Using generalised dissimilarity modelling and targeted field surveys to gap-fill an ecosystem surveillance network

Greg R. Guerin^{1,2*}, Kristen J. Williams³, Emrys Leitch^{1,2}, Andrew J. Lowe¹, Ben Sparrow^{1,2}

¹School of Biological Science, The University of Adelaide, Adelaide, South Australia 5005,
Australia

²Terrestrial Ecosystem Research Network

³CSIRO Land and Water, Canberra, Australian Capital Territory 2601, Australia

*Email: greg.guerin@adelaide.edu.au

Abstract

1. When considering which sites or land parcels complement existing conservation or monitoring networks, there are many strategies for optimising ecological coverage in the absence of ground observations. However, such optimisation is often implemented theoretically in conservation prioritisation frameworks and real-world implementation is rarely assessed, particularly for networks of monitoring sites.
2. We assessed the performance of adding new survey sites informed by predictive modelling in gap-filling the ecological coverage of the Terrestrial Ecosystem Research Network's (TERN) continental network of ecosystem surveillance plots, Ausplots. Using plant cover observations from 531 sites, we constructed a generalised dissimilarity model (GDM) in which species composition was predicted by environmental parameters. We combined predicted nearest-neighbour ecological distances for locations across Australia with practical considerations to select regions for gap-filling surveys of 181 new plots across 18 trips. We tracked the drop in mean nearest-neighbour distances in GDM space, and increases in the actual sampling of ecological space through cumulative multivariate dispersion.
3. GDM explained 34% of deviance in species compositional turnover and retained geographic distance, soil P, aridity, actual evapotranspiration and rainfall seasonality among 17 significant predictors.
4. Key bioregions highlighted as gaps included Cape York Peninsula, Brigalow Belt South, South Eastern Queensland, Gascoyne and Dampierland.
5. We targeted identified gap regions for surveys in addition to opportunistic or project-based gap-filling over two years. Approximately 20% of the land area of Australia received increased servicing of biological representation, corresponding to a drop in mean nearest-neighbour ecological distances from 0.38 to 0.33 in units of compositional dissimilarity. The gain in sampled ecological space was 172% that from the previous 181 plots. Notable gaps

were filled in northern and south-east Queensland, north-east New South Wales and northern Western Australia.

6. Biological scaling of environmental variables through GDM supports practical sampling decisions for ecosystem monitoring networks. Optimising putative survey locations via ecological distance to a nearest neighbour rather than to all existing sites is useful when the aim is to increase representation of habitats rather than sampling evenness *per se*. Iterations between modelled gaps and field campaigns provide a pragmatic compromise between theoretical optima and real-world decision-making.

Introduction

Conservation planners have long grappled with the problem of how to maximise ecological representation in the absence of perfect information on biodiversity (Albuquerque & Beier 2018). For example, finite budgets mean the acquisition of incompletely surveyed land parcels to complement reserve systems must be prioritised. Indeed, much of the literature concerning spatial representation deals with conservation planning (Ferrier 2002; Williams et al. 2006; Arponen et al. 2008; Penniford et al. 2017), and, in particular, reserve selection (Faith & Walker 1996; Araújo et al. 2001; Engelbrecht et al. 2016).

Fewer studies have addressed optimising representation for targeted biological surveys (Funk et al. 2005; Hortal & Lobo 2005; Ware et al. 2018) or the design of monitoring networks (Stohlgren et al. 2011; Rose et al. 2016). Networks of ecosystem observation plots have traditionally been designed to maximise spatial or environmental coverage using stratification (Goring et al. 2016; Guerin et al. 2020). Even so, the problems of complementary reserve selection and gap-filling monitoring networks are, in essence, the same (Hopkins & Nunn 2010); both seek to maximise ecological coverage in the absence of fine-scale biodiversity survey data. For this reason, biodiversity surrogates are often employed to estimate and compare representativeness (Grantham et al. 2010).

While many easily measured indicators of biological pattern could serve as surrogates, spatial information is most commonly employed, due to its relative ease of access and high spatial coverage, and to enable interpolation between sparse data locations (Rodrigues & Brook 2007). For example, environmental layers relating to soil type, climate and land cover are frequently used as surrogates for ecological community composition, as are maps of classified vegetation types or bioregions (Ware et al. 2018). The rationale for environmental layers as surrogates is the assumption that species compositional turnover occurs as a result of environmental heterogeneity, due to differences in realised species

niches (Araújo et al. 2001). However, while environmental layers represent the broad environmental setting of ecological communities, they do not directly capture other key drivers of finer scale species compositional turnover, such as local disturbance regimes, microclimates, human impacts, and biotic interactions (Araújo et al. 2001; Peres-Neto et al. 2012; Goring et al. 2016). The spatial structure of ecological similarity (i.e., the degree to which species are shared between locations in space or time) is well known and is the reason that unstratified, systematic sampling performs reasonably at sampling ecological variation (Peres-Neto et al. 2012; Goring et al. 2016; Guerin et al. 2020). While local (patch level) drivers are opaque at large scales, ideally the purely spatial component of turnover would be included in assessments of surrogates.

In terms of selecting additional sites that best complement the ecological pattern of an existing set of sites, a range of optimisation strategies have been proposed that vary in terms of effectiveness and computational efficiency. Stohlgren et al. (2011) used Maxent modelling of the environmental coverage of monitoring site locations to select the most dissimilar putative sampling sites. Another approach is to select new sites that maximise the pairwise distances among all sample sites, so-called ‘Environmental Diversity’, which results in even sampling across environmental space (Faith & Walker 1996; Faith et al. 2004; Albuquerque & Beier 2018).

One realisation of environmental surrogate approaches is that ecological turnover is rarely a linear function of environmental turnover. The relationship is often nonstationary, with varying rates of turnover or even abrupt boundaries between compositionally distinct communities or geographic ecotones between structural vegetation types (Ferrier et al. 2007; Gibson et al. 2015; Goring et al. 2016). Generalised dissimilarity modelling (GDM; Ferrier et al 2007) has become the standard in community ecology for deriving biotically-scaled environmental surrogates (ecological environments) where sufficient biological training data

are available (Ashcroft et al. 2010; Rose et al. 2016), and is an appropriate surrogate measure for Faith's ED (Faith et al. 2004; Faith 2011). Because GDM explicitly accounts for nonlinearity, it more realistically scales environmental heterogeneity (and geographic distance) to actual species compositional turnover (Ferrier 2002; Ware et al. 2018). In doing so, GDM conveniently predicts ecological dissimilarity (and therefore the inverse measure, ecological similarity). The biotically-scaled environmental variables then characterise the pattern of ecological environments (Williams et al. 2014).

