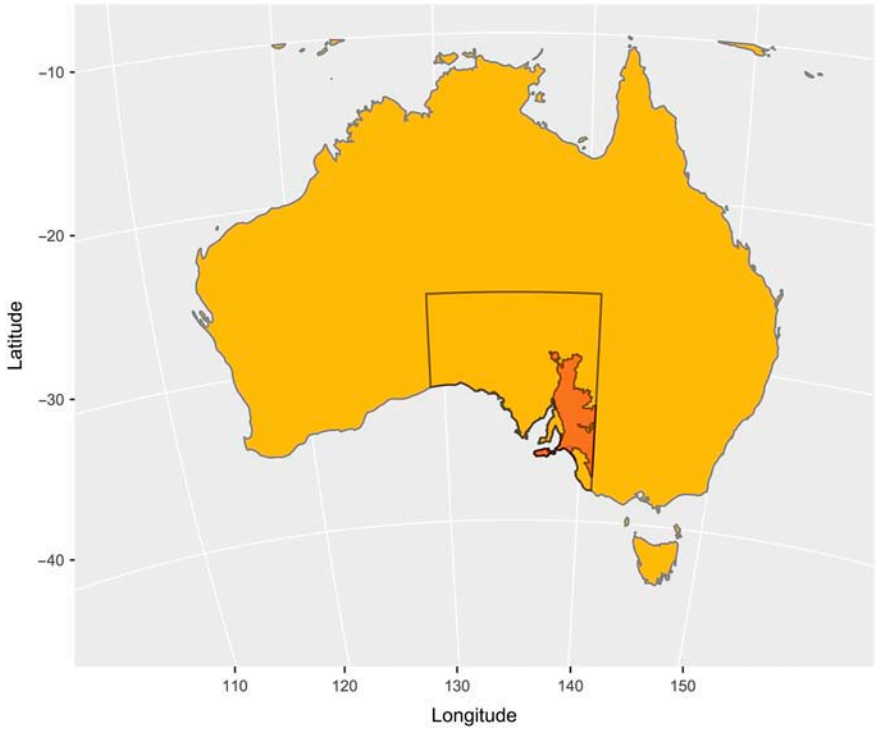

Figure 1. Location of the study area in the context of the State of South Australia and the Australian continent.

Datasets

We compared a set of macroecological metrics calculated from the vascular plant species occurrence records of two datasets: 1) 'herbarium': herbarium records obtained via Australia's virtual herbarium that were filtered for spatial validity and certain species level identification, giving a dataset of 140 624 records of 3586 species across all habitat types (AVH 2014) and; 2) 'plots': systematically collected species inventories from field survey plots of the Biological Survey of South Australia supplemented by those from Terrestrial Ecosystem Research Network AusPlots (Dept of Environment, Water and Natural Resources 2012, Guerin et al. 2017), giving a dataset of 125 085 records of 2144 species from 4740 plots, collected between 1977 and 2012. Most of the plots could be described as 'preferential' as they were neither strictly randomly located nor stratified to a sampling grid and therefore sample habitats somewhat subjectively. However, the plots were generally arranged to deliberately sample the diversity of environments and vegetation types, and during surveys all observable plant species were recorded within the plot, including both dominant and rare species. Collectively, the surveys are expected to be a good representation of the species diversity present in an area (Robinson et al. 1988).

Bioclimatic predictor variables for macroecological modelling were sourced from WorldClim (Hijmans et al. 2005) and soil and landscape variables from the Soil and Landscape Grid of Australia (Grundy et al. 2015). Variables were selected that covered the major gradients in climate and soil properties expected a priori to be important for predicting plant species composition, including rainfall, temperature, seasonality, soil fertility, pH and landscape factors (Table 1).

Influence of sampling intensity, data type and choice of metric on turnover estimates

Mapping framework and examination of sampling influences

To address our aim of informing selection of β metrics that are robust to sampling for macroecological applications, we compared species turnover patterns to test for differences in estimates due to sampling intensity and data type (i.e. whether 'herbarium' or 'plots'). Because herbarium records are spatially scattered, they are only comparable to plot-based records at coarse spatial scales rather than specific sites. The high intensity of sampling in the study region allowed gridded map analysis at a resolution of 0.1° (x, y) in which grid cells, each covering an area of approximately 103 km^2 (depending on latitude), were treated as community sample units. This resulted in species data for 974 cells for 'plots' and 1389 cells for 'herbarium'. Initially, sampling intensity (number of specimens or plots) and species richness were mapped to assess the potential for sampling influences over β estimates. Scatterplots and correlation tests were used to relate variables. An expected monotonic relationship between number of samples and species richness would indicate that any β metric sensitive to richness is biased whenever sampling is incomplete.

Table 1. Details of variables included in generalised dissimilarity models (GDM) of plant beta diversity in South Australia (Hijmans et al. 2005, Grundy et al. 2015).

Variable	Code	Resolution
Spatial variables		
1. Geographic distance	Geo	
Bioclimatic variables:		
2. Mean annual temperature (°C)	Bio1	30 arc seconds
3. Temperature seasonality	Bio4	"
4. Mean maximum temperature of the warmest month (°C)	Bio5	"
5. Mean minimum temperature of the coolest month (°C)	Bio6	"
6. Mean temperature of the wettest month (°C)	Bio8	"
7. Mean annual rainfall (mm)	Bio12	"
8. Precipitation seasonality	Bio15	"
9. Mean rainfall of the driest month (mm)	Bio17	"
Soil and landscape variables:		
10. Total nitrogen (%)	N	3 arc seconds
11. Total phosphorus (%)	P	"
12. Total carbon (%)	C	"
13. pH ($CaCl_2$)	pH	"
14. Topographic relief within 1000 m (m)	Relief	"
15. Available water capacity (%)	AWC	"
16. Sand (%)	Sand	"
17. Effective cation exchange capacity (mEq per 100 g)	ECE	"
18. Depth of soil (m)	Depth	"
19. Topographic slope (%)	Slope	"

β metrics

Species occurrence data collected via systematic sampling schemes, including plot-based species inventories, enable the use of non-parametric estimators of true species richness and β (Gotelli and Colwell 2011, Cardoso et al. 2015). β estimation methods can also be applied to grid cells by either: 1) comparing cumulative samples within cells for pairwise cell–cell β comparisons that account for incomplete sampling within cells; or 2) treating cells themselves as cumulative samples to estimate gamma diversity (i.e. the total number of species in a set of grid cells) and therefore estimate Whittaker's beta diversity by dividing estimated gamma diversity by mean cell alpha diversity.

