Environmental Conservation

cambridge.org/enc

Research Paper

Cite this article: McKenzie PF et al. (2021) Partitioning tree diversity patterns to prioritize conservation investments. *Environmental Conservation* **48**: 75–83. doi: 10.1017/S0376892921000060

Received: 1 June 2020
Revised: 7 January 2021
Accepted: 19 January 2021
First published online: 11 March 2021

Keywords:
conservation; diversity partitioning; protected areas

Author for correspondence:
Patrick F McKenzie,
Email: p.mckenzie@columbia.edu

© The Author(s), 2021. Published by Cambridge University Press on behalf of Foundation for Environmental Conservation.

CAMBRIDGE UNIVERSITY PRESS

Thematic Section: Biodiversity Revisited

Partitioning tree diversity patterns to prioritize conservation investments

Patrick F McKenzie[1,2], Gwenllian D Iacona[1,3], Eric R Larson[1,4] and Paul R Armsworth[1]

[1]Department of Ecology and Evolutionary Biology, University of Tennessee, TN, USA; [2]Department of Ecology, Evolution, and Environmental Biology, Columbia University, NY, USA; [3]Resources for the Future, Washington, DC, USA and [4]Department of Natural Resources and Environmental Sciences, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Summary

The available tools and approaches to inform conservation decisions commonly assume detailed distribution data. We examine how well-established ecological concepts about patterns in local richness and community turnover can help overcome data limitations when planning future protected areas. To inform our analyses, we surveyed tree species in protected areas in the southern Appalachian Mountains in the eastern USA. We used the survey data to construct predictive models for alpha and beta diversity based on readily observed biophysical variables and combined them to create a heuristic that could predict among-site richness in trees (gamma diversity). The predictive models suggest that site elevation and latitude in this montane system explain much of the variation in alpha and beta diversity in tree species. We tested how well resulting protected areas would represent species if a conservation planner lacking detailed species inventories for candidate sites were to rely only on our alpha, beta and gamma diversity predictions. Our approach selected sites that, when aggregated, covered a large proportion of the overall species pool. The combined gamma diversity models performed even better when we also accounted for the cost of protecting sites. Our results demonstrate that classic community biogeography concepts remain highly relevant to conservation practice today.

Introduction

Habitat destruction remains a key driver of biodiversity loss (Wilcove et al. 1998, Maxwell et al. 2016), and preventing biodiversity loss is a primary goal of conservation efforts (Johnson et al. 2017). To protect biodiversity, conservation organizations often set aside land from development with the goal of keeping its natural habitat intact (Chape et al. 2008, Le Saout et al. 2013), but they often face important funding shortfalls (McCarthy et al. 2012). As such, organizations active in land protection must prioritize their investments to maximize their progress towards protecting biodiversity (Moilanen et al. 2009, Groves & Game 2016). This requires both protecting places rich in species and also ensuring that different parts of the protected area network cover locations that are important for complementary sets of species (Gaston et al. 2006). Community biogeography has long focused on understanding such patterns of species distribution, both in local richness and in community turnover (Koleff & Gaston 2002, Mena & Vázquez-Domínguez 2005, Qian & Ricklefs 2012). Here, we re-examine the potential for using ideas from community biogeography to help prioritize areas for protection.

Historically, total area of protected land has been cited as a metric of progress towards preserving biodiversity (Chape et al. 2005). However, because of inconsistency in the quality of protected areas and in the type of habitat protected, total protected area coverage alone is not an adequate indicator of biodiversity preservation or of habitat protection (Nagendra 2008, Joppa & Pfaff 2009, Nelson & Chomitz 2011). Simply maximizing protected area coverage favours protecting 'cheap' areas – land with few alternative competing uses. Indeed, historical protection biases have led to high-elevation, rugged environments with poor-quality soils being disproportionately represented among protected areas (Joppa & Pfaff 2009, Aycrigg et al. 2013). So how should a conservation organization approach prioritizing sites for protection in order to conserve biodiversity? Very commonly, such organizations will not have the luxury of performing detailed on-site inventories of species before committing resources for each land acquisition they are pursuing. Instead, they must rely on models or other estimation methods to decide what species a site is likely to contain (Le Saout et al. 2013).

One common current practice is first to model the distributions of each individual species within some taxon of interest (Early et al. 2008, de Pous et al. 2011, Villero et al. 2017). Typically, these are modelled independently (but see Pollock et al. 2014, Rougier et al. 2015). The

individual species distributions are then usually combined to estimate biodiversity patterns, something we will refer to as a stacked distribution model approach. Stacked distribution models are used to prioritize locations for protection based on how well they sample the combined set of estimated distributions for species in the community. However, modelled species distributions are known to suffer from commission errors (reporting false occurrences). It can be very challenging to understand and account for how these errors accumulate and affect resulting prioritizations when combining distribution maps for many different species (Araujo & Guisan 2006, Feeley & Silman, 2011 Franklin 2013). Dissatisfaction with these methods has led researchers to consider ways to prioritize land acquisition that do not require species data at all. Instead, priorities could be based on habitat types or ecosystem properties (e.g., soils, topography). The idea behind this is established in a body of literature that describes 'conserving nature's stage' as a way to strategize conservation based on more easily observed physical properties of land (Anderson et al. 2015, Beier et al. 2015, Lawler et al. 2015). This framework emphasizes the protection of geodiversity, or different types of landscapes, to set aside land in a way that is robust to a changing environment. The 'actors' – the species supported on those landscapes – can then move around those landscapes as needed.

Here, we explore a third approach, one sourced in well-established ideas from community biogeography. Specifically, we explore the performance of prioritization methods that are based on trying to predict community-level properties (local richness and turnover). This is in contrast to an approach of first disaggregating to predict individual species distributions and then re-aggregating predictions to understand biodiversity patterns, or to one of ignoring species information altogether.