Here, we combined the predictive power of GDM with practical considerations to strategically gap-fill an established network of ecosystem surveillance plots – TERN Ausplots (Sparrow et al. 2019a). Ausplots were originally stratified to sample environmentally distinct bioregions across the Australian rangelands (Sparrow et al. 2019b; Guerin et al. 2020). However, due to an expansion in scope to all terrestrial bioregions and a practical limitation on the total number of sites that can be feasibly monitored, we aimed to select regions for survey that would efficiently increase ecological representation.

Because we aimed for representation and inclusion rather than even sampling across ecosystems (Schröder et al. 2006), we used a 'nearest-neighbour' approach (Faith & Norris 1989), in which each ecosystem is ideally 'serviced' by a representative site, with a preference to add new sites in habitats that are currently the least well represented (Hargrove & Hoffman 2004). We did not aim to make sampling even across all environments, not least because part of the original rationale behind Ausplots was that the Australian rangelands were more poorly monitored than mesic coastal areas (Guerin et al. 2017).

Here, we present how GDM was used to guide the placement of gap-filling surveys and assess the performance of strategic field sampling over time to fill identified gaps and increase the diversity (but not necessarily evenness) of the sampling in environmental space. While most of the literature on this topic selects putative, optimal sites (Arponen et al. 2008;

Hoffman et al. 2013; Rose et al. 2016; Albuquerque & Beier 2018), it is because we implemented the approach via field surveys that we use real new sites and observations of species turnover to assess performance. This assessment of biological survey gaps in the TERN Ausplots network was guided by the following questions:

Generalised dissimilarity model:

- What are the main macro-scale predictors of plant species turnover across Australia?
- Is extrapolation needed to predict turnover at national scale?

Site selection:

- Surveys of which putative regions would ecologically complement the existing Ausplots network, based on GDM predictions?
- How well did gap-filling surveys 'service' under-sampled ecosystems in practice?

Methods

Ausplots method and data

TERN Ausplots are a distributed network of one-hectare, fixed location monitoring plots with scope covering all major terrestrial ecosystems in Australia. Field methods (described at length elsewhere) focus on sampling vegetation and soils (White et al. 2012; Sparrow et al. 2019b). We calculated percent cover of plant species from 1010 point-intercepts per plot, with identifications of all species observed anywhere in the plot made through herbarium determination of vouchers. We accessed data for 763 Ausplots, including 582 plots that had been established prior to commencement of our gap-filling objectives (Guerin et al. 2018; TERN 2020). For convenience, we refer to the first 582 plots hereafter as ‘establishment plots’ and the subsequent 181 plots surveyed over 18 trips as ‘gap plots’.

Generalised dissimilarity modelling

We fit a GDM where the response variable was compositional turnover in plant species among establishment plots using species percentage cover as the abundance measure in calculating Bray-Curtis dissimilarity (Taft 2014; Manion et al. 2017; R Core Team 2016). We included repeat surveys (representing 7% of the samples) to maximise ecological information content, resulting in 572 surveys from 531 unique plots that had full vegetation data available at that time. Revisits capture observed variation in species composition within the same environmental setting and are handled numerically through spatial covariables, in that temporal replicates will have zero spatial distance.

Spatial layers with a resolution of 9 arcseconds were used as predictors, comprising 44 candidate variables related to climate, soil and landscape (Grundy et al. 2015; Harwood et al. 2016; Gallant et al. 2018; Table 1). Geographic distance was also included as a predictor. The number of predictors was reduced to a subset of 25 for which variance inflation factors were less than 10 by removing highly correlated variables that had weak importance in exploratory models. A statistical variable selection process for model fitting was applied to these 25 candidate variables.

To build the model, we used a backward selection procedure. Variables were removed when model coefficients summed to zero, which indicates no effect on species turnover. Variables were also removed if they were not statistically significant, using the method of Leitão et al. (2017), applying 1000 matrix permutations to test whether the contribution of each variable in the GDM was greater than expected at random.

Selection and appraisal of gap-filling surveys

Gaps in the ecological coverage of establishment plots were assessed spatially. Using predictor layers that had been GDM-transformed according to corresponding modelled *I*-splines (Manion 2009), we calculated the Euclidean distance of each grid cell to the nearest establishment plot in GDM space, and produced heat maps of 'ecological distance to nearest neighbour' (NND), which spatially maps representativeness (Hoffman et al. 2013). To evaluate reliability, we mapped a model extrapolation index, calculated as the sum of GDM-transformed values outside the input data range (i.e., the total amount of extrapolated turnover; Gibson et al. 2015).

Various procedures can be used to optimise the selection of putative sites with the aim of reducing mean NND, indicating better 'servicing' by the plot network (Faith & Norris 1989). For example, in the *p*-median approach, the mean drop in NND across a set of demand points is optimised (Engelbrecht et al. 2016). A simpler 'greedy' approach is to sequentially select sites with the highest NND, though this does not optimise the servicing of other sites *per se*. In the real world, however, the selection of sites for surveys combines quantitative information with a set of practical considerations. Field expeditions are planned around appropriate seasonal conditions for floristic sampling, logistical constraints (such as time in the field, availability of road infrastructure) and access to suitable sampling locations. Using modelled information on gaps from the GDM, we targeted regions rather than specific sites. During these field trips, an average of 10 (range 2–16) plots were established to cover local variation and provide some level of replication. In some instances, gap surveys were combined with local spatial gap-filling around establishment sites, which were opportunistically conducted during surveys focused on revisits, or were established in collaboration with other ecosystem monitoring or surveillance survey programs.

To improve confidence in decision-making, we supplemented the modelling and NND heat maps presented here to gaps highlighted using a pre-existing GDM of plant

species composition based on natural history records ('ANHAT' data; Williams et al. 2013; Ware et al. 2018). By applying the same set of establishment sites, *p*-median was used to map the highest priority areas for gap surveys (using the 'Survey Gap Analysis Tool'; Funk et al. 2005; Manion & Ridges 2009). Confidence in the mapping of gaps was raised where the two sources of modelled gap information (i.e., GDMs based on Ausplots or ANHAT data) were similar. We do not report this existing modelling in full here as it has been described elsewhere and formed only one aspect of the decision-making process. Notwithstanding, the major gaps identified by either model were similar.

The identification and surveying of gaps continued iteratively. Here, we evaluate two years of gap surveys in two ways. Firstly, we add real field plots clustered in regional 'trips' to the establishment plots in the chronological order they were surveyed and calculate the drop in NND across Australia. We refitted the GDM and repeated the appraisal of gaps incorporating the additional gap-filling survey data. The results were highly similar in terms of responses to gradients, and spatial mapping of predicted remaining gaps (qualitatively identical), hence are not reported here in further detail.

Secondly, we use the species composition percentage cover data recorded during the field surveys to calculate cumulative multivariate dispersion (MVD) in ecological space (Bray-Curtis dissimilarities in species composition), as individual plots were surveyed in chronological order. MVD is defined as mean distance of plots to their centroid in Principal Coordinates space (MDS of pairwise site distances; Oksanen et al. 2018), and is a measure of beta diversity (Anderson et al. 2006). If gap-filling surveys have been successful, we would expect NND to decrease over time, while we would expect MVD to increase. The difference between successive NND heat maps visualises gap-filling spatially.