Estimation approaches are particularly problematic when applied to herbarium data, because herbarium sampling violates several key assumptions (Tobler et al. 2007, Guerin et al. 2016, Gioia and Hopper 2017), especially that samples are random and independent (Gotelli and Colwell 2011). In theory, gamma diversity estimators could be applied to collective herbarium records within a set of grid cells. However, subsequent estimation of beta diversity requires a good estimate of alpha diversity (i.e. the mean of expected species richness within each grid cell), which is not robust for incomplete herbarium sampling. Due to the inherent limitations of obtaining valid β estimates from herbarium datasets in particular, we used raw scores in order to compare the direct effects of sampling between the two data types.

Calculations of β used binary species presence/absence within a moving 3×3 cell window. Scores for individual cells were calculated based on comparison with the neighbouring eight cells, so that high scores indicated high spatial turnover across that location. Two approaches were used to calculate turnover. Firstly, the mean of pairwise dissimilarities among cells within the window was calculated based on metrics known to have different sensitivities to richness differences: Sørensen ($\beta_{SØR}$; includes a richness difference component), Harrison et al. β_{-2} (reportedly robust to under-sampling), and Simpson (β_{SIM}; species replacement index) (Harrison et al. 1992, Jost 2007, Cardoso et al. 2009, Baselga 2010, Legendre and De Cáceres 2013). Calculations were repeated for selected beta diversity metrics with the contribution of individual cells in the moving window to the mean weighted by observed species richness, to minimise the influence of very poorly sampled cells. Secondly, multi-site dissimilarity was calculated within the window using the multi-site versions of $\beta_{SØR}$ (for total turnover) and β_{SIM} (replacement only component) (Baselga and Orme 2012, Baselga 2013).

The correlation of β scores for each metric to observed mean species richness within the window was calculated using Pearson's r. Metrics were also compared to one another, as were scores from 'herbarium' versus 'plots'. A β metric that is robust to sampling would be expected to have no relationship with richness and to produce results that are highly correlated between 'herbarium' and 'plots'.

Influence of data type on turnover models

To address the question of whether modelling species turnover along environmental gradients provides more congruent results between datasets by excluding sampling noise, we used generalised dissimilarity modelling (GDM; Ferrier et al. 2007). GDM is a correlative modelling procedure in which turnover is estimated via non-linear transformations of predictor variables that recognise the curvilinear relationship

between ecological and environmental distances, and that turnover rates vary along gradients. We compared GDMs for 'herbarium' and 'plots' with species composition in grid cells as the response variable.

High colinearity among predictor variables can create statistical artifacts (Zuur et al. 2010). There are several valid methods for assessing variable colinearity, including factor analysis, bivariate scatterplots, and correlation coefficients (Zuur et al. 2010). We elected to use raw predictor variables rather than uncorrelated PCA axes representing environmental variation because we were interested in the influence of particular variables. However, we assessed variable pairs and excluded the weakest performing variable of pairs that were highly colinear (Pearson's $r > 0.8$; Dahlgren and Ehrlén 2009, Belmaker and Jetz 2011, Carlson et al. 2016).

We allowed five splines in the non-linear responses and included geographic distance as a co-variable to account for spatial effects on species turnover and spatial autocorrelation. The number of variables in the model was reduced following a backwards selection procedure. Variables were initially removed if their coefficients summed to zero, indicating they do not contribute to species turnover. The reduced set of variables was further vetted by removing those not deemed statistically significant according to the non-parametric