In summarizing community-level properties, we will refer throughout to the ecological principles of alpha, beta and gamma diversity, which are well-known to conservation practitioners. Alpha diversity refers to the value of biodiversity contained within a local area. Species richness within a protected area, for example, is a measure of alpha diversity (Gering & Crist 2002). Beta diversity, in contrast, describes the turnover in diversity between discrete sites (Whittaker 1960, 1972, Legendre & De Cáceres 2013, Socolar et al. 2016). For example, two sites containing many of the same species might each be high in alpha diversity, but their beta diversity value would be low because of the overlap in species between the sites. Measures such as the Jaccard index and Sorensen–Dice index are commonly used for quantifying beta diversity between sites, but many methods for calculating beta diversity exist in the literature. These methods are summarized in reviews by Ellison (2010) and Tuomisto (2010a, 2010b), and a comparison is made by Legendre and De Cáceres (2013). New methods continue to be proposed, such as those described in Baselga (2010), Podani and Schmera (2011) and Carvalho et al. (2012). Given that so many different approaches exist, working with beta diversity can be confusing in practice. However, all beta diversity methods place a common emphasis on dissimilarity in species representation across multiple sites. Conservation planning approaches emphasizing representation (e.g., Chadés et al. 2015, Chauvenet et al. 2017) are in effect seeking to account for some measure of beta diversity, and so calculating some metric for it is therefore important (Jost 2007).

Alpha and beta diversity are important to consider jointly for most conservation objectives, given that neither alpha nor beta diversity values alone guarantee adequate species coverage by protected areas. Gamma diversity is the number of species contained in a region and therefore considers both the biodiversity within individual sites (alpha diversity) and the turnover in biodiversity between them (beta diversity), which are independent from one another, to describe the total biodiversity across a landscape (Gering et al. 2003, Jost 2007, Tuomisto 2010a, 2010b). Gamma diversity is traditionally formulated as being the product of alpha and beta diversity (Whittaker 1960, 1972), although more recent approaches have considered an additive relationship (Lande 1996, Veech et al. 2002). Regardless, we suggest that predicted values of alpha and beta diversity could be a useful approach for predicting total diversity among a set of sites (gamma diversity) to then prioritize areas for protection.

In this paper, we test this approach by applying it to data obtained from surveying tree diversity on a set of sites protected by The Nature Conservancy (TNC), a large non-profit land trust (Birchard 2005). The study sites in question were protected in part to conserve forested regions of the southern Appalachian Mountains. This region contains many tree species that are disproportionately under-protected and that will be important for future biodiversity conservation (Jenkins et al. 2015). It also contributes an elevational gradient and a north–south movement corridor that could act as a buffer for species responses to climate change (Lawler et al. 2013). We use the ecological properties of the sites to build two predictive models, evaluating both alpha and beta diversity. We then demonstrate how the predictions from the two models can be used jointly to estimate gamma diversity, or total biodiversity, across groups of sites. We show that these gamma diversity estimates can improve protected area selection under two different selection regimes, one focused only on representing species within protected areas and one that also considers the cost of securing that protection.

Methods

Data collection

We surveyed tree species diversity on 27 TNC protected areas (hereafter 'sites') distributed across 10 states within the Central Appalachian Forest, Southern Blue Ridge, and Cumberland and Southern Ridge and Valley ecoregions of the eastern USA (Fig. 1). These ecoregions constitute part of the Appalachian Mountains of the eastern USA. This region is a conservation priority and is known for its diversity of tree species (Currie & Paquin 1987, Stein et al. 2000, Jenkins et al. 2015).

We used a spatial random point generator (ArcMap version 10.1, ESRI, Redlands, CA, USA) to identify 20 survey points prior to visiting each protected area – 10 of which were near (within 100 m of) the protected area edge and 10 of which were in the core. We visited each sampling site between May and September of 2013. At each point, location data were recorded using handheld GPS (Garmin E-Trex 20). We then identified the 10 trees nearest to each point to the species level (5400 trees in total). Resampled and averaged accumulation curves from each site are provided in Supplementary Fig. S1 (available online) to show that our sampling procedure adequately characterized the relative contributions of each site to tree diversity in the study. We considered trees that had a diameter at breast height (dbh) greater than 10 cm, but counted any dbh scrub oak (*Quercus ilicifolia*) as a tree because it is the dominant tree species in the scrub oak–heath community that is a conservation priority at some of our protected areas (The Nature Conservancy 1997), and it rarely exceeds 10 cm dbh. All species nomenclatures follow Kirkman et al. (2007), except species

https://doi.org/10.1017/S0376892921000060 Published online by Cambridge University Press