Results

The GDM included geographic distance and 16 environmental variables as predictors, and explained 34% of deviance in pairwise dissimilarity (Table 1; Figs 1–2). The highest turnover was predicted in response to geographic distance, followed by soil P, aridity, actual evapotranspiration and rainfall seasonality. The GDM extrapolation index was below the suggested maximum of 0.1 (Gibson et al. 2015) for all but a small area in South Australia that was not highlighted as a gap (Fig. 1b).

Prior to gap-filling surveys, the NND map highlighted the Cape York Peninsula (CYP), Brigalow Belt South (BBS) and South Eastern Queensland (SEQ) bioregions as notable gaps (Thackway & Creswell 1995; Figs 3a, 4). Additional gaps included the Gascoyne (GAS) and Dampierlands (DAL) bioregions, as well as the north-west deserts. Field campaigns over two years in 2018, 2019 and 2020 targeted these and other successive gaps regions. Eighteen trips, surveying a total of 181 plots, were conducted (Table 2; Figs 3b, 4).

Considering all surveyed locations as of 2020 and NND, the major remaining gap was the north-west desert region, and to a lesser extent the far south-east, the coast sandplains of south-west Western Australia and the southern Queensland coastline (Figs 3b; 4). The difference between before and after NND maps spatially highlights the gaps that were filled by the surveys (Fig. 3c). Many of the major gaps identified initially via NND were filled to some extent. The magnitude of the improvement in NND for those gap areas was high, with NND drops of up to ~1, equivalent to a shift from no coverage to complete servicing of species composition based on the GDM.

Mean NND across Australia dropped from 0.38 to 0.33, in units of compositional dissimilarity (Fig. 5a). The proportion of Australia that was serviced through gap-filling plots via decreased NND was 18% at a minimum of 0.05 drop, 14% at a minimum of 0.1 drop, 6% at a minimum drop of 0.25, and 1% at a minimum 0.5 drop (Fig. 5b).

The rate at which mean MVD increased with additional surveys (representing accumulation of ecological space in the network) rose sharply during the early establishment phase of surveys, levelled off from approximately 200 plots, but was then higher during the gap-filling phase, with a gain of 172% that of the previous 181 plots, and an 8% increase in overall accumulation over the entire establishment phase (Fig. 5c).

Discussion

Biodiversity surrogates provide a powerful approach to gap-filling when applied to networks of both conservation reserves and surveillance monitoring sites (Rodrigues & Brook 2007). The utility of environmental layers as surrogates increases when these are biotically scaled by modelling their correlation with species turnover (Ware et al. 2018). We chose generalised dissimilarity modelling (GDM) for that purpose because it accounts appropriately for nonlinear responses, including varying rates of turnover at different positions along environmental gradients (Ferrier 2002), and has been extensively applied and proven as suitable for biodiversity assessment with hundreds of citations (Ferrier et al. 2007).

We were able to use the comprehensive species composition and relative abundance data recorded in the large set of plots already established across Australia to inform further gap-filling surveys through compositional models. This approach contrasts with the original environmental stratification used to establish the TERN Ausplots network, which used environmental layers in the absence of ground data at that stage (Guerin et al. 2020).

Although species occurrence data held in natural history databases provide greater spatial and taxonomic coverage, the advantage of TERN Ausplots species composition data for modelling ecological distance is its ability to capture locally complete sets of co-occurring species at patch scales, and is the only standardised dataset of species identity and relative abundance within one-hectare plots across all State and Territory jurisdictions in Australia.

Alternative datasets for this modelling are either presence-only (requiring assumptions about aggregation to represent assemblages) or require harmonisation of plot-based surveys with eclectic methods (e.g., plot sizes, affecting species richness) and different accuracies for scoring relative abundance, necessitating development of covariates to characterise and explain systematic error.

By mapping predicted ecological distance to the most similar site in the establishment network of permanent plots, we were able to identify major sampling gaps. These gaps were addressed via survey expeditions to the Australian Alps (AUA), Cape York Peninsula (CYP), Brigalow Belt South (BBS), Yalgoo (YAL), Dampierland (DAL), Tasmanian Northern Midlands (TNM) and Tasmanian Central Highlands (TCH) bioregions (Table 2; Fig. 4; Thackway & Creswell 1995). Our retrospective assessment confirms that the combination of gap modelling and practical considerations with more traditional local stratification of survey plots was able to achieve the goal of efficiently representing terrestrial Australian ecosystems.

Predictors of plant species turnover across Australia

In the GDM of species composition data from Ausplots, the highest species turnover was fitted to geographic distance, soil P, aridity (monthly minimum and maximum aridity index, potential evaporation and actual evapotranspiration) and rainfall seasonality. However, our aim was to model ecological turnover rather than determine the most important drivers, and it is likely that covariance among competing predictors is involved in some of the observed patterns. Indeed, variance partitioning (Gilbert & Bennet 2010) shows total deviance explained comprised 3.8% independent spatial effect, 34.6% independent environmental effect and 61.7% spatial–environmental covariance, while the partitioning for environmental

variables was 31.9% independent climate effect, 21.3% independent soil effect and 46.9% climate–soil covariance.

The deviance explained by the model at 34% suggests that a number of additional factors not accounted for in the model influence observed species composition, as expected. Examples of such factors include fire and grazing regimes, human disturbance, proximity to regional climate refugia, seasonal phenology and climatic condition, and biotic interactions (Araújo et al. 2001; Gibson et al. 2015; Keppel et al. 2017), and these are likely to weaken relationships between macroclimate and biodiversity, particularly at local scales (Bruehlheide et al. 2018; Harrison et al. 2020).

The other published national-scale model of turnover in plant species composition for Australia is the GDM based on occurrence data extracted from the Australian Natural Heritage Assessment Tool (‘ANHAT’; Williams et al. 2013; Ware et al. 2018). The ANHAT model was also fitted with a backwards selection procedure that finished with 11 climate and six substrate predictors. The ANHAT and Ausplots GDMs are similar, with rainfall seasonality, evaporation and actual evapotranspiration as the most important predictors. The importance of aridity and rainfall seasonality are intuitive as key variables at large scales that explain differences between tropical, desert and temperate biomes (as visualised in Fig. 1a). However, it is likely that the importance of particular variables, and the nature of the ecological response to them, varies from region to region (Burley et al. 2012; Guerin et al. 2019).