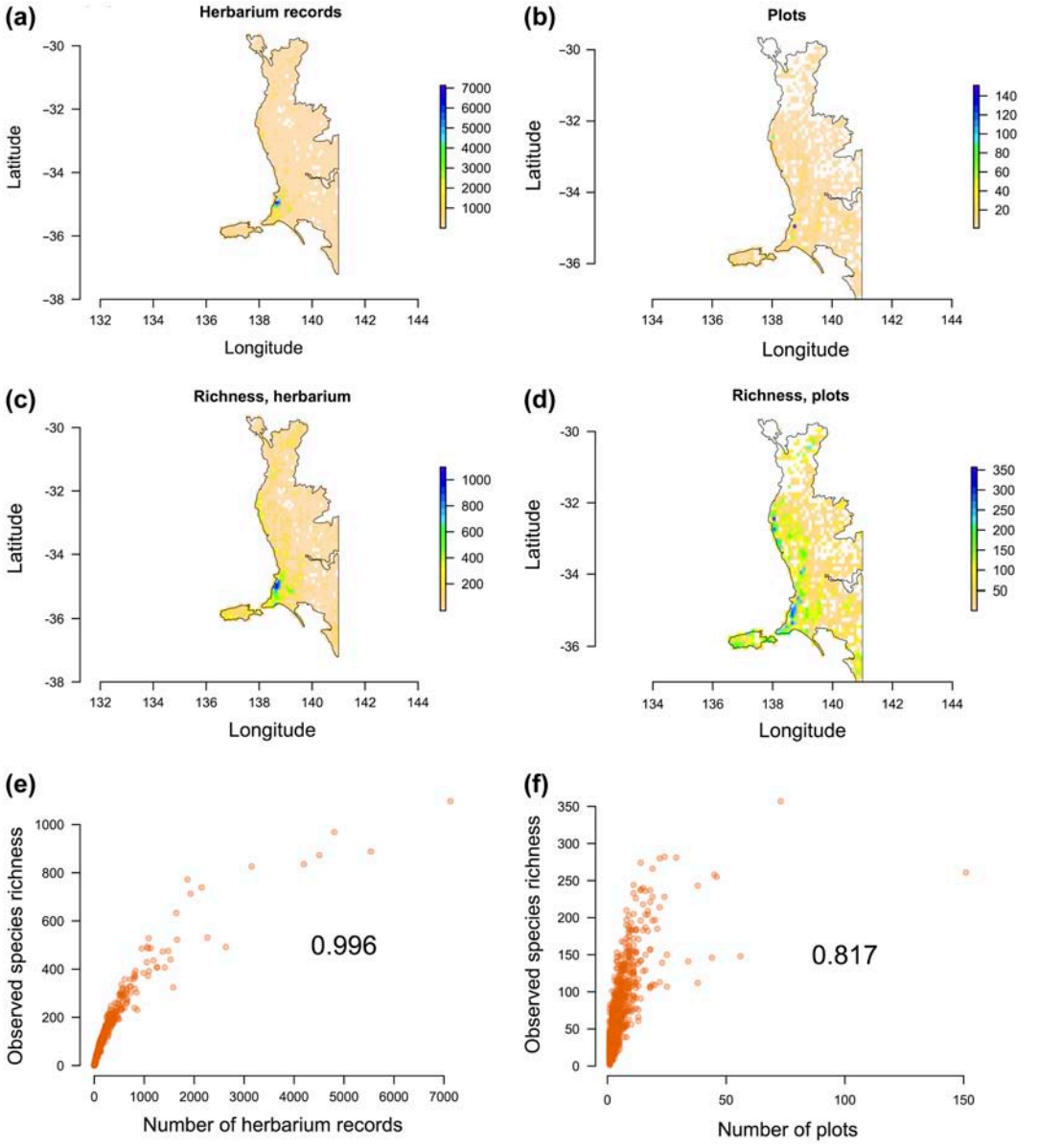

Figure 2. Study area in South Australia showing vascular plant sampling intensity for; (a) herbarium records (number of records); (b) vegetation survey plots (number of plots); and associated species richness for; (c) herbarium and; (d) plots. Panels (e) and (f) show the relationship between observed species richness and sampling intensity in grid cells for; (e) herbarium and; (f) plots. Spearman's ρ is shown to quantify the correlation between number of samples and richness.

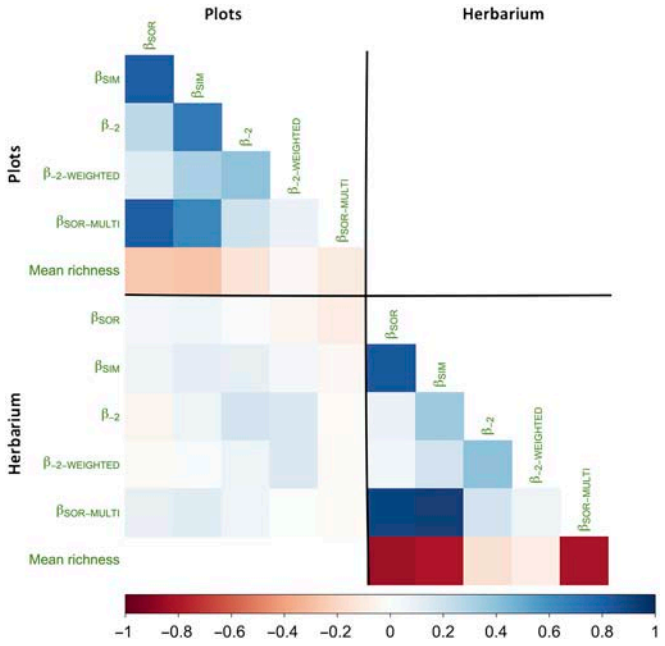

Figure 3. Pairwise correlations (Pearson's r) among beta diversity metrics and species richness for plant species occurrences in two datasets, 'herbarium' and 'plots', from South Australia, based on moving windows. Metrics for 'herbarium' were more correlated with species sampling, and correlated poorly with 'plots'. The least biased metric overall was Harrison et al. β_{-2}.

test of Leitão et al. (2017). Variables remaining in the final models were ranked by importance according to summed coefficients, which is equivalent to the height of the associated function and represents the total amount of species turnover along that gradient.

We repeated GDM using individual field plots (i.e. 'local sites' instead of grid cells) as the response communities, to compare model performance. We randomly selected 1000 plots for model training to increase efficiency and to train the model with approximately the same number of sites as cells in the gridded analyses (i.e. 974).

Final models were compared according to the independent selection and ranking of variables, the shape of response curves for common predictors (height and gradient subintervals with the highest turnover rate as assessed through rolling calculations of slope across gradient segments), and overall patterns of predicted compositional similarity (based on visualisation of transformed environmental variables). A method completely robust to input data type should result in the same selection of variables, zones of rapid turnover along gradients and pattern of composition. However, some variation in models could be due to differences in the spatial extent of grid cells included (resulting in different points and subintervals along gradients being sampled) and to the inclusion of different covariates. Therefore, we repeated GDM for both datasets with the training area restricted to the same set of common grid cells and using a reduced set of variables common to both models.

Software

Analyses were conducted using custom programming in R (ver. 3.1.3 and 3.3.0; R Core Team) using functions from packages 'wCorr', 'betapart', 'vegan', 'gdm', 'sgdm', 'raster' and 'simba' (Baselga and Orme 2012, Jurasinski and Retzer 2012, Hijmans 2015, Manion et al. 2016, Oksanen et al. 2016, Emad and Bailey 2017, Leitão et al. 2017).

Results

Influence of sampling intensity, data type and choice of metric on turnover estimates

Species richness was highly dependent upon sampling intensity for 'herbarium' and 'plots', with an almost perfect rank correlation for 'herbarium' ($\rho = 0.996$; Fig. 2). Sampling intensity and species richness revealed idiosyncratic spatial patterns between datasets with herbarium records skewed towards population centres in the south-west of the study area (Fig. 2).

Pairwise and mutli-site $\beta_{SØR}$ and β_{SIM} were linearly related to mean species richness for 'herbarium' ($-0.85 < r < -0.80$) but less so for 'plots' ($-0.28 < r < -0.11$; Fig. 3; Fig. 4). Harrison et al. β_{-2} was less correlated with species richness for both datasets (Fig. 4), indeed the correlation with richness was very low when sparsely sampled cells were sequentially down-weighted (Fig. 3). However, Harrison et al. β_{-2} was somewhat heteroskedastic for 'herbarium', with less variance among well-sampled cells (Fig. 4). 'Herbarium' correlated poorly with 'plots' for each metric and therefore produced strongly contrasting spatial patterns (Fig. 3; Fig. 4). The highest correlation between datasets was for Harrison et al. β_{-2} ($r = 0.18$) and the lowest was for mutli-site $\beta_{SØR}$.

Influence of data type on turnover models

GDMs fitted independently with a variable selection process resulted in overlapping but not identical predictors and importance ranks (Table 2). Final models explained 49% of deviance for 'plots' (41% for local sites) and 25% for 'herbarium'. For each model, mean annual rainfall was the most important predictor followed by geographic space, while all other predictors were ranked differently.

