META-ANALYSIS

Island species–area relationships and species accumulation curves are not equivalent: an analysis of habitat island datasets

Thomas J. Matthews1,2*, Kostas A. Triantis1,3, François Rigal2,4, Michael K. Borregaard1,5, François Guilhaumon2,6 and Robert J. Whittaker1,5

1 Conservation Biogeography and Macroecology Programme, School of Geography and the Environment, University of Oxford, Oxford OX1 3QY, UK, 2CE3C – Centre for Ecology, Evolution and Environmental Changes / Azorean Biodiversity Group and Universidade dos Açores, Departamento de Ciências Agrárias, Angra do Heroísmo, Açores 9700-042, Portugal, 3Department of Ecology and Taxonomy, Faculty of Biology, National and Kapodistrian University, Athens GR-15784, Greece, 4Environment and Microbiology Team, Université de Pau et des Pays de l’Amour, Pau Cedex 64013, France, 5Center for Macroecology, Evolution and Climate, Department of Biology, University of Copenhagen, Copenhagen Ø DK-2100, Denmark, 6IRD UMR 9190 MARBEC, IRD-CNRS-IFREMER-UM, Université de Montpellier, Montpellier 34095, France

ABSTRACT

Aim The relationship between species number and area is of fundamental importance in macroecology and conservation science, yet the implications of different means of quantitative depiction of the relationship remain contentious. We set out (1) to establish the variation in form of the relationship between two distinct methods applied to the same habitat island datasets, (2) to explore the relevance of several key dataset properties for variation in the parameters of these relationships, and (3) to assess the implications for application of the resulting models.

Locations Global.

Methods Through literature search we compiled 97 habitat island datasets. For each we analysed the form of the island species–area relationship (ISAR) and several versions of species accumulation curve (SAC), giving priority to a randomized form (Ran-SAC). Having established the validity of the power model, we compared the slopes (z-values) between the ISAR and the SAC for each dataset. We used boosted regression tree and simulation analyses to investigate the effect of nestedness and other variables in driving observed differences in z-values between ISARs and SACs.

Results The Ran-SAC was steeper than the ISAR in 77% of datasets. The differences were primarily driven by the degree of nestedness, although other variables (e.g. the number of islands in a dataset) were also important. The ISAR was often a poor predictor of archipelago species richness.

Main conclusions Slopes of the ISAR and SAC for the same data set can vary substantially, revealing their non-equivalence, with implications for applications of species–area curve parameters in conservation science. For example, the ISAR was a poor predictor of archipelagic richness in datasets with a low degree of nestedness. Caution should be employed when using the ISAR for the purposes of extrapolation and prediction in habitat island systems.

Keywords Boosted regression trees, conservation biogeography, fragmentation, habitat islands, island biogeography, island species–area relationship, macroecology, nestedness, species accumulation curve, species–area relationship.

INTRODUCTION

The increase in number of species with area is one of the few laws of ecology (Scheiner, 2003). While those working with species–area curves have long recognized different forms of curve, confusion in terminology and usage has persisted (for use herein see Table 1), as evidenced by an exchange of papers in this journal disputing the number of fundamental types of...
relationship and the key distinctions between them (Scheiner, 2003, 2004; Gray et al., 2004a,b). In their contributions, Gray et al. (2004a,b) argued that the most important distinction was between species accumulation curves (SACs), which present cumulative counts of increased species number with sampling area, and island-type species–area relationships (herein termed ISARs), in which the function fitted is based on how many species are found in each sampled area (or island). By virtue of their mode of construction, SACs (which can be constructed in several distinct ways; see Appendix S1 in the Supporting Information) must be rising functions, whereas in principle an ISAR can have a negative slope: a large island may have fewer species than a smaller one (Scheiner, 2003), suggesting the possibility that the form of the SAC and ISAR may vary substantially. However, in the island literature, and in its application in conservation science, it is sometimes assumed that the form of the SAC and that of the ISAR are directly comparable (as noted by Ulrich & Buszko, 2007, p. 55). Two classic illustrations of this are to be found in the seminal works of MacArthur & Wilson (1967) and Rosenzweig (1995, 2003).