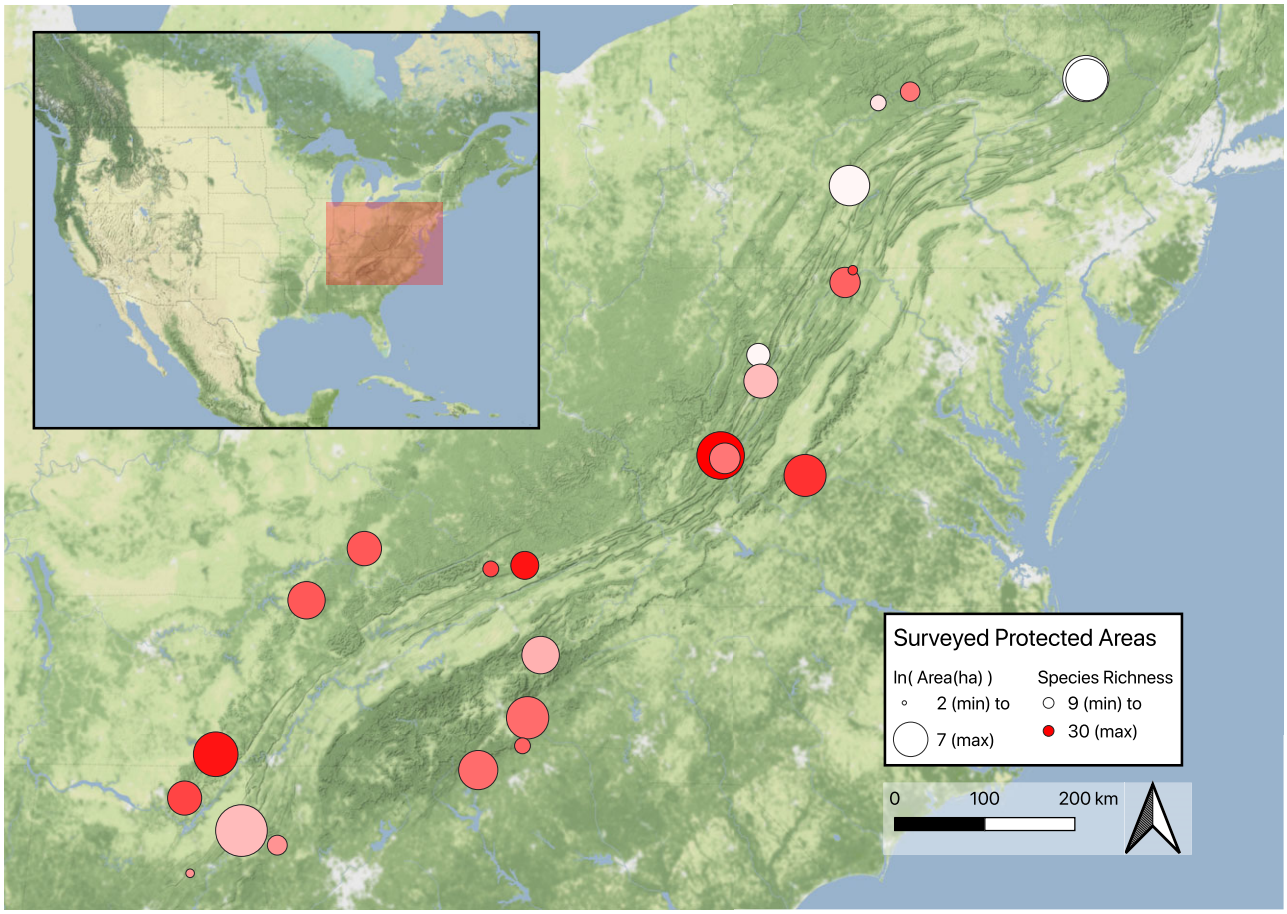


Fig. 1. Sites that were surveyed for tree diversity. Area (ha) is scaled using the natural logarithm, and the relative scaled values are represented on the map by dot diameters, ranging on a continuous scale from small to large. Surveyed species richness at each site is represented by the colour of each dot, ranging on a continuous scale from light to dark.

only found north of Maryland, which follow Elias (1980). All *Crataegus* and *Amelanchier* were identified to the genus level only.

The tree sampling data were paired with GPS location data for each point. Elevation data for each sampled location were obtained from the NASA Shuttle Radar Topography Mission (Rodriguez et al. 2006). For our analysis, all data were aggregated at the level of the protected area using the mean values of the 20 points in the protected area to determine the protected area mean elevation and mean latitude. The distance between the highest and lowest sampling points was used to calculate the elevation range within each protected area.

Across sites, mean site elevation ranged from 130 to 1700 m; latitude ranged from 34° to 41.5°; within-site elevation change varied from 30 to 400 m; and parcel area ranged from 10 ha to nearly 900 ha (Table 1).

Data on protected area parcel sizes and costs were provided by TNC (Table 1). The protected areas were acquired by TNC at different times between 2000 and 2009. As such, we needed to correct the cost information for inflation. Therefore, we converted this cost information to 2000 US$ equivalents using a state-level housing price index.

Models of alpha, beta and gamma diversity

To model alpha diversity, we fit a generalized linear model to test the predictive capacities of site latitude, mean site elevation, within-site elevation range and site area for estimating within-site richness. We limited our analysis to these variables to avoid overparameterizing our model given our limited sample size (Zuur et al. 2007). Because we present alpha diversity as counts of site species richness, we assumed the dependent variable to be Poisson-distributed. We tested predictor variables for collinearity and found them to be sufficiently independent from one another to proceed. We added a quadratic term for site latitude because an examination of residuals when only linear terms were included strongly suggested a non-linear effect of latitude. Introduction of this quadratic term resulted in large reductions in Akaike information criterion (AIC) values for the alpha diversity model and in a higher pseudo-R^2 value, indicating that the model was explaining more variance. We did not consider other interaction effects, having no a priori reason to expect any particularly strong interaction effects from among the many that are possible.

To model beta diversity, we performed a multiple regression on matrices of pairwise distances between sites' ecological variables (MRM; see Lichstein 2007). MRM, an extension of the Mantel test (Mantel 1967) to a regression framework, is a flexible multivariate method for examining relationships between a series of predictive matrices and a response matrix. A MRM can be constructed to accommodate linear as well as more complex relationships (Lichstein 2007). In our analysis, the response variable matrix consisted of a Sorensen–Dice dissimilarity index, which is one of the most common ways to represent beta diversity among pairs of sites

https://doi.org/10.1017/S0376892921000060 Published online by Cambridge University Press

Table 1. Summary of site characteristics. For each variable, we report the mean with SD in parentheses. Alpha diversity is shown in units of species and beta diversity is represented by the averaged pairwise Sorensen–Dice dissimilarity indices. Costs are adjusted to year 2000 US$ and represent the acquisition cost for each site.

Variable	Mean (SD)
Elevation	571 m (330 m)
Latitude	37.6° (2.5°)
Elevation change	156 m (109 m)
Area	189 ha (207 ha)
Alpha diversity	19.2 species (6.0 species)
Beta diversity	0.58 (0.09)
Acquisition cost per site	$793 000 ($943 000)

(Chao et al. 2006, Anderson et al. 2011, Legendre & De Cáceres 2013). We chose this particular metric because it is well-known, easy to calculate and therefore very interpretable, particularly within our prioritization that emphasizes species representation. The Sorensen–Dice dissimilarity index describes turnover in species richness between sites and is calculated as follows for each cell in the distance matrix, where sets X and Y are composed of species observed in each of two sites:

$$Dissimilarity = 1 - \frac{2|X \cap Y|}{|X| + |Y|}$$

Each cell in the response variable matrix therefore represents the magnitude of turnover among species present between two sites. The independent variables chosen were the same as in the alpha diversity model, although now structured as pairwise distance matrices: site latitude, site elevation, site area and within-site elevation range. In these matrices, each cell represents the difference in values for the variable in question between two sites. Model predictors were again examined for collinearity. Pairwise distances between site areas were natural-log-transformed, and distances in terms of latitude squared were not included in this regression because the distribution of the residuals did not warrant including the extra term.