For GDMs repeated using a common spatial extent and predictors, the following variables were included: geographic space, mean annual rainfall, mean maximum temperature of the warmest month, mean temperature of the wettest quarter, soil percent sand and topographic relief. For both standardised models ('herbarium' and 'plots'), mean annual rainfall was the most important predictor followed by geographic space, while the ranking of the remaining variables differed (Table 2). The functions fitted to each variable were similar in height and shape (Fig. 5). Differences in fitted functions between 'herbarium' and 'plots' were comparable in magnitude to differences between independent versus

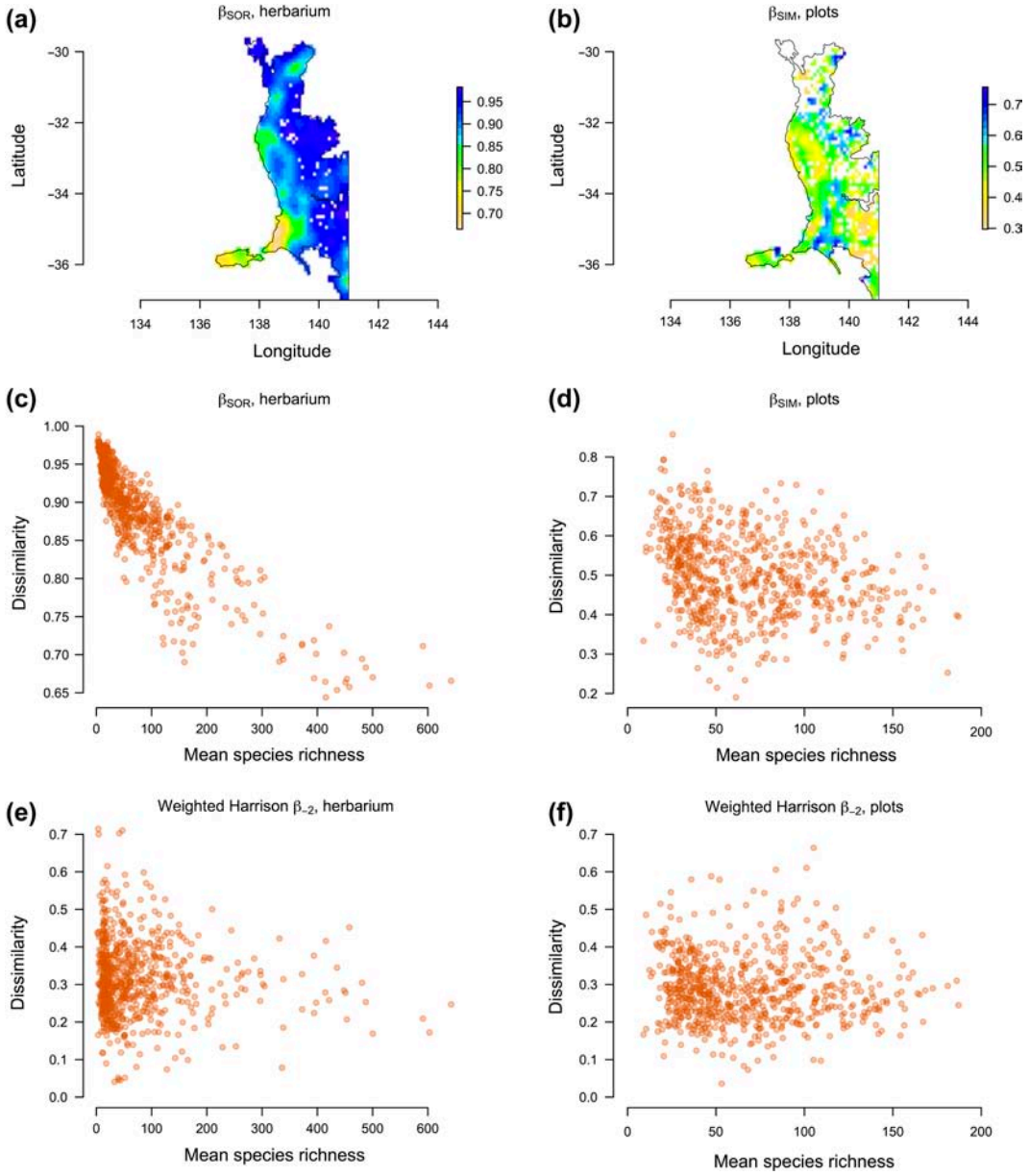

Figure 4. Relationship between estimated vascular plant beta diversity and observed species richness for metrics with the lowest and highest correlations (Fig. 3). (a)–(b) raster maps of study area in South Australia: (a) $\beta_{SØR}$ (herbarium); (b) β_{SIM} (plots); (c)–(f) scatterplots: (c) $\beta_{SØR}$ against richness (herbarium); (d) β_{SIM} against richness (plots); (e) Harrison et al. β_{-2} weighted by richness (herbarium); (f) Harrison et al. β_{-2} weighted by richness (plots).

common variables for each dataset, and to the differences between plots fitted to grid cells versus local sites.

The similarity in functions resulted in similar predicted compositional similarity (Fig. 6). Mapping of transition zones (gradient subintervals with the highest species turnover rates) revealed overlapping zones for most variables, excepting differences for Bio5 (mean maximum temperature of the warmest month) and Bio8 (mean temperature of wettest quarter) (Table 2; Fig. 5). For mean annual rainfall, all models revealed a rapid transition zone at the arid extreme of the gradient leading into higher rainfall areas (Fig. 5). However, 'herbarium' models predicted the transition to be steepest at lower rainfall subintervals of approximately 200–300 mm compared to approximately 250–350 mm for 'plots'.

Discussion

Are turnover estimates influenced by sampling intensity and type?

A key finding of this study was that the intensity of species sampling in grid cells biases turnover estimates. Most turnover estimates were strongly negatively correlated with the

Table 2. Statistics for generalised dissimilarity models fitted to herbarium records versus plot-based plant species inventory data within grid cells in South Australia. Models were fitted independently or using a common set of variables and spatial extent ('standardised'). Bold values indicate function height (amount of species turnover), values in brackets indicate rank importance and ranges indicate transition zones defined as gradient subintervals with the highest turnover rate.