MacArthur & Wilson (1967), in their Fig. 3, provide two data series—one a cumulative species–area curve of contiguous sampling areas on mainland New Guinea and the other a set of points representing the richness of each island in the region. They comment (p. 10) that ‘[t]he line and the cluster of points illustrate the principle that the increase in number of species with area is more rapid in the case of isolated islands or archipelagos than in expanding sample areas on a single land mass.’ Similarly, Rosenzweig (1995, p. 10) notes, in developing the three scales of species–area relationship model on which his reconciliation ecology (Rosenzweig, 2003) is largely based, that if scattered areas (or islands) are used to construct a species–area curve this will produce a steeper slope (a higher z-value) than if contiguous subplots are used: moreover, the greatest disparity occurs between the richness estimated for the smallest areas, while with increasing area of sample units what we term the SAC and ISAR converge towards the regional richness value. Rosenzweig (1995, p. 19) poses the question as to whether the difference in z-values might be a consequence of the different mode of construction of the two forms of function, but having provided a single demonstration of similarity of values he goes on to state (Rosenzweig, 1995, p. 19) that ‘The data suggest that we do not have a problem here. The question of equivalence of SACs and ISARs is not trivial: if we assume that the different mode of construction of SACs and ISARs is of no great consequence, it follows that we can, for example, base projections of the consequences of habitat loss on comparisons of z-values for contiguous habitat (assessed by SAC models) with z-values for islands or habitat islands (assessed by ISARs). That this might be a dangerous assumption has previously been argued by several authors (e.g. Ulrich & Buszko, 2007; Whittaker & Fernández-Palacios, 2007; Hui, 2008; Dengler, 2009), while the significance of how SACs are constructed for projections of species extinctions has also seen lively recent debate (e.g. He & Hubbell, 2011, 2013; Axelsen et al., 2013), revealing that our understanding of the behaviour of different forms of species–area curve remains incomplete (Whittaker & Matthews, 2014).

Our purpose herein is to advance understanding of the empirical differences between SAC and ISAR data structures by means of a systematic comparison for a large set of habitat island datasets. Specifically, we use a randomization procedure to construct a Scheiner type IIIb curve (termed Ran-SAC; Table 1), and

### Table 1. A glossary of the terms used in this study.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>–</td>
<td>Species–area curve/relationship</td>
<td>Here used as general terms for the relationship between sample area and species richness/number</td>
</tr>
<tr>
<td>ISAR</td>
<td>Island species–area relationship</td>
<td>Whereby the number of species occurring within each of a set of islands is analysed as a function of the area of each island. Equivalent to Scheiner (2003) type IV curves</td>
</tr>
<tr>
<td>SAC</td>
<td>Species accumulation curve</td>
<td>Plots of increasing cumulative species number with increasing sampling effort/area</td>
</tr>
<tr>
<td>SL-SAC</td>
<td>Small–large SAC</td>
<td>The order of island incorporation into the SAC starts with the smallest island and increases up to the largest</td>
</tr>
<tr>
<td>Ran-SAC</td>
<td>Random SAC</td>
<td>The order of island accumulation when constructing the SAC is random</td>
</tr>
<tr>
<td>z</td>
<td>–</td>
<td>The slope parameter of the power species–area relationship model</td>
</tr>
<tr>
<td>zDif</td>
<td>–</td>
<td>The difference between the SAC z-value and the ISAR z-value</td>
</tr>
<tr>
<td>BRT</td>
<td>Boosted regression tree analysis</td>
<td>A regression method which combines a large number of tree models</td>
</tr>
<tr>
<td>Nestedness</td>
<td>–</td>
<td>The situation in which depauperate island faunas constitute proper subsets of the species in richer islands (see Matthews et al., 2015a). In the current study we are interested in species nestedness between isolates</td>
</tr>
<tr>
<td>Anti-nestedness</td>
<td>–</td>
<td>A dataset which is significantly less nested than expected by chance (Matthews et al., 2015a)</td>
</tr>
<tr>
<td>NODF</td>
<td>Nestedness metric based on overlap and decreasing fill</td>
<td>A nestedness metric based on the twin properties of standardized differences in row and column fills and the overlap of presences in two adjacent columns</td>
</tr>
<tr>
<td>ArcRes</td>
<td>Archipelagic residual</td>
<td>The standardized absolute difference between the observed number of species across all islands in a dataset and the number of species predicted by the log-log power model (log transformed)</td>
</tr>
</tbody>
</table>
compare the $z$-values with a Scheiner type IV curve (the ISAR) based on fits using the power model to test the hypothesis that they are non-equivalent. Strictly speaking, our Ran-SAC is a variant of a Scheiner type IIIb curve, because although the Ran-SAC is constructed using mean richness obtained by randomization the observational units are habitat islands of varying area rather than non-contiguous plots of equal size. However, despite this difference, our Ran-SAC is still clearly a SAC, allowing us to compare SAC and ISAR data structures. Next, we explore the relevance of several key dataset properties for variation in the parameters of these relationships, testing the hypothesis that compositional nestedness (see Table 1) will be a significant determinant of variation in $z$ (slope), such that in anti-nested systems (Table 1) the Ran-SAC will be steeper than the ISAR and as the degree of nestedness increases this situation will gradually reverse. While it has long been appreciated that the form of species–area curves from non-contiguous samples is a consequence of both alpha (local) and beta (differentiation) diversity this is, to our knowledge, the first systematic attempt to quantify the interrelationship between ISAR and SAC form and system nestedness across a series of habitat islands.