We used our models of alpha and beta diversity to explore how estimated gamma diversity could be used to inform selection of protected areas. The following heuristic equation for estimating gamma diversity between two sites incorporates alpha and beta diversity into the site selection process:

$$\gamma_{ab} = (\alpha_a + \alpha_b) \times \beta_{ab}$$

Here, α_a and α_b refer to individual site richness values and β_{ab} refers to the Sorensen–Dice dissimilarity index for the two sites. We chose this gamma diversity metric again because of its simple transparency and because it adequately captures common understanding about the relationships between alpha, beta and gamma diversity. This choice of index makes clear that as the alpha diversity indices of the two sites increase, or as the dissimilarity between the sites increases, the gamma diversity index between the two sites will also increase (see Veech et al. 2002 for a discussion of alternative ways to combine alpha and beta diversity into measures of gamma diversity).

We constructed a matrix of pairwise gamma diversity estimates using predicted values for each site from the alpha and beta models. The average of all pairwise gamma diversity indices including a specific site was used as a proxy for the site's overall contribution to gamma diversity.

Protected area selection strategies

After building models predicting alpha, beta and gamma diversity, we examined how these models could be used to select sites for protection. In reality, all 27 sites included in our data collection were actually protected. For our analysis, we focused on hypothetical scenarios of what would have happened had TNC not had sufficient funding to protect all 27 sites. If the organization were more limited in funding, which of these sites should be the very top priorities? We also set a goal of ensuring as many species as possible received at least some protection within a selected protected area network. Other protection goals, such as ensuring a particular number of occurrences of each species within selected protected area networks, are also possible (Willis et al. 1996, Rodrigues et al. 2000).

We benchmarked the relative performance of different strategies for selecting protected areas against the best possible choice and a random selection strategy. To calculate the maximum possible coverage of species by a protected area network of a given number of sites, we iterated through all possible choice combinations (exhaustive search). The maximal possible coverage of species by a reserve network of a given size was found by counting the number of species present on at least one protected site for every combination of selected sites. We also calculated the expected species coverage if choosing protected areas at random from within the set of 27 sites. One hundred randomly selected sets of sites were drawn for each budget level and the average number of species included in the set of protected sites was calculated.

Next, we used our models of alpha, beta and gamma diversity to guide the selection of sites to protect. The alpha diversity model was used to predict which site would have the highest alpha diversity based on predictor variables, and new sites were added in order of decreasing alpha diversity. For beta diversity, sites were selected in order of decreasing mean value across the predicted matrix of pairwise Sorensen–Dice values, so that sites that were predicted to be most different from the other sites were generally selected first. To account for variance in model predictions, this process was repeated for 100 random samples through the variance of predictions for each of the two models. As detailed above, we combined our alpha and beta diversity models to estimate a site's overall contribution to gamma diversity. To test how this could be used to inform decisions about which sites to protect, we selected sites in order of decreasing mean predicted gamma diversity value across all of each site's pairwise gamma diversity estimates, so that priority was given to sites that were predicted to have high richness and also not to overlap much with other sites. This was repeated for 100 random samples through the mean squared error of predictions by the alpha and beta diversity models.

In addition to considering the performance of different approaches given specific numbers of sites selected in a protected area network, we also considered the ability of our models to perform well when considering the cost of the protected areas. To include cost information, we considered the one-time acquisition costs of different sites and considered a range of budget intervals (23 values spread logarithmically from US$18 000 to US$20 000 000) to select sites for the protected area network. For our analyses of the role of costs, we focused on the gamma diversity strategy, as

https://doi.org/10.1017/S0376892921000060 Published online by Cambridge University Press

it was the best performing of those based on the diversity models when considering only biodiversity data.

Results

Tree diversity survey results

Our field sampling effort identified a total of 94 species across the 5400 individual trees on 27 sites. The mean sampled species richness (alpha diversity) of these sites was 19.2 tree species, with a minimum of 9 tree species and a maximum of 30 tree species (Table 1). The average Sorensen–Dice dissimilarity (beta diversity) across the 27 sites was 0.58, reflecting substantial turnover among sites (Table 1). Between individual sites, the turnover was highly variable; for instance, the highest turnover was between a site sampled in Alabama and a site sampled in North Carolina. Of the 39 species observed across the two sites, only 1 was recorded at both sites (Sorensen–Dice dissimilarity = 0.95). In contrast, the lowest turnover was observed between a site sampled in Tennessee and one in Virginia, where for the 30 species observed across both sites, 22 were shared (Sorensen–Dice dissimilarity = 0.15).

A few tree species were very frequent across all sites, but infrequent species contributed to our sampled diversity, with 30 of the 94 observed tree species recorded just five times or fewer. The most commonly observed species was *Acer rubrum*, accounting for about a tenth (531) of all observations; however, it was the most common tree at only one site, in Tennessee. Across sites, dominant species included those in the genera *Pinus*, *Liriodendron*, *Prunus*, *Betula*, *Nyssa*, *Robinia*, *Picea*, *Fraxinus*, *Celtis* or *Tsuga*, mirroring an expectation of highly heterogeneous community composition and high diversity across Appalachia.

Diversity models

Model coefficients for the two latitude terms indicate that alpha diversity increases with latitude at the lowest latitudes, but decreases with latitude at higher latitudes (Table 2). Additionally, site elevation has a negative relationship with site alpha diversities. Altogether, the alpha diversity model explained some of the variation in tree richness values of sites ($R^2 = 0.33$). Perhaps surprisingly, neither site area nor within-site elevation range contributed significant predictive capacity to the models.