Variable	Herbarium (grid cells)	Plots (grid cells)	Plots (local sites)	Herbarium (grid cells) – standardised	Plots (grid cells) – standardised
Deviance explained	25%	49%	41%	41%	52%
Geographic distance	**1.63** (2nd) 0–108.11	**1.18** (2nd) 0–98.85	**1.43** (2nd) 0–89.16	**1.07** (2nd) 0–98.85	**1.10** (2nd) 0–98.85
Mean temperature of the warmest month (Bio5)	**0.47** (5th) 35.53–37.30	**0.34** (5th) 34.46–36.10	**0.84** (5th) 26.64–28.28	**0.48** (5th) 30.15–31.79	**0.67** (4th) 34.46–36.10
Mean minimum temperature of the coldest month (Bio6)	–	**0.18** (8th) 4.63–5.67	**0.14** (11th) 4.93–6.02	NA	NA
Mean temperature of the wettest quarter (Bio8)	**0.20** (7th) 20.98–23.46	**0.32** (6th) 25.05–27.40	**0.22** (10th) 9.81–12.22	**0.11** (6th) 13.58–15.94	**0.26** (6th) 25.05–27.40
Mean annual rainfall (Bio12)	**1.74** (1st) 198.80–302.28	**2.00** (1st) 255.22–357.50	**2.19** (1st) 255.03–363.69	**2.11** (1st) 199.83–302.10	**1.95** (1st) 242.44–344.71
Soil sand (%)	**0.67** (3rd) 76.13–82.16	**0.69** (3rd) 76.13–82.16	**0.78** (6th) 71.20–77.17	**0.48** (4th) 76.63–82.66	**0.75** (3rd) 76.13–82.16
Soil available water capacity (%)	–	**0.24** (7th) 10.53–11.79	**0.37** (8th) 10.87–11.96	NA	NA
Topographical relief within 1000 m (m)	**0.32** (6th) 2.71–55.89	**0.41** (4th) 2.71–55.07	**0.32** (9th) 2.16–71.87	**0.62** (3rd) 2.71–55.07	**0.55** (5th) 2.71–55.07
Topographic slope (%)	**0.11** (8th) 0.04–9.02	–	**0.85** (4th) 0.025–9.32	NA	NA
Soil total phosphorus (% w/w)	–	**0.05** (9th) 2.05×10^{-2} – 2.48×10^{-2}	**0.60** (7th) 4.36×10^{-2} – 4.83×10^{-2}	NA	NA
Soil effective cation exchange capacity (mEq per 100 g)	**0.51** (4th) 14.61–17.54	–	–	NA	NA
Soil pH ($CaCl_2$)	–	–	**1.37** (3rd) 4.55–4.93	NA	NA

number of species sampled for herbarium data but less so for plot data. A linear relationship between richness and turnover is a critical problem for macroecological application of these data types because observed richness is an artifact of sampling intensity. Herbarium sampling breaks key assumptions of richness estimators (Tobler et al. 2007), which makes robust corrections for sampling biases problematic.

Importantly, the confounding relationship between turnover and sampling held for metrics that do not quantify richness differences per se, such as Simpson β, which should be more robust to under-sampling. Of candidate β metrics tested, Harrison et al. β_{-2} was the least correlated with species richness, suggesting independence from sampling intensity. The correlation between herbarium and plot data was also the highest for Harrison et al. β_{-2}, although relatively low at 0.18. We therefore verify the findings of Cardoso et al. (2009) for sparse plant data that Harrison et al. β_{-2} is more robust than other metrics to under-sampling. However, the failure of the metric to reproduce the same turnover patterns between datasets implies sampling still overrides the real ecological pattern.

Plot-based species records from vegetation databases are an important resource for ecological analysis (Roleček et al. 2007) but may also suffer from biases of subjectivity. Plots located preferentially have been found to inflate estimates of beta diversity compared to stratified-random sampling (Michalcová et al. 2011) and to bias species richness estimates (Chytrý 2001). However, they also capture more diversity in species and environments (Roleček et al. 2007). Additionally, species identifications may be difficult to verify and cross-validate across surveys unless specimens are vouchered. For our plot data, $\beta_{SØR}$ and β_{SIM} both correlated somewhat negatively with observed species richness, which could represent both real patterns and inflated turnover estimation where sampling was sparse. Wherever species richness estimates may be biased by sampling intensity or preferential location, we recommend Harrison et al. β_{-2} or multi-site $\beta_{SØR}$.

The incongruence in raw beta diversity estimates between herbarium and plot datasets and, specifically, the influence of spatial collection intensity of herbarium samples on observed richness (Engemann et al. 2015) and beta diversity suggest that herbarium records – and, by extension, other natural history collection records – should be used with caution for inferring macroecological patterns (Hortal et al. 2007, Burley et al. 2016). Although potential biases are noted regularly (Vale and Jenkins 2012, Engemann et al. 2015, Haque 2015), these datasets are used frequently for macroecological applications, with varying consideration of sampling influences. Our results suggest that even the use of metrics insensitive to richness differences (González-Orozco et al. 2014, Gioia and Hopper 2017), and reputedly robust to under-sampling (Cardoso et al. 2009), may not alleviate the problem.

Dissimilarity can be inflated for under-sampled grid cells (Burley et al. 2016) but the issue is not only sampling evenness, which can be improved by sub-sampling highly sampled cells and excluding poorly sampled ones. For example, Gioia