**MATERIALS AND METHODS**

**Data collection, formatting and species–area curve construction**

Datasets were sourced via a comprehensive search of the literature (Appendix 1) following steps and criteria for evaluating suitability set out in Appendix S2 (see also Matthews et al., 2014a; Matthews, 2015). For each selected dataset we recorded the following: geographical location, taxon, habitat island type (forested or non-forested), latitude and longitude of the study extent (for some datasets this was an estimate as the data were not presented in the source papers), range of species richness and island sizes, and a classification of the predominant matrix type (as per Appendix S2).

For each dataset we constructed (1) the standard ISAR using the raw island area and richness values and (2) several forms of SAC, using different rules for combining islands into the sequence, namely small–large, large–small, poor–rich, rich–poor, random (Table S1). Except where stated, we present results for the smoothed random form of SAC (Ran-SAC) on the grounds that this is now standard practice for constructing such curves (e.g. Ugland et al., 2003) (but see Fig. S1 in Appendix S1 for an example of the variation in form that can be observed using the full set of SAC rules). The Ran-SAC was constructed using a simple bootstrap procedure, randomly selecting the order of island addition into the SAC at each iteration. For example, for the second data point (i.e. two islands) of the SAC of a dataset we randomly selected two islands, noting down the combined area and combined richness of this pair. We repeated this 5000 times, and used the mean of the 5000 $x$- and $y$-values as the data point representing two islands in the analysis.

Our aim was to compare the form of the ISAR with that of Ran-SAC for the same dataset. The simplest way to do this is to compare the $z$-values of the power model ($S = cA^z$) fitted to both sets of data. We focus principally on $z$ as this parameter has attracted far more attention and application than has $c$ (e.g. Rosenzweig, 1995; Tjørve & Tjørve, 2008; Triantis et al., 2012). However, this method is arguably only appropriate if the power model provides a reasonably good fit to both sets of data. Thus, we devised a set of criteria to select suitable datasets for analysis. First, we fitted the power (nonlinear) model to the two different data types for each dataset using nonlinear regression and the ‘mmSAR’ R package (Guilhaumon et al., 2010), and recorded the two parameters ($c$, $z$), $R^2$ and whether the $z$-value was significantly different from zero. A dataset was deemed satisfactory according to this process if the $z$-value of the power model was significant for both the SAC and ISAR structures, and the $R^2$ was $\geq 0.5$ in both cases. This threshold was chosen arbitrarily to eliminate datasets in which the power model explained only a small amount of variation in the SAR. Second, while the observed shape of the power (nonlinear) model is generally convex, we wished to determine whether a model with a different shape provided a better fit to our data. Thus, we fitted a set of 20 ISAR models (including the power model) to the ISAR and SAC data from each dataset using a modified version of the fitting algorithm within the R package mmSAR (see Triantis et al., 2012; Matthews et al., 2015b). We considered the model with the lowest Akaic information criterion corrected for small sample size (AICc) as providing the best fit (Burnham & Anderson, 2002). AICc weights (wAICc; the probability of each model being the best-fitting model given the set of models) were computed for each model fit. Because of the non-nullity constraint on the denominator of the formula, AICc could not be calculated for datasets with fewer than seven islands, and thus these datasets were discarded as were models with inadequate fits (e.g. due to non-normal residuals; cf. Triantis et al., 2012). We then determined the observed shape of the best model fit (convex, sigmoid or linear) according to the algorithm outlined in Triantis et al. (2012) (and see Appendix S3). For each dataset, if the observed shape of either of the best model fits (i.e. to the ISAR and SAC data) was linear or sigmoid (a check was made for linear power model fits and none were observed), the dataset was discarded as the inclusion of $z$-values from these fits may bias our comparisons. In sum, to be classified as satisfactory, both the ISAR and SAC forms of each dataset needed to have: (1) significant $z$-values from the power model fits, (2) $R^2$ values $\geq 0.5$ from the power model fits, and (3) observed convex fits, when considering the best model from a set of twenty competing models.