The beta diversity model used pairwise differences in latitude, elevation, within-site elevation ranges and ln(area) values to predict pairwise Sorensen–Dice dissimilarity values (Table 3). Pairwise distances between site latitude values and site elevation values were significant ($p < 0.001$) predictors of Sorensen–Dice dissimilarity indices. The positive coefficient for distance between latitudes implies that differences in latitude contribute to high values for species turnover, as does the positive coefficient for distance between site elevations. The McFadden R^2 value of 0.25 indicates that a moderate amount of the variance in beta diversity is explained by just the few predictors included in the model. As in the alpha diversity model, pairwise distances between site areas and within-site elevation ranges were not significant predictors of species turnover between sites.

Protected area selection strategies

We compared the number of species covered by different protected area strategies when limited by the number of sites that can be protected (Fig. 2). The filled circles in Fig. 2 are the maximum coverage solutions and represent the 'best possible' answer, namely the combination of sites that would be selected to maximize species richness if we had perfect knowledge of species presence. The distance between these solutions and the results from the other methods is a consequence of 'hiding' the species presence data and instead making predictions based on site characteristics alone (e.g., latitude, elevation, etc.) from our alpha, beta and gamma diversity models. Of these methods, our gamma diversity estimation method performed best. Selection by predicted alpha diversity alone performed well at low thresholds, but lost its advantage at selection sizes of 15 sites or greater. Selection by predicted beta diversity performed significantly worse than random selection at low selection sizes.

We then also integrated information on the cost of protecting sites and repeated site prioritizations for a range of budget thresholds. We first used known information about species richness at each site to sample sites under each price threshold and to produce a 'near maximum richness' set of site selections. Because we could not exhaustively sample all possibilities at each threshold, each value might be slightly underestimated. We then used the predicted number of species at each site (using only the gamma diversity prediction model) to make site selections under each threshold, and we compared these results to that of random selection. Gamma diversity-estimated site selection performed much better than random site selection at each price threshold (Fig. 3).

With both prioritization scenarios (limited by site number or limited by cost), the gamma diversity selection method performed better than random selection. The gamma diversity model worked particularly well in comparison to random site selection when price thresholds were considered, covering over 98% of the area under the maximum solution curve while random selection only covered *c.* 90% of the area under the curve.

Table 2. Alpha diversity regression results.

Variable	B	P-value	R^2
			0.33
Constant	0.41	1.00	
Latitude Dist	0.03	1.00e–6	
Elev Dist	2.52e–4	6.22e–4	
ElevRange Dist	6.68e-6	0.96	
ln(Area) Dist	–5.96e–4	0.96	

Discussion

We aimed to determine whether abiotic factors such as latitude and elevation would help prioritize the purchasing of a protected area network that maximizes tree species coverage. We surveyed 27 sites recently acquired by TNC in the southern Appalachian Mountains for tree diversity, and we paired these data with data on the physical characteristics of the protected areas. We then revisited the community biogeography principles of alpha, beta and gamma diversity to separately model local patterns of species richness and regional patterns of species turnover among the tree diversity data. We modelled alpha and beta diversity separately. We then combined predictions from alpha and beta diversity models into an estimate of gamma diversity. Our gamma diversity estimate performed well under two scenarios for prioritizing protected area selections: one scenario in which we were limited by the number of protected areas in the network and one in which we were limited by total cost.

https://doi.org/10.1017/S0376892921000060 Published online by Cambridge University Press

Table 3. Beta diversity regression results.

Variable	B	SE	P-value	AIC	McFadden R^2
				150.40	0.25
Constant	−60.61	14.42	2.63e−5		
AvgSiteLatitude	3.46	0.77	7.00e−6		
AvgSiteLatitude^2	−0.05	0.01	5.34e−6		
AvgPtElevation	−6.98e−4	1.63e−4	1.82e−5		
ElevationRange	6.66e−4	4.98e−4	0.18		
ln(Area)	−0.04	0.04	0.32		

AIC = Akaike information criterion.

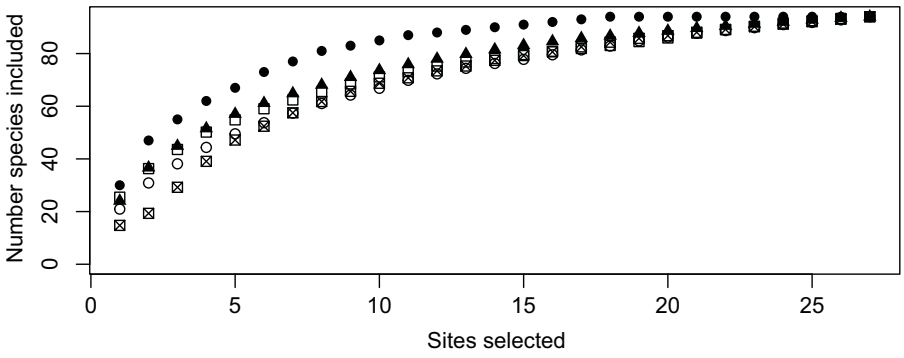


Fig. 2. Site selection efficacy by number of sites. Filled circles: maximum coverage solutions. Triangles: average coverage by site selection using pairwise gamma diversity estimates from alpha/beta model predicted values. Empty circles: random site selection. Empty squares: site selection by predicted alpha diversity alone. Crossed squares: site selection by predicted beta diversity alone.

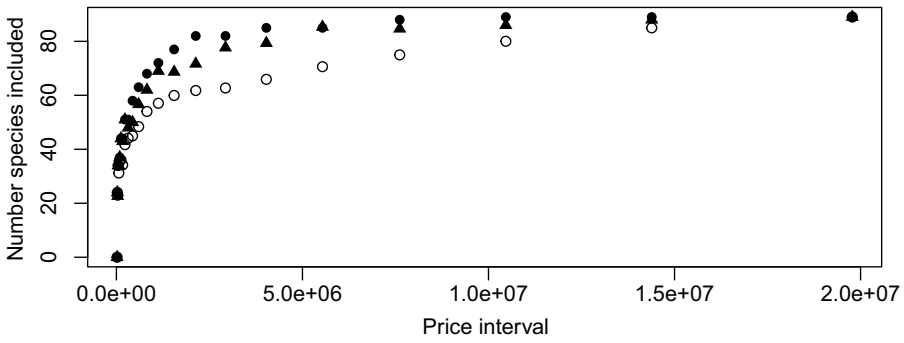


Fig. 3. Site selection efficacy with price thresholds. Filled circles: maximum richness possible at each price interval. Black triangles: average number of species included at each interval if selecting sites by gamma approximation. Empty circles: average number of species included at each interval by random selection.