Edaphic predictors are harder to compare between the two models because the ANHAT model used an earlier version of the soil parameters before the full TERN data release (Viscarra-Rossel & Chen 2011; Viscarra Rossel et al. 2014a; Grundy et al. 2015; Viscarra Rossel et al. 2015), whereas our Ausplots model used a more comprehensive variable suite aggregated to 9-arcsecond grids from the 3-arcsecond source data (Viscarra

Rossel et al. 2014a; Gallant et al. 2018). The ANHAT GDM incorporates a larger dataset of presence-only records aggregated within adequately sampled grid cells (as described in Williams et al. 2010), whereas the Ausplots GDM was based on comprehensive floristic surveys of 531 plots, with full inventories of co-occurring species at patch level along with robust measures of percent cover.

Complementing the existing TERN Ausplots network with strategic gap-filling surveys

Strategic gap-filling surveys guided by the Ausplots GDM and practical considerations were successful in complementing sampling of the Australian vegetation over two years. The mean nearest-neighbour distance (NND) across the whole of Australia dropped from 0.38 to 0.33 in units of compositional dissimilarity. Almost one fifth of the continent was thereby additionally serviced by gap-filling to some degree, while 6% of the land area of Australia had a drop in NND of at least 0.25 in units of compositional dissimilarity. The size of the additional ecological space sampled was 172% that of the previous 181 plots, adding 8% over all establishment phase sampling, as measured by multivariate dispersion, which is significant given the law of diminishing returns and that coverage of the Australian terrestrial environment by establishment plots was already quite comprehensive (Fig. 5c; Guerin et al. 2020).

Although we generated sets of putative survey sites optimised in one way or another during the course of the gap-filling process, we found it more enlightening to evaluate how gap-filling performed in practice. While optimisation methods are useful, it is not always feasible to access pre-defined locations. Logistical considerations also influence survey locations. The *modus operandi* of the TERN ecosystem surveillance field team is to target a region with a set of surveys over an expedition for cost-efficiency reasons, to capture local heterogeneity and to ensure a level of replication.

Planning of field surveys is iterative and often involves a series of logistical contingencies. To achieve all of this, an optimal approach would need to iteratively trade-off survey ‘benefit’ due to greater biological coverage against survey ‘cost’ (e.g. Rodewald et al. 2019). In covering the diverse ecosystems of Australia, field work needed to be flexible, practical and pragmatic to maximise useful data collection given huge variation in seasonal climatic conditions and plant phenology responses. In addition, opportunities for sampling due to collaborative projects and the time and cost of reaching distant and remote sampling locations, were among many factors that must be considered. Sampling of the diverse Australian environment has been remarkably efficient for a relatively small number of plots, due to strategic selection of monitoring sites.

The never-ending story

Strategic field surveying of identified gaps is an iterative process leading to gap-filling at finer levels of detail over time. A practical end-point to this process is imposed by resource constraints that limit the number of field plots feasible to establish and regularly revisit, considered against the desired level of gap-filling and trade-offs between spatial and temporal sampling coverage. The goal of TERN Ausplots has been to establish a representative network of ecosystem surveillance plots to remeasure at least once per decade. Gap-filling will continue to some degree into the future, with increasing emphasis on revisits.

In terms of the GDM developed here, a feasible end-point could be to continue gap-filling until all NND are below one (see Fig. 3), meaning each grid cell is predicted to share some species with the plot network, based on the most up-to-date data available (Ferrier 2002). This end-point has almost been reached already, with only five 100,000ths of the land area over that cut-off. A more ambitious end-point would be a maximum NND of 0.5. This end-point requires increased servicing of an additional 12% of the Australian land area (a

reduction from 21% prior to gap-filling; Fig. 3). Current indications are that a network with an additional 250–750 plots will be adequate to cover the major Australian terrestrial environments with sufficient replication of major vegetation types.

The largest remaining gap identified in the TERN Ausplots network from the GDM is a large region in the north-west deserts (Fig. 2b). This area was already an obvious spatial gap in the distribution of field plots. However, staging an expedition to this region with the expectation of performing revisits requires careful, long-term planning, as it is a remote and challenging environment far from population centres and with extremely limited road infrastructure. In addition, long-term access to sites for monitoring would need to be secured through partnerships with land managers and traditional owners which we aim to progress in the coming years.

The never-ending pursuit of ecological sampling gaps in the ecosystem surveillance network also continually shifts priority sites for regular revisits. For example, the subset of Ausplots that is most representative of the network, and of Australian ecosystems collectively, will evolve as gap-filling surveys proceed. This iterative approach to representation is one consideration in sequencing repeat visits. The existence of climate change and other drivers of compositional dynamics is another consideration. Gaps in representativeness may therefore change at different rates across the region. For this reason, sensitivity and exposure to climate change are now also being considered empirically to ensure ecosystems potentially undergoing rapid change are more regularly resampled.

Limitations

While we implemented a quantitative yet practical solution to ecological gap-filling surveys, a range of other factors will also determine the efficacy of the network in observing spatial and temporal patterns in Australian terrestrial ecosystems. The frequency of revisits,

precision of measurements relative to rates of change and degree of replication are factors that contribute to the power to detect compositional change in species and infer drivers of observed patterns (Guillera-Aroita & Lahoz-Monfort 2012). Remotely derived environmental layers, even when biotically scaled, are imperfect surrogates for biodiversity turnover (Araújo et al. 2001) for reasons already discussed. Other aspects of ecological stratification and gap-filling could include sampling along gradients of disturbance and human influence, for example.

Conclusions

Gap-filling surveys for ecosystem monitoring or conservation networks are guided by quantitative assessments that help to prioritise putative target sites within practical constraints. Ecological coverage of TERN Ausplots was significantly increased by targeted surveys guided by biotically scaled environmental surrogates for biodiversity. Optimising putative survey locations by modelling distance to a nearest neighbour in GDM space was an efficient analytical method that enabled regular updates in near real time. Filling ever-finer gaps becomes a trade-off with temporal coverage as well as addressing different kinds of gaps, such as relative climate change vulnerability and consideration of logistical constraints. Modelling combined with field campaigns in near real time offers a pragmatic compromise for real-world decision-making.

Acknowledgements

We thank all members past and present of the TERN Ecosystem Surveillance field and data teams and TERN, supported by the Australian Government through the National Collaborative Research Infrastructure Strategy.