For these satisfactory datasets, the difference between the Ran-SAC $z$-value and the ISAR $z$-value ($z$Dif; Table 1) was used as our response variable. We also re-ran our analyses using the $z$Dif values from all datasets (i.e. those deemed satisfactory and unsatisfactory).

**Relative influence of the explanatory variables**

To determine the factors that could potentially account for between-dataset differences in $z$Dif, we fitted a boosted regression trees model (BRT; Table 1) using the functions provided in
Elith et al. (2008). In contrast to traditional regression, for which inference is drawn from a single best model, BRT combines a large number of tree models in order to increase predictive performance (Elith et al., 2008). Recent work has shown that the use of weight of evidence values (Burnham & Anderson, 2002) within a linear modelling framework to determine the relative importance of predictor variables is flawed (Galipaud et al., 2014). BRT offers a useful intuitive and robust alternative (see Elith et al., 2008).

In the BRT analysis, zDif was used as the response variable and the dataset characteristics listed above were used as predictor variables. As a further predictor variable, we also calculated the degree of nestedness in each dataset using the NODF metric (Almeida-Neto et al., 2008), implemented in the ‘vegan’ R package (Oksanen et al., 2013). NODF was calculated using the maximally packed matrix. Sites containing zero species were removed prior to calculation, as by definition such sites have zero nestedness and act to depress the NODF value for the full set of sites (see Matthews et al., 2015a). Cross-validation functionality was used to determine the optimum number of trees. We experimented with varying tree complexity, learning rate and bag fraction parameter values in order to minimize the minimum predictive errors, selecting final values of 5 (tree complexity), 0.001 (learning rate; 0.01 when using all datasets) and 0.5 (bag fraction). These values resulted in an optimal number of 7050 fixed trees using only the satisfactory datasets and 4500 trees when all datasets were used. The error structure was considered to be Gaussian. No outliers were removed when running the BRT analysis using the satisfactory datasets, but two outliers had to be removed when using all datasets to induce normality in zDif. The predictive power of the model was assessed through cross-validation, whereby the fitted model was tested on withheld portions of the data (10 subsets were used).

The relative influence of the predictor variables was again calculated using the functions in Elith et al. (2008), which weigh the number of times a predictor is chosen for splitting by the squared improvement to the model due to each split. Partial dependence plots were also used to assess and visualize the effect of a predictor after accounting for the effects of all other model predictors (Elith et al., 2008). In partial dependence plots the y-axis shows the effect of a predictor variable on the response variable after accounting for the effects of the other model predictors (i.e. the marginal effect of the predictor).

Simulation analyses

As the BRT analyses revealed NODF to be the most important variable (see Results), we undertook a simulation analysis to further examine the relevance of nestedness for the difference between the z-values of the two forms of dataset. First, we simulated a set of six islands with area and species richness values conforming roughly to a convex ISAR (area = 1, 5, 10, 15, 20 and 30; containing 10, 12, 14, 16, 17 and 18 species, respectively; herein ‘Sim1’). We fitted the power model (nonlinear) to the ISAR-structured form of Sim1 and recorded the model parameters. We then simulated a perfectly nested (i.e. NODF = 100) presence–absence matrix using the data characteristics of Sim1 (i.e. island number and species richness) and used this matrix to