Our results use classic ideas about diversity partitioning to make predictions about species coverage (Gering & Crist 2002, Jost 2007, Socolar et al. 2016). The simplicity of the models demonstrates the predictive power of just a handful of variables. We found that variables such as latitude and elevation of the protected areas influenced both the alpha diversity of individual sites and the beta diversity between pairs of sites. Incorporating predictions of the separate models for alpha and beta diversity for protected area prioritization allowed us to optimize our choices to include sites with high local species richness while also favouring sites that were likely to include species different from those already in the network. Including additional predictors could further increase the explanatory power of the analysis (e.g., see Albuquerque & Beier 2015). However, fitting models that include more variables would require survey data from more sites.

Including cost thresholds improved the performance of our method relative to the selection scenario that was limited by number of sites. This suggests a disconnect between the cost of individual parcels and their contribution to species coverage in a protected area network, which our site selection procedure leveraged to accumulate many species compared to random selection as price thresholds increased. Costs are one of many socioeconomic factors that could be considered, and our method would likely improve further under other socioeconomic constraints (e.g., threats or proximity to urban areas; Armsworth et al. 2017). Additionally, further analyses might benefit from also considering management costs, since previous work has found that management costs did not correlate with acquisition costs (Armsworth et al. 2011).

A conventional approach to predicting species diversity might be to make predictions about occupancy from stacked species distribution models (SDMs) (Villero et al. 2017). However, conservation planners anticipating that species diversity data may not be available, or may change in the future, have sought to use abiotic characteristics to guide site prioritizations (Anderson et al. 2015, Beier et al. 2015). We show here that very simple abiotic site characteristics can meaningfully inform site selection in comparison to empirical data on species coverage. Moreover, the predictions made by the models presented here are specifically informed by data at the site level, and they might not be as susceptible to problems at fine scales as SDM-based methods (Hernandez et al. 2006, Lobo et al. 2010).

We made a number of choices with our research design. First, we examined a particular conservation context. Our data were obtained from a sample of sites protected by a single conservation

https://doi.org/10.1017/S0376892921000060 Published online by Cambridge University Press

organization, albeit one that operated a similar land protection mission to many other land trusts and regional public agencies during the time these sites were protected. Even when working with one organization, many factors contribute to determining just which particular properties end up being protected. Acquisitions of the particular set of protected areas we consider were guided by a collaborative ecoregional planning process, but were constrained by what land was available for purchase and what funding was available (Groves 2003, McDonald-Madden et al. 2008, Lennox et al. 2017). The outcome of this process is encapsulated in TNC's ecoregional plans for the study area (e.g., The Nature Conservancy 2000) covering the period of these acquisitions. In identifying our sample of sites for this work, we also examined internal TNC documents detailing why each individual parcel was acquired and we consulted with individual TNC staff involved in the relevant land deals. All of the parcels in question were acquired at least partly to protect forested ecosystems, and the diversity of these forest ecosystems was commonly highlighted. A related point was that the data came from sites that were actually protected. As such, our analysis is most relevant when considering sites that are candidates for selection and, again, differences in diversity patterns may have resulted if we included sites at random from across the landscape.

Next, we focused on a conservation goal of species coverage. Conservation organizations, including TNC, pursue a multiplicity of objectives when protecting land, of which protecting biodiversity is often one. When considering how well protected area networks perform in terms of biodiversity, species coverage – our focus here – is one particularly relevant criteria (Church et al. 1996, Veach et al. 2017). However, we recognize that there are others that also warrant consideration. For example, if prioritizing rare species, the models presented here could also be adapted to maximize metrics weighted for representation of species of higher conservation priority (Arponen et al. 2005, Veach et al. 2017). Alternatively, one might assign greater weight to ensuring persistence of species covered somewhere within the protected area network (e.g., by valuing repeat occurrences of the species at different protected sites). Other design choices we made should also be considered to contextualize our results, including limiting our sampled sites to the southern Appalachian Mountains (a region known for high levels of tree diversity and that spans a substantial range of elevations) and sampling from randomly selected points within protected areas. Lastly, our approach here was limited specifically to considering diversity of tree species and might not be as effective for other species groups, such as those that are very mobile or are especially sensitive. Other metrics for site prioritization that do not require extensive on-the-ground surveying have been shown to be effective for other groups of species, such as predicted importance, predicted rarity-weighted richness and metrics related to environmental diversity and velocity (Albuquerque & Beier 2015, 2016, Carroll et al. 2017).

Biodiversity partitioning using alpha, beta and gamma diversity is rooted in fundamental concepts about patterns in species distributions. Here, we have revisited biodiversity partitioning as a tool for conservation planning and have demonstrated its continued relevance for explaining patterns in diversity and for guiding conservation decision-making.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/S0376892921000060

Acknowledgements. We thank The Nature Conservancy for providing relevant data and colleagues at University of Tennessee for useful discussions.

Financial support. This material is based on work partially supported by the National Science Foundation (grant number 121141).

Conflict of interest. None.

Ethical standards. None.

References

Albuquerque F, Beier P (2015) Using abiotic variables to predict importance of sites for species representation. *Conservation Biology* 29: 1390–1400.

Albuquerque F, Beier P (2016) Predicted rarity-weighted richness, a new tool to prioritize sites for species representation. *Ecology and Evolution* 6: 8107–8114.

Anderson MG, Comer PJ, Beier P, Lawler JJ, Schloss CA, Buttrick S et al. (2015) Case studies of conservation plans that incorporate geodiversity. *Conservation Biology* 29: 680–691.

Anderson MJ, Crist TO, Chase JM, Vellend M, Inouye BD, Freestone AL et al. (2011) Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters* 14: 19–28.