References

- Albuquerque, F., & Beier, P. (2018) Improving the use of environmental diversity as a surrogate for species representation. *Ecology and Evolution*, 8, 852–858.
- Anderson, M.J., Ellingsen, K.E. & McArdle, B.H. (2006) Multivariate dispersion as a measure of beta diversity. *Ecology Letters* 9:683–693.
- Araújo, M.B., Humphries, C.J., Densham, P.J., Lampinen, R., Hagemer, W.J.M., Mitchell-Jones, A.J. & Gasc, J.P. (2001) Would environmental diversity be a good surrogate for species diversity? *Ecography*, 24, 103–110.
- Arponen, A., Moilanen, A., & Ferrier, S. (2008) A successful community-level strategy for conservation prioritization. *Journal of Applied Ecology*, 45, 1436–1445.
- Ashcroft, M.B., Gollan, J.R., Faith, D.P., Carter, G.A., Lassau, S.A., Ginn, S.G., Bulbert, M.W. & Cassis, G. (2010) Using generalised dissimilarity models and many small samples to improve the efficiency of regional and landscape scale invertebrate sampling. *Ecological Informatics*, 5, 124–132.
- Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S.M., Botta-Dukát, Z., Chytrý, M., Field, R., Jansen, F., Kattge, J., Pillar, V.D., Schrod, F., Mahecha, M.D., Peet, R.K., Sandel, B., van Bodegom, P., Altman, J., Alvarez Davila, E., Arfin Khan, M.A.S., Attorre, F., Aubin, I., Baraloto, C., Barroso, J.G., Bauters, M., Bergmeier, E., Biurrun, I., Bjorkman, A.D., Blonder, B., Čarni, A., Cayuela, L., Černý, T., Cornelissen, J.H.C., Craven, D., Dainese, M., Derroire, G., De Sanctis, M., Díaz, S., Doležal, J., Farfan-Rios, W., Feldpausch, T.R., Fenton, N.J., Garnier, E., Guerin, G.R., Gutiérrez, A.G., Haider, S., Hattab, T., Henry, G., Hérault, B., Higuchi, P., Hölzel, N., Homeier, J., Jentsch, A., Jürgens, N., Kački, Z., Karger, D.N., Kessler, M., Kleyer, M., Knollová, I., Korolyuk, A.Y., Kühn, I., Laughlin, D.C., Lens, F., Loos, J., Louault, F., Lyubenova, M.I., Malhi, Y., Marcenò, C., Mencuccini, M.,

- Müller, J.V., Munzinger, J., Myers-Smith, I.H., Neill, D.A., Niinemets, Ü., Orwin, K.H., Ozinga, W.A., Penuelas, J., Pérez-Haase, A., Petřík, P., Phillips, O.L., Pärtel, M., Reich, P.B., Römermann, C., Rodrigues, A.V., Sabatini, F.M., Sardans, J., Schmidt, M., Seidler, G., Silva Espejo, J.E., Silveira, M., Smyth, A., Sporbert, M., Svenning, J., Tang, Z., Thomas, R., Tsiripidis, I., Vassilev, K., Violle, C., Virtanen, R., Weiher, E., Welk, E., Wesche, K., Winter, M., Wirth, C., Jandt, U. (2018) Global trait–environment relationships of plant communities. *Nature Ecology and Evolution*, 2, 1906–1917.
- Burley, H.M., Laffan, S.W. & Williams, K.J. (2012) Spatial non-stationarity and anisotropy of compositional turnover in eastern Australian Myrtaceae species. *International Journal of Geographical Information Science*, 26, 2065-2081.
- Department of the Environment (2012) Interim Biogeographic Regionalisation for Australia (Subregions) v. 7 (IBRA) [ESRI shapefile] Available from <http://intspat01.ris.environment.gov.au/fed/catalog/search/resource/details.page?uuid=%7BDFB470D2-FA7E-4792-B524-0250EC3AF5B3%7D>
- Engelbrecht, I., Robertson, M., Stoltz, M., & Joubert, J.W. (2016) Reconsidering environmental diversity (ED) as a biodiversity surrogacy strategy. *Biological Conservation*, 197, 171–179.
- Faith, D.P. & Norris, R.H. (1989). Correlation of environmental variables with patterns of distribution and abundance of common and rare freshwater macroinvertebrates. *Biological Conservation*, 50, 77–98.
- Faith, D.P. & Walker, P.A. (1996) Environmental diversity: on the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. *Biodiversity & Conservation*, 5, 399–415.

- 469 Faith, D.P., Ferrier, S. & Walker, P.A. (2004) The ED strategy: how species-level surrogates
- 470 indicate general biodiversity patterns through an 'environmental diversity' perspective.
- 471 *Journal of Biogeography*, 31, 1207–1217.
- 472 Faith, D.P. (2011) Attempted tests of the surrogacy value of the ED environmental diversity
- 473 measures highlight the need for corroboration assessment of surrogacy hypotheses.
- 474 *Ecological Indicators*, 11, 745–748.
- 475 Ferrier, S. (2002) Mapping spatial pattern in biodiversity for regional conservation planning:
- 476 where to from here? *Systematic Biology*, 51, 331-363.
- 477 Ferrier, S., Manion, G., Elith, J. & Richardson, K. (2007) Using generalized dissimilarity
- 478 modelling to analyse and predict patterns of beta diversity in regional biodiversity
- 479 assessment. *Diversity & Distributions*, 13, 252–264.
- 480 Funk, V.A., Richardson, K.S. & Ferrier, S. (2005) Survey-gap analysis in expeditionary
- 481 research: where do we go from here? *Biological Journal of the Linnean Society*, 85,
- 482 549–567.
- 483 Gallant, J. & Austin, J. (2012) Contributing Area - Multiple Flow Direction (Partial) derived
- 484 from 1" SRTM DEM-H. v2. CSIRO. Data Collection.
- 485 <https://doi.org/10.4225/08/57511C42603DF>
- 486 Gallant, J., Austin, J., Williams, K., Harwood, T., King, D., Nolan, M. & Mokany, K. (2018)
- 487 9s soil and landform for continental Australia analysis of biodiversity pattern:
- 488 aggregated from 3s data. V1. In: CSIRO (ed.) Data Collection. CSIRO,
- 489 <https://doi.org/10.4225/08/5b285fd14991f>.
- 490 Gibson, L.A., Williams, K.J., Pinder, A.M., Harwood, T.D., McKenzie, N.L., Ferrier, S. &
- 491 Manion, G. (2015) Compositional patterns in terrestrial fauna and wetland flora and
- 492 fauna across the Pilbara biogeographic region of Western Australia and the

representativeness of its conservation reserve system. *Records of the Western Australian Museum, Supplement*, 78, 515–545.

Gilbert, B. & Bennett, J.R. (2010) Partitioning variation in ecological communities: do the numbers add up? *Journal of Applied Ecology*, 47, 1071–1082.

Goring, S.J., Mladenoff, D.J., Cogbill, C.V., Record, S., Paciorek, C.J., Jackson, S.T., Dietze, M.C., Dawson, A., Matthes, J.H., McLachlan, J.S. & Williams, J.W. (2016) Novel and lost forests in the upper Midwestern United States, from new estimates of settlement-era composition, stem density, and biomass. *PloS ONE*, 11, e0151935.

Grantham, H.S., Pressey, R.L., Wells, J.A. & Beattie, A.J. (2010) Effectiveness of biodiversity surrogates for conservation planning: different measures of effectiveness generate a kaleidoscope of variation. *PLoS ONE*, 5, e11430.