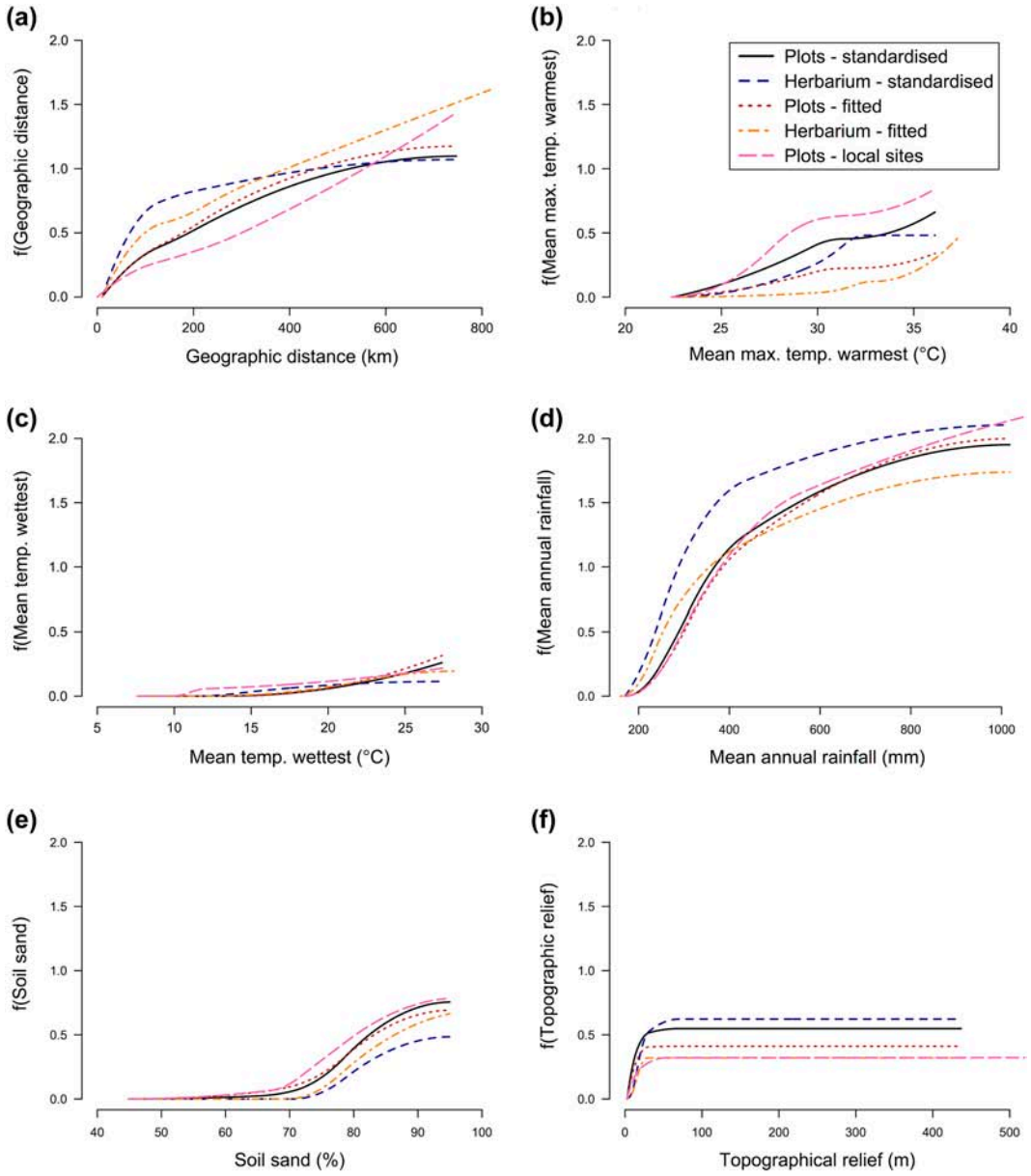

Figure 5. Predicted plant species turnover among grid cells in South Australia from fitted generalised dissimilarity models (Table 2) for predictors common to all models. Slope represents the instantaneous rate of species turnover while function height on the y-axis (same scale for all) equals the total amount of species turnover along each gradient. Response are overlaid for models fitted independently ('fitted') or 'standardised' by fitting to a common subset of variables and geographic extent for plots and herbarium (legend): (a) geographic distance; (b) mean maximum temperature of the warmest month (Bio5); (c) mean temperature of the wettest quarter (Bio8); (d) mean annual rainfall (Bio12); (e) soil sand; (f) topographical relief within 1000 m.

and Hopper (2017) dealt with uneven herbarium sampling effort by limiting β_{SIM} comparisons to cells with 50 or more species and sub-sampling those with more than 500 species. However, for our data, a strong linear sampling effect on β_{SIM} remained for cells with mean richness of 50–500 species, and scores did not correlate with those estimated from plot data. However, the implications for applications such as clustering for bioregionalisation based on cell-cell dissimilarities (González-Orozco et al. 2014, Gioia and Hopper 2017) require further exploration, for example as to optimal cell size.

Do models of composition along environmental gradients predict the same transitions?

Model-based gradient analysis produced more similar results between herbarium and plot datasets but with exceptions. Comparisons had three components: model selection, variable importance and location of transition zones. For models fitted using a flexible but objective procedure, there were differences in included variables for herbarium and plot data as well as for plots fitted as local sites. For example, mean

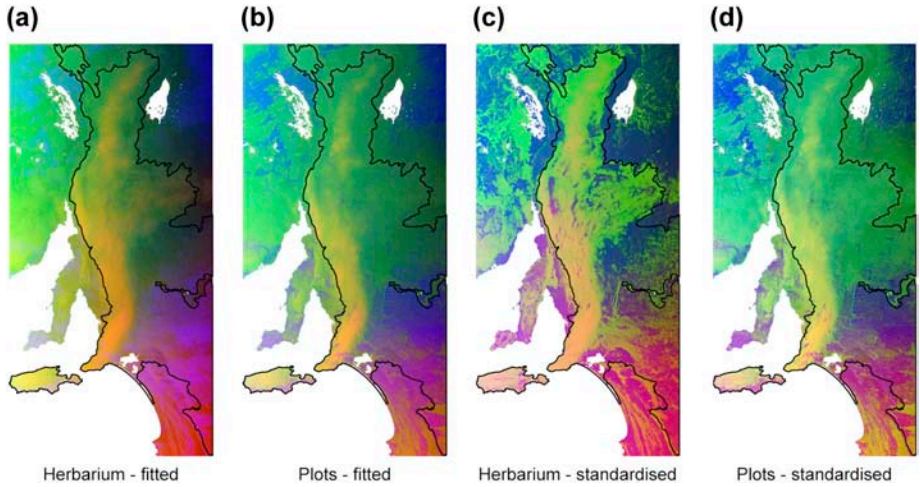

Figure 6. Compositional similarity in vascular plant species predicted by generalised dissimilarity models in South Australia represented as the first three axes of a PCA of model-transformed predictors on an RBG colour-scale (similarity in colour represents similarity in species composition): (a)–(b) models fitted independently for: (a) herbarium; (b) plots; (c)–(d) models standardised by fitting to a common subset of variables and geographic extent for: (c) herbarium; (d) plots.

minimum temperature of the coldest month, soil available water capacity and soil total phosphorus were retained as predictors for plot but not herbarium data. Nevertheless, six predictors were retained in all models and mean annual rainfall was consistently the most important, followed by geographic space. Rank importance for variables otherwise differed among models.