Araujo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography* 33: 1677–1688.

Armsworth PR, Cantú-Salazar L, Parnell M, Davies ZG, Stoneman R (2011) Management costs for small protected areas and economies of scale in habitat conservation. *Biological Conservation* 144: 423–429.

Armsworth PR, Jackson HB, Cho S-H, Clark M, Fargione JE, Iacona GD et al. (2017) Factoring economic costs into conservation planning may not improve agreement over priorities for protection. *Nature Communications* 8: 1–10.

Arponen A, Heikkinen RK, Thomas CD, Moilanen A (2005) The value of biodiversity in reserve selection: representation, species weighting, and benefit functions. *Conservation Biology* 19: 2009–2014.

Aycrigg JL, Davidson A, Svancara LK, Gergely KJ, McKerrow A, Scott JM (2013) Representation of ecological systems within the protected areas network of the continental United States. *PLoS One* 8: e54689.

Baselga A (2010) Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography* 19: 134–143.

Beier P, Sutcliffe P, Hjort J, Faith DP, Pressey RL, Albuquerque F (2015) A review of selection-based tests of abiotic surrogates for species representation. *Conservation Biology* 29: 668–679.

Birchard B (2005) *Nature's Keepers: The Remarkable Story of How the Nature Conservancy Became the Largest Environmental Group in the World*. Hoboken, NJ, USA: John Wiley & Sons.

Carroll C, Roberts DR, Michalak JL, Lawler JJ, Nielsen SE, Stralberg D et al. (2017) Scale-dependent complementarity of climatic velocity and environmental diversity for identifying priority areas for conservation under climate change. *Global Change Biology* 23: 4508–4520.

Carvalho JC, Cardoso P, Gomes P (2012) Determining the relative roles of species replacement and species richness differences in generating beta-diversity patterns. *Global Ecology and Biogeography* 21: 760–771.

Chadés I, Nicol S, van Leeuwen S, Walters B, Firn J, Reeson A et al. (2015) Benefits of integrating complementarity into priority threat management. *Conservation Biology* 29: 525–536.

Chao A, Chazdon RL, Colwell RK, Shen T (2006) Abundance-based similarity indices and their estimation when there are unseen species in samples. *Biometrics* 62: 361–371.

Chape S, Harrison J, Spalding M, Lysenko I (2005) Measuring the extent and effectiveness of protected areas as an indicator for meeting global biodiversity targets. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360: 443–455.

Chape S, Spalding M, Jenkins M (2008) *The World's Protected Areas: Status, Values and Prospects in the 21st Century*. Berkeley, CA, USA: University of California Press.

Chauvenet ALM, Kuempel CD, McGowan J, Beger M, Possingham HP (2017) Methods for calculating Protection Equality for conservation planning. *PLoS one* 12: e0171591.

Church RL, Stoms DM, Davis FW (1996) Reserve selection as a maximal covering location problem. *Biological Conservation* 76: 105–112.

https://doi.org/10.1017/S0376892921000060 Published online by Cambridge University Press