Grundy, M.J., Rossel, R.A.V., Searle, R.D., Wilson, P.L., Chen, C. & Gregory, L.J. (2015) Soil and Landscape Grid of Australia. *Soil Research*, 53, 835–844.

Guerin, G.R., Sparrow, B., Tokmakoff, A., Smyth, A., Leitch, E., Baruch, Z., Lowe, A.J. (2017) Opportunities for integrated ecological analysis across inland Australia with standardised data from AusPlots Rangelands. *PloS ONE*, 12, e0170137.

Guerin, G.R., Saleeba, T. & Tokmakoff, A. (2018) ausplotsR: TERN AusPlots analysis package. R package version 1.0. <https://github.com/GregGuerin/ausplotsR>

Guerin, G.R., Andersen, A.N., Rossetto, M., van Leeuwen, S., Byrne, M., Sparrow, B., Rodrigo, M. & Lowe, A.J. (2019) Consistent sorting but contrasting transition zones in plant communities along bioclimatic gradients. *Acta Oecologica* 95, 74–85.

Guerin, G.R., Williams, K.J., Sparrow, B. & Lowe, A.J. (2020) Stocktaking the environmental coverage of a continental ecosystem observation network. *BioRxiv* DOI: 10.1101/2020.02.20.958769

- Guillera-Aroita, G. & Lahoz-Monfort, J.J. (2012) Designing studies to detect differences in species occupancy: power analysis under imperfect detection. *Methods in Ecology and Evolution*, 3, 860–869.
- Hargrove, W. W., & Hoffman, F. M. (2004). Potential of multivariate quantitative methods for delineation and visualization of ecoregions. *Environmental management*, 34, S39–S60.
- Harrison, S., Spasojevic, M.J. & Li, D. (2020) Climate and plant community diversity in space and time. *PNAS*, 117, 4464–4470.
- Harwood, T., Donohue, R., Harman, I., McVicar, T., Ota, N., Perry, J., and Williams, K. (2016) 9s climatology for continental Australia 1976-2005: Summary variables with elevation and radiative adjustment. V1. In: CSIRO (ed.) Data Collection. CSIRO, <https://doi.org/10.4225/08/5afa9f7d1a552>.
- Hoffman, F.M., Kumar, J., Mills, R.T. & Hargrove, W.W. (2013) Representativeness-based sampling network design for the State of Alaska. *Landscape Ecology*, 28, 1567–1586.
- Hopkins, M.E. & Nunn, C.L. (2010) Gap analysis and the geographical distribution of parasites. *The Biogeography of Host–Parasite Interactions*. Oxford University Press: Oxford, pp.129-42.
- Hortal, J. & Lobo, J.M. (2005) An ED-based protocol for optimal sampling of biodiversity. *Biodiversity & Conservation*, 14, 2913–2947.
- Keppel, G., Robinson, T.P., Wardell-Johnson, G.W., Yates, C.J., Van Niel, K.P., Byrne, M. & Schut, A.G. (2017) A low-altitude mountain range as an important refugium for two narrow endemics in the Southwest Australian Floristic Region biodiversity hotspot. *Annals of Botany*, 119, 289–300.

- Leitão, P.J., Schwieder, M. & Senf, C. (2017) Sgdm: an R package for performing sparse generalized dissimilarity modelling with tools for gdm. *ISPRS International Journal of Geo-Information*, 6, 23. Doi:10.3390/ijgi6010023
- Manion, G. (2009) A technique for constructing monotonic regression splines to enable non-linear transformation of GIS rasters. *In*: Anderssen RS, Braddock RD and Newham LTH (eds) 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation - Interfacing Modelling and Simulation with Mathematical and Computational Sciences, 13–17 July 2009, Cairns, Australia. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, Cairns, Australia, 2507-2513.
- Manion, G. & Ridges, M. (2009) An optimisation of the survey gap analysis technique to minimise computational complexity and memory resources in order to accommodate fine grain environmental and site data. *In*: Anderssen RS, R.D. Braddock and L.T.H. Newham (ed.) 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, 2514-2520.
- Manion, G., Lisk, M., Ferrier, S., Nieto-Lugilde, D., Mokany, K. & Fitzpatrick, M.C. (2017) gdm: Generalized Dissimilarity Modeling. R package version 1.3.0. <https://CRAN.R-project.org/package=gdm>
- Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E. & Wagner, H. (2018) vegan: Community Ecology Package. R package version 2.5-3. <https://CRAN.R-project.org/package=vegan>

564 Penniford, M. G., Williams, K. J., Pinder, A. M., Harwood, T. D., Manion, G., & Ferrier, S.
565 (2017) Whole-landscape modelling of compositional turnover in aquatic invertebrates
566 informs conservation gap analysis: An example from south-western
567 Australia. *Freshwater Biology*, 62, 1359–1376.

568 Peres-Neto, P.R., Leibold, M.A., Dray, S. (2012) Assessing the effects of spatial contingency
569 and environmental filtering on metacommunity phylogenetics. *Ecology*, 93, S14–S30.

570 R Core Team (2016) R: A language and environment for statistical computing. R Foundation
571 for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

572 Rodewald, A.D., Strimas-Mackey, M., Schuster, R. & Arcese, P. (2019) Tradeoffs in the
573 value of biodiversity feature and cost data in conservation prioritization. *Scientific*
574 *Reports*, 9, 15921.

575 Rodrigues, A. S., & Brooks, T. M. (2007) Shortcuts for biodiversity conservation planning:
576 the effectiveness of surrogates. *Annual Review of Ecology, Evolution &*
577 *Systematics*, 38, 713–737.

578 Rose, P. M., Kennard, M. J., Sheldon, F., Moffatt, D. B., & Butler, G. L. (2016) A data-
579 driven method for selecting candidate reference sites for stream bioassessment
580 programs using generalised dissimilarity models. *Marine and Freshwater Research*,
581 67, 440–454.

582 Schröder, W., Pesch, R., & Schmidt, G. (2006) Identifying and closing gaps in environmental
583 monitoring by means of metadata, ecological regionalization and geostatistics using
584 the UNESCO biosphere reserve Rhoein (Germany) as an example. *Environmental*
585 *Monitoring and Assessment*, 114, 461–488.