Clearly, some of the differences in variable selection and importance in independently fitted models were due to differences in the spatial extent of the datasets and the subsequent sampling of gradients, as well as the influence of different covariables. However, the results for models repeated using common spatial extents and predictor variables were reasonably similar to independently fitted models, with mean annual rainfall and geographic space the top predictors, and rank importance of the remaining variables differing among models.

A key finding was that modelled transition zones matched well between datasets, even when the relative contribution of predictors differed. Three transition zones that particularly stand out were between 200–350 mm of mean annual rainfall, from plains to topographically complex areas with higher relief and, finally, from loam to very sandy soils. While transitions were more poorly matched for mean temperature of the wettest quarter, this variable consistently contributed relatively little to turnover.

Deviance explained by independent models was lower for herbarium (25%) than plot (49%) data. This may indicate more 'noise' not explained by environmental patterns, although deviance explained was higher for 'standardised' herbarium models. The relatively high deviance explained for plots suggests that assessing beta diversity with respect to environment can bypass some of the issues associated with historical species occurrence data. However, the model-based approach left about half of species turnover unexplained, which prevents mapping zones of high total turnover and places limits on predictive power. Also, the model predictions are transformations of environmental variables that should be interpreted as ecologically scaled gradient-space rather than predictions of actual species composition (Prober et al. 2012). Unexplained variance may be due to factors such as incomplete sampling, disturbance, history or chance. Therefore the relevance of beta diversity estimates made with respect to gradients depends on the application.

Conclusions

Untransformed herbarium data may be unsuitable for detecting true macroecological transitions because turnover metrics are linearly related to sampling intensity and do not correlate with those calculated from systematic surveys. Such approaches should be interpreted conservatively because reported biodiversity patterns are only meaningful if they hold regardless of sampling. We found that even metrics that ignore richness differences remain biased by sampling and no metric provided a robust correlation between herbarium and plot data. Although plot-based species inventories are not necessarily unbiased, they appear to provide turnover estimates more robust to sampling. Turnover estimates, and their sampling biases, are expected to be somewhat scale-dependent, and selection of an appropriate resolution for analysis is also a consideration.

Nevertheless, the collection of species occurrence data is costly to replicate, and given the need for macroecological data for conservation planning, we can ill-afford to ignore collections as a resource. Incongruence between herbarium and plot data can be alleviated with model-based approaches such as GDM because environment and space are important determinants of turnover. GDMs appear to robustly indicate ecological transition zones, although they can only be

interpreted as ecologically-scaled environmental space and do not explain all spatial turnover. We therefore recommend this as an approach for detecting macroecological transitions along major bioclimatic gradients. Further exploration of optimising spatial resolution and sub-sampling techniques to improve estimates form herbarium data is also warranted.

Acknowledgements – We thank the Terrestrial Ecosystem Research Network (TERN) supported by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS), the Australian Transect Network and the original collectors of the data used.

References

Albuquerque, F. and Beier, P. 2016. Predicted rarity-weighted richness, a new tool to prioritize sites for species representation. – Ecol. Evol. 6: 8107–8114.

AVH 2014. Australia's virtual herbarium. – Council of Heads of Australasian Herbaria, <http://avh.chah.org.au> accessed 31 October 2014.

Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. – Global Ecol. Biogeogr. 19: 134–143.

Baselga, A. 2013. Multiple site dissimilarity quantifies compositional heterogeneity among several sites, while average pairwise dissimilarity may be misleading. – Ecography 36: 124–128.

Baselga, A. and Orme, C. D. L. 2012. betapart: an R package for the study of beta diversity. – Methods Ecol. Evol. 3: 808–812.

Belmaker, J. and Jetz, W. 2011. Cross-scale variation in species richness–environment associations. – Global Ecol. Biogeogr. 20: 464–474.

Burley, H. M. et al. 2016. Primary productivity is weakly related to floristic alpha and beta diversity across Australia. – Global Ecol. Biogeogr. 25: 1294–1307.

Caddy-Retalic, S. et al. 2017. Bioclimatic transect networks: powerful observatories of ecological change. – Ecol. Evol. 7: 4607–4619.

Cardoso, P. et al. 2009. Testing the performance of beta diversity measures based on incidence data: the robustness to undersampling. – Divers. Distrib. 15: 1081–1090.

Cardoso, P. et al. 2015. BAT – biodiversity assessment tools, an R package for the measurement and estimation of alpha and beta taxon, phylogenetic and functional diversity. – Methods Ecol. Evol. 6: 232–236.

Carlson, C. J. et al. 2016. An ecological assessment of the pandemic threat of Zika virus. – PLoS Negl. Trop. Dis. 10: e0004968.

Chytrý, M. 2001. Phytosociological data give biased estimates of species richness. – J. Veg. Sci. 12: 441–444.

Dahlgren, J. P. and Ehrlén, J. 2009. Linking environmental variation to population dynamics of a forest herb. – J. Ecol. 97: 666–674.

Dept of Environment, Water and Natural Resources 2012. Biological survey of South Australia – vegetation survey, biological database of South Australia, version 1/2014. – Obtained via the ÆKOS Data Portal (<www.aekos.org.au/>), made available by the State of South Australia (Dept of Environment, Water and Natural Resources), Adelaide, South Australia, accessed 7 July 2014.

DeWalt, R. E. et al. 2005. Just how imperiled are aquatic insects? A case study of stoneflies (Plecoptera) in Illinois. – Ann. Entomol. Soc. Am. 98: 941–950.

Emad, A. and Bailey, P. 2017. wCorr: weighted correlations. – R package ver. 1.9.1, <https://CRAN.R-project.org/package=wCorr>.

Engemann, K. et al. 2015. Limited sampling hampers "big data" estimation of species richness in a tropical biodiversity hotspot. – Ecol. Evol. 5: 807–820.

Engemann, K. et al. 2016. Patterns and drivers of plant functional group dominance across the Western Hemisphere: a macroecological re-assessment based on a massive botanical dataset. – Bot. J. Linn. Soc. 180: 141–160.

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.

Fitzpatrick, M. C. et al. 2013. Environmental and historical imprints on beta diversity: insights from variation in rates of species turnover along gradients. – Proc. R. Soc. B 280: 20131201.

Franklin, J. et al. 2017. Big data for forecasting the impacts of global change on plant communities. – Global Ecol. Biogeogr. 26: 6–17.

Gavin, D. G. 2009. The coastal-disjunct mesic flora in the inland Pacific Northwest of USA and Canada: refugia, dispersal and disequilibrium. – Divers. Distrib. 15: 972–982.