Currie DJ, Paquin V (1987) Large-scale biogeographical patterns of species richness of trees. *Nature* 329: 326–327.
de Pous P, Beukema W, Weterings M, Dümmer I, Geniez P (2011) Area prioritization and performance evaluation of the conservation area network for the Moroccan herpetofauna: a preliminary assessment. *Biodiversity and Conservation* 20: 89–118.
Early R, Anderson B, Thomas CD (2008) Using habitat distribution models to evaluate large-scale landscape priorities for spatially dynamic species. *Journal of Applied Ecology* 45: 228–238.
Elias TS (1980) *The Complete Trees of North America. Field Guide and Natural History*. New York, NY, USA: Van Nostrand Reinhold Company & Times Mirror Magazines, Inc.
Ellison AM (2010) Partitioning diversity 1. *Ecology* 91: 1962–1963.
Feeley KJ, Silman MR (2011) Keep collecting: accurate species distribution modelling requires more collections than previously thought. *Diversity and Distributions* 17: 1132–1140.
Franklin J (2013) Species distribution models in conservation biogeography: developments and challenges. *Diversity and Distributions* 19: 1217–1223.
Gaston KJ, Charman K, Jackson SF, Armsworth PR, Bonn A, Briers RA et al. (2006) The ecological effectiveness of protected areas: the United Kingdom. *Biological Conservation* 132: 76–87.
Gering JC, Crist TO (2002) The alpha–beta–regional relationship: providing new insights into local–regional patterns of species richness and scale dependence of diversity components. *Ecology Letters* 5: 433–444.
Gering JC, Crist TO, Veech JA (2003) Additive partitioning of species diversity across multiple spatial scales: implications for regional conservation of biodiversity. *Conservation Biology* 17: 488–499.
Groves C (2003) *Drafting a Conservation Blueprint: A Practitioner's Guide to Planning for Biodiversity*. Washington, DC, USA: Island Press.
Groves C, Game ET (2016) *Conservation Planning: Informed Decisions for a Healthier Planet*. New York, NY, USA: W. H. Freeman.
Hernandez PA, Graham CH, Master LL, Albert DL (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29: 773–785.
Jenkins CN, Van Houtan KS, Pimm SL, Sexton JO (2015) US protected lands mismatch biodiversity priorities. *Proceedings of the National Academy of Sciences of the United States of America* 112: 5081–5086.
Johnson CN, Balmford A, Brook BW, Buettel JC, Galetti M, Guangchun L, Wilmshurst JM (2017) Biodiversity losses and conservation responses in the Anthropocene. *Science* 356: 270–275.
Joppa LN, Pfaff A (2009) High and far: biases in the location of protected areas. *PLoS One* 4: e8273.
Jost L (2007) Partitioning diversity into independent alpha and beta components. *Ecology* 88: 2427–2439.
Kirkman KL, Brown CL, Leopold DJ (2007) *Native Trees of the Southeast: An Identification Guide*. Protland, OR, USA: Timber Press.
Koleff P, Gaston KJ (2002) The relationships between local and regional species richness and spatial turnover. *Global Ecology and Biogeography* 11: 363–375.
Lande R (1996) Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* 76: 5–13.
Lawler JJ, Ackerly DD, Albano CM, Anderson MG, Dobrowski SZ, Gill JL et al. (2015) The theory behind, and the challenges of, conserving nature's stage in a time of rapid change. *Conservation Biology* 29: 618–629.
Lawler JJ, Ruesch AS, Olden JD, McRae BH (2013) Projected climate-driven faunal movement routes. *Ecology Letters* 16: 1014–1022.
Le Saout S, Hoffmann M, Shi Y, Hughes A, Bernard C, Brooks TM et al. (2013) Protected areas and effective biodiversity conservation. *Science* 342: 803–805.
Legendre P, De Cáceres M (2013) Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters* 16: 951–963.
Lennox GD, Fargione J, Spector S, Williams G, Armsworth PR (2017) The value of flexibility in conservation financing. *Conservation Biology* 31: 666–674.
Lichstein JW (2007) Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecology* 188: 117–131.
Lobo JM, Jiménez-Valverde A, Hortal J (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography* 33: 103–114.
Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 209–220.
Maxwell SL, Fuller RA, Brooks TM, Watson JEM (2016) Biodiversity: the ravages of guns, nets and bulldozers. *Nature News* 536: 143.
McCarthy DP, Donald PF, Scharlemann JPW, Buchanan GM, Balmford A, Green JMH et al. (2012) Financial costs of meeting global biodiversity conservation targets: current spending and unmet needs. *Science* 338: 946–949.
McDonald-Madden E, Bode M, Game ET, Grantham H, Possingham HP (2008) The need for speed: informed land acquisitions for conservation in a dynamic property market. *Ecology Letters* 11: 1169–1177.
Mena JL, Vázquez-Domínguez E (2005) Species turnover on elevational gradients in small rodents. *Global Ecology and Biogeography* 14: 539–547.
Moilanen A, Possingham H, Wilson K (2009) *Spatial Conservation Prioritization: Quantitative Methods and Computational Tools*. Oxford, UK: Oxford University Press.
Nagendra H (2008) Do parks work? Impact of protected areas on land cover clearing. *AMBIO* 37: 330–337.
Nelson A, Chomitz KM (2011) Effectiveness of strict vs. multiple use protected areas in reducing tropical forest fires: a global analysis using matching methods. *PLoS One* 6: e22722.
Podani J, Schmera D (2011) A new conceptual and methodological framework for exploring and explaining pattern in presence–absence data. *Oikos* 120: 1625–1638.
Pollock LJ, Tingley R, Morris WK, Golding N, O'Hara RB, Parris KM et al. (2014) Understanding co-occurrence by modelling species simultaneously with a joint species distribution model (JSDM). *Methods in Ecology and Evolution* 5: 397–406.
Qian H, Ricklefs RE (2012) Disentangling the effects of geographic distance and environmental dissimilarity on global patterns of species turnover. *Global Ecology and Biogeography* 21: 341–351.
Rodrigues AS, Orestes Cerdeira J, Gaston KJ (2000) Flexibility, efficiency, and accountability: adapting reserve selection algorithms to more complex conservation problems. *Ecography* 23: 565–574.
Rodriguez E, Morris CS, Belz JE (2006) A global assessment of the SRTM performance. *Photogrammetric Engineering & Remote Sensing* 72: 249–260.
Rougier T, Lassalle G, Drouineau H, Dumoulin N, Faure T, Deffuant G et al. (2015) The combined use of correlative and mechanistic species distribution models benefits low conservation status species. *PLoS One* 10: e0139194.
Socolar JB, Gilroy JJ, Kunin WE, Edwards DP (2016) How should beta-diversity inform biodiversity conservation? *Trends in Ecology & Evolution* 31: 67–80.
Stein BA, Kutner LS, Adams JS (2000) *Precious Heritage: The Status of Biodiversity in the United States*. Oxford, UK: Oxford University Press on Demand.
The Nature Conservancy (1997) *A Natural Areas Inventory of Lackawanna County, Pennsylvania. Technical Report*. Arlington County, VA, USA: The Nature Conservancy.
The Nature Conservancy (2000) *Southern Blue Ridge Ecoregional Conservation Plan: Summary and Implementation Document*. Southern Appalachian Forest Coalition and The Nature Conservancy (eds). Arlington County, VA, USA: The Nature Conservancy.
Tuomisto H (2010a) A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography* 33: 2–22.
Tuomisto H (2010b) A diversity of beta diversities: straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography* 33: 23–45.
Veach V, Di Minin E, Pouzols FM, Moilanen A (2017) Species richness as criterion for global conservation area placement leads to large losses in coverage of biodiversity. *Diversity and Distributions* 23: 715–726.
Veech JA, Summerville KS, Crist TO, Gering JC (2002) The additive partitioning of species diversity: recent revival of an old idea. *Oikos* 99: 3–9.
Villero D, Pla M, Camps D, Ruiz-Olmo J, Brotons L (2017) Integrating species distribution modelling into decision-making to inform conservation actions. *Biodiversity and Conservation* 26: 251–271.
Whittaker RH (1960) Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30: 279–338.

https://doi.org/10.1017/S0376892921000060 Published online by Cambridge University Press

Whittaker RH (1972) Evolution and measurement of species diversity. *Taxon* 21: 213–251.

Wilcove DS, Rothstein D, Dubow J, Phillips A, Losos E (1998) Quantifying threats to imperiled species in the United States: assessing the relative importance of habitat destruction, alien species, pollution, overexploitation, and disease. *Bioscience* 48: 607–615.

Willis CK, Lombard AT, Cowling RM, Heydenrych BJ, Burgers CJ (1996) Reserve systems for limestone endemic flora of the Cape lowland fynbos: iterative versus linear programming. *Biological Conservation* 77: 53–62.

Zuur A, Ieno EN, Smith GM (2007) *Analyzing Ecological Data*. Berlin, Germany: Springer.

https://doi.org/10.1017/S0376892921000060 Published online by Cambridge University Press