586 Sparrow, B.D, Edwards, W., Munroe, S.E.M., Wardle, G.M., Guerin, G.R., Bastin, J-F,
587 Morris, B., Christensen, R., Phinn, S. and Lowe, A.J. (2019a) Effective ecological
588 monitoring requires a multi-scaled approach. *EcoEvoRxiv*. doi:10.32942/osf.io/3q5v2

- Sparrow, B., Foulkes, J., Wardle, G., Leitch, E., Caddy-Retalic, S., van Leeuwen, S., Tokmakoff, A., Thurgate, N., Guerin, G. and Lowe, A. (2019b) A Vegetation and soil survey method for surveillance monitoring of rangeland environments. *Preprints* 2019110178. Doi: 10.20944/preprints201911.0178.v1.
- Stohlgren, T. J., Kumar, S., Barnett, D. T., & Evangelista, P. H. (2011) Using maximum entropy modeling for optimal selection of sampling sites for monitoring networks. *Diversity*, 3, 252–261.
- Taft, B.N. (2014) dissUtils: Utilities for making pairwise comparisons of multivariate data. R package version 1.0. <https://CRAN.R-project.org/package=dissUtils>
- TERN (2020) AusPlots ecosystem surveillance monitoring dataset (URL: <http://aekos.org.au/collection/adelaide.edu.au/ausplotsrangelands>). Obtained via the ausplotsR package (URL: <https://github.com/ternaustralia/ausplotsR>), accessed 27 March 2020.
- Thackway, R., & Creswell, I. (1995) An interim biogeographic regionalisation for Australia: a framework for establishing the national system of reserves, Version 4.0. Australian Nature Conservation Agency, Canberra.
- Viscarra-Rossel, R.A. & Chen, C. (2011) Digitally mapping the information content of visible-near infrared spectra of surficial Australian soils. *Remote Sensing of Environment*, 115, 1443–1455.
- Viscarra Rossel, R., Chen, C., Grundy, M., Searle, R. & Clifford, D. (2014a) Soil and Landscape Grid Australia-Wide 3D Soil Property Maps (3" resolution) - Release 1. v3. CSIRO. Data Collection. <https://doi.org/10.4225/08/5aaf553b63215>
- Viscarra Rossel, R., Chen, C., Grundy, M., Searle, R., Clifford, D., Odgers, N., Holmes, K., Griffin, T., Liddicoat, C. & Kidd, D. (2014b): Soil and Landscape Grid National Soil

613 Attribute Maps - Total Phosphorus (3" resolution) - Release 1. v5. CSIRO. Data
614 Collection. <https://doi.org/10.4225/08/546F617719CAF>

615 Viscarra Rossel, R., Chen, C., Grundy, M., Searle, R., Clifford, D., Odgers, N., Holmes, K.,
616 Griffin, T., Liddicoat, C. & Kidd, D. (2014c): Soil and Landscape Grid National Soil
617 Attribute Maps - Total Nitrogen (3" resolution) - Release 1. v5. CSIRO. Data
618 Collection. <https://doi.org/10.4225/08/546F564AE11F9>

619 Viscarra Rossel, R., Chen, C., Grundy, M., Searle, R., Clifford, D., Odgers, N., Holmes, K.,
620 Griffin, T., Liddicoat, C. & Kidd, D. (2014d): Soil and Landscape Grid National Soil
621 Attribute Maps - pH - CaCl₂ (3" resolution) - Release 1. v3. CSIRO. Data Collection.
622 <https://doi.org/10.4225/08/546F17EC6AB6E>

623 Viscarra Rossel, R., Chen, C., Grundy, M., Searle, R., Clifford, D., Odgers, N., Holmes, K.,
624 Griffin, T., Liddicoat, C. & Kidd, D. (2014e): Soil and Landscape Grid National Soil
625 Attribute Maps - Bulk Density - Whole Earth (3" resolution) - Release 1. v5. CSIRO.
626 Data Collection. <https://doi.org/10.4225/08/546EE212B0048>

627 Viscarra Rossel, R., Chen, C., Grundy, M., Searle, R., Clifford, D., Odgers, N., Holmes, K.,
628 Griffin, T., Liddicoat, C. & Kidd, D. (2014f): Soil and Landscape Grid National Soil
629 Attribute Maps - Clay (3" resolution) - Release 1. v5. CSIRO. Data Collection.
630 <https://doi.org/10.4225/08/546EEE35164BF>

631 Viscarra Rossel, R., Chen, C., Grundy, M., Searle, R., Clifford, D., Odgers, N., Holmes, K.,
632 Griffin, T., Liddicoat, C. & Kidd, D. (2014g): Soil and Landscape Grid National Soil
633 Attribute Maps - Available Water Capacity (3" resolution) - Release 1. v4. CSIRO.
634 Data Collection. <https://doi.org/10.4225/08/546ED604ADD8A>

635 Viscarra Rossel, R., Chen, C., Grundy, M., Searle, R., Clifford, D., Odgers, N., Holmes, K.,
636 Griffin, T., Liddicoat, C. & Kidd, D. (2014h): Soil and Landscape Grid National Soil

637 Attribute Maps - Organic Carbon (3" resolution) - Release 1. v2. CSIRO. Data
638 Collection. <https://doi.org/10.4225/08/547523BB0801A>

639 Viscarra Rossel, R., Chen, C., Grundy, M., Searle, R., Clifford, D., Odgers, N., Holmes, K.,
640 Griffin, T., Liddicoat, C. & Kidd, D. (2014i): Soil and Landscape Grid National Soil
641 Attribute Maps - Effective Cation Exchange Capacity (3" resolution) - Release 1. v4.
642 CSIRO. Data Collection. <https://doi.org/10.4225/08/546F091C11777>

643 Viscarra Rossel, R.A., Chen, C., Grundy, M.J., Searle, R., Clifford, D. & Campbell, P.H.
644 (2015) The Australian three-dimensional soil grid: Australia's contribution to the
645 GlobalSoilMap project. *Soil Research*, 53, 845–864.

646 Ware, C., Williams, K.J., Harding, J., Hawkins, B., Harwood, T., Manion, G., Perkins, G.C.
647 & Ferrier, S. (2018) Improving biodiversity surrogates for conservation assessment: A
648 test of methods and the value of targeted biological surveys. *Diversity and*
649 *Distributions*, 24, 1333–1346.

650 White, A., Sparrow, B., Leitch, E., Foulkes, J., Flitton, R., Lowe, A.J. & Caddy-Retalic, S.
651 (2012) AusPlots Rangelands survey protocols manual. Version 1.2.9. The University
652 of Adelaide Press, Adelaide, South Australia. ISBN 978-1-922064-38-7.

653 Williams, P., Faith, D., Manne, L., Sechrest, W. & Preston, C. (2006) Complementarity
654 analysis: Mapping the performance of surrogates for biodiversity. *Biological*
655 *Conservation*, 128, 253–264.

656 Williams, K.J., Belbin, L., Austin, M.P., Stein, J. & Ferrier, S. (2012) Which environmental
657 variables should I use in my biodiversity model? *International Journal of Geographic*
658 *Information Sciences*, 26, 2009–2047.

659 Williams, K., Harwood, T., Manion, G., Ferrier, S., Perry, J., Rosauer, D. & Laffan, S. (2013)
660 VAS_v5_r11: Generalised dissimilarity model of compositional turnover in vascular
661 plant species for continental Australia at 9 second resolution using ANHAT data

extracted April 2013. V1. CSIRO. Data Collection.

<https://doi.org/10.4225/08/557FB520465F7>

Williams, K.J., Prober, S.M., Harwood, T.D., Doerr, V.A.J., Jeanneret, T., Manion, G. & Ferrier, S. (2014) Implications of climate change for biodiversity: a community-level modelling approach (a guide for use with the datasets and maps). CSIRO Land and Water Flagship, Canberra, www.AdaptNRM.org.

Williams, K., Ferrier, S., Rosauer, D., Yeates, D., Manion, G., Harwood, T., Stein, J., Faith, D., Laity, T. & Whalen, A. (2020) Harnessing Continent-Wide Biodiversity Datasets for Prioritising National Conservation Investment. Canberra: CSIRO Ecosystem Sciences. <https://doi.org/10.4225/08/584c42ba9662a>

Table 1. Climate, soil and landform variables used in a generalised dissimilarity model (GDM) of plant species composition based on TERN Ausplots 'establishment' sites (source: Harwood et al. 2016; Gallant et al. 2018). Variables ranked by importance according to summed model coefficients (height of the *I*-spline on the y-axis as a measure of relative importance).

Code	Name and origin data citation	Unit	Summed GDM coefficients
GEO	Geographic distance	degrees	2.72
PTO	Total Phosphorus (Viscarra Rossel et al. 2014b)	%	2.22
EAA	Annual total actual evapotranspiration (Harwood et al. 2016)	mm	2.10
ADI	Minimum monthly aridity index (Williams et al. 2012)	proportion	1.74
ADX	Maximum monthly aridity index (Williams et al. 2012)	proportion	1.45
PTS2	Precipitation seasonality 2 (Williams et al. 2012)	ratio	1.36
TRX	Maximum monthly mean diurnal temperature range (Harwood et al. 2016)	°C	1.02
NTO	Total Nitrogen (Viscarra Rossel et al. 2014c)	%	0.88
PHC	pH – CaCl ₂ (Viscarra Rossel et al. 2014d)	None	0.78
BDW	Bulk Density – Whole Earth (Viscarra Rossel et al. 2014e)	g/cm ³	0.73
CLY	Clay (Viscarra Rossel et al. 2014f)	%	0.60
AWC	Available Water Capacity (Viscarra Rossel et al. 2014g)	%	0.57
PTS1	Precipitation seasonality 1 (Williams et al. 2012)	ratio	0.47
SOC	Organic Carbon (Viscarra Rossel et al. 2014h)	%	0.38
EPX	Maximum monthly potential evaporation (Harwood et al. 2016)	mm	0.26
CONAREA	Contributing area (Gallant & Austin 2012)	index	0.11
ECE	Effective Cation Exchange Capacity (Viscarra Rossel et al. 2014i)	meq/100g	0.04

679 Table 2. Survey trips during the gap-filling period of 2018–2020 in chronological order,
680 noting the location and number of 181 new plots. Some trips also included revisits. Bold
681 entries indicate bioregions (Fig. 4) that were not sampled prior to gaps surveys.

Trip	Year	Bioregions sampled	Number of gap plots
1	2018	Australian Alps (AUA)	15
2	2018	Cobar Peneplain (COP) , Murray Darling Depression (MDD), Mulga Lands (MUL), Channel Country (CHC), Broken Hill Complex (BHC)	16
3	2018	NSW South Western Slopes (NSS), Brigalow Belt South (BBS) , Nandewar (NAN) , New England Tablelands (NET)	11
4	2018	Cape York Peninsula (CYP)	12
5	2018	Cape York Peninsula (CYP)	12
6	2018	Flinders Lofty Block (FLB)	2
7	2018	Flinders Lofty Block (FLB)	2
8	2018	Kanmantoo (KAN)	12
9	2019	Brigalow Belt South (BBS) , New England Tablelands (NET) , South Eastern Queensland (SEQ)	12
10	2019	Arnhem Plateau (ARP), Pine Creek (PCK)	10
11	2019	Brigalow Belt South (BBS) , Brigalow Belt North (BBN) , South Eastern Queensland (SEQ)	12
12	2019	Gawler (GAW), Stony Plains (STP)	7
13	2019	Northern Kimberley (NOK) , Dampierland (DAL)	12
14	2019	Carnarvon (CAR) , Yalgoo (YAL) , Geraldton Sandplains (GES)	11
15	2019	Murchison (MUR), Yalgoo (YAL), Avon Wheatbelt (AVW)	12
16	2019	Broken Hill Complex (BHC)	2
17	2019	Tasmanian Northern Midlands (TNM)	6
18	2020	Ben Lomond (BEL), Tasmanian Central Highlands (TCH)	15

682

Fig. 1. a) Visual representation of the generalised dissimilarity model (GDM) based on vascular plant species composition in TERN Ausplots 'establishment' plots. RGB colour scale represents position along the first three PCA axes of GDM-transformed predictors included in the fitted model. Similarity in colour represents similarity in species composition; b) GDM extrapolation index: sum of transformed grid values outside the input data range (total amount of extrapolated turnover).

Fig. 2. GDM *I*-spline partial functions, depicting the rate of species turnover along the 16 environmental gradients retained in the model, in order of variable importance, based on height of the *I*-spline function. See Table 1 for variable descriptions and importance values.

Fig. 3. a) Map of Australia showing ecological distance (Euclidean distance in biotically-scaled environmental space) of each pixel to the most similar Ausplots site (NND), as predicted from a GDM of plant species composition. Based on 582 'establishment' plots (points). High scoring regions represent ecological gaps in the sampling coverage that were targeted in subsequent surveys; b) NND as in previous panel, based on 763 plots (points), including 181 gap-filling plots; c) drop in ecological distance to most similar Ausplots location (NND) after 181 gap-filling plots were surveyed over 18 trips; calculated as the difference between maps before (a) and after (b) gap-filling.

Fig. 4. Australia, highlighting Australian States and Territories and the IBRA bioregions targeted by 'gap plots' in 2018–2020 (Table 2; Department of the Environment 2012).

Fig. 5. a) Drop in mean ecological distance to nearest sampled Ausplots site (NND) with the addition of 18 gap-filling trips in chronological order. First point is establishment plots mean. Trips were expeditions targeting particular regions in which an average of 10 (range 2–16) new plots were surveyed. Some trips focused on revisits to established sites, with smaller numbers of local gap-filling plots being established (e.g. 7th and 8th points above, corresponding to trips 6–7 in Table 2); b) Cumulative proportion of Australia's land area serviced by decreased distance to nearest Ausplots site in ecological space (NND) due to 18 gap-filling surveys 2018–2020 across a range of minima. 18% of the continent was serviced with a minimum NND drop of 0.05; c) Changes in the real sampling of ecological space as visualised via cumulative mean multivariate dispersion (y-axis; MVD, mean distance to centroid in Principal Coordinates Space, based on MDS of Bray-Curtis dissimilarities). The x-axis represents additional surveys of plots in chronological order.