Gioia, P. and Hopper, S. D. 2017. A new phytogeographic map for the Southwest Australian Floristic Region after an exceptional decade of collection and discovery. – Bot. J. Linn. Soc. 184: 1–15.

González-Orozco, C. E. et al. 2014. Biogeographical regions and phytogeography of the eucalypts. – Divers. Distrib. 20: 46–58.

Gotelli, N. J. and Colwell, R. K. 2011. Estimating species richness. – In: Magurran, A. and McGill, B. (eds), Biological diversity: frontiers in measurement and assessment. Oxford Univ. Press, pp. 39–54.

Graham, C. H. et al. 2004. New developments in museum-based informatics and applications in biodiversity analysis. – Trends Ecol. Evol. 19: 497–503.

Grundy, M. J. et al. 2015. Soil and landscape grid of Australia. – Soil Res. 53: 835–844.

Guerin, G. R. et al. 2014. A spatially predictive baseline for monitoring multivariate species occurrences and phylogenetic shifts in mediterranean southern Australia. – J. Veg. Sci. 25: 338–348.

Guerin, G. R. et al. 2016. Identifying centres of plant biodiversity in South Australia. – PLoS One 11: e0144779.

Guerin, G. R. et al. 2017. Opportunities for integrated ecological analysis across inland Australia with standardised data from AusPlots Rangelands. – PLoS One 12: e0170137.

Hallgren, W. et al. 2016. The biodiversity and climate change virtual laboratory: where ecology meets big data. – Environ. Model. Softw. 76: 182–186.

Harris, J. A. et al. 2006. Ecological restoration and global climate change. – Restor. Ecol. 14: 170–176.

Harrison, S. et al. 1992. Beta diversity on geographic gradients in Britain. – J. Anim. Ecol. 61: 151–158.

Haque, M. M. 2015. A legacy of sampling: exploring spatial patterns among occurrence records in Australia's virtual herbarium. – Master of Research thesis, Macquarie Univ., Sydney, Australia.

Hijmans, R. J. 2015. raster: geographic data analysis and modeling. – R package ver. 2.4-18, <http://CRAN.R-project.org/package=raster>.
Hijmans, R. J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – Int. J. Climatol. 25: 1965–1978.
Hortal, J. et al. 2007. Limitations of biodiversity databases: case study on seed–plant diversity in Tenerife, Canary Islands. – Conserv. Biol. 21: 853–863.
Jácome, J. et al. 2007. A human-induced downward-skewed elevational abundance distribution of pteridophytes in the Bolivian Andes. – Global Ecol. Biogeogr. 16: 313–318.
Jost, L. 2007. Partitioning diversity into independent alpha and beta components. – Ecology 88: 2427–2439.
Jurasinski, G. and Retzer, V. 2012. simba: a collection of functions for similarity analysis of vegetation data. – R package ver. 0.3-5, <http://CRAN.R-project.org/package=simba>.
Kooyman, R. M. et al. 2013. Landscape patterns in rainforest phylogenetic signal: isolated islands of refugia or structured continental distributions? – PLoS One 8: e80685.
Legendre, P. and De Cáceres, M. 2013. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. – Ecol. Lett. 16: 951–963.
Leitão, P. J. et al. 2017. sgdm: an r package for performing sparse generalized dissimilarity modelling with tools for gdm. – ISPRS Int. J. Geo-Inform. 6: 23.
Maldonado, C. et al. 2015. Estimating species diversity and distribution in the era of Big Data: to what extent can we trust public databases? – Global Ecol. Biogeogr. 24: 973–984.
Manion, G. et al. 2016. gdm: functions for generalized dissimilarity modeling. – R package ver. 1.1.7, <http://CRAN.R-project.org/package=gdm>.
Michalcová, D. et al. 2011. Bias in vegetation databases? A comparison of stratified-random and preferential sampling. – J. Veg. Sci. 22: 281–291.
Mokany, K. et al. 2012. Dynamic macroecology and the future for biodiversity. – Global Change Biol. 18: 3149–3159.
Nualart, N. et al. 2017. Assessing the relevance of herbarium collections as tools for conservation biology. – Bot. Rev. 83: 303–325.
Oksanen, J. et al. 2016. vegan: community ecology package. – R package ver. 2.3-4, <http://CRAN.R-project.org/package=vegan>.
Prober, S. M. et al. 2012. Combining community-level spatial modelling and expert knowledge to inform climate adaptation in temperate grassy eucalypt woodlands and related grasslands. – Biodivers. Conserv. 21: 1627–1650.
Robertson, M. P. and Barker, N. P. 2006. A technique for evaluating species richness maps generated from collections data. – S. Afr. J. Sci. 102: 77–84.
Robertson, M. P. et al. 2016. Biogeo: an R package for assessing and improving data quality of occurrence record datasets. – Ecography 39: 394–401.
Robinson, A. C. et al. 1988. A biological survey of the Gawler Ranges, South Australia in October 1985. – National Parks and Wildlife Service, Dept of Environment and Planning, South Australia and South Australian Museum, Dept for the Arts, South Australia, Adelaide. D. J. Woolman, Government Printer.
Roleček, J. et al. 2007. Sampling design in large-scale vegetation studies: do not sacrifice ecological thinking to statistical purism! – Folia Geobot. 2: 199–208.
Rowe, R. J. 2005. Elevational gradient analyses and the use of historical museum specimens: a cautionary tale. – J. Biogeogr. 32: 1883–1897.
Thackway, R. and Cresswell, I. D. 1995. An interim biogeographic regionalization for Australia: a framework for establishing the national system of reserves, version 4.0. – Australian Nature Conservation Agency, Canberra.
Tobler, M. et al. 2007. Implications of collection patterns of botanical specimens on their usefulness for conservation planning: an example of two neotropical plant families (Moraceae and Myristicaceae) in Peru. – Biodivers. Conserv. 16: 659–677.
Vale, M. M. and Jenkins, C. N. 2012. Across-taxa incongruence in patterns of collecting bias. – J. Biogeogr. 39: 1744–1748.
Wernberg, T. et al. 2011. Seaweed communities in retreat from ocean warming. – Curr. Biol. 21: 1828–1832.
Zuur, A. F. et al. 2010. A protocol for data exploration to avoid common statistical problems. – Methods Ecol. Evol. 1: 3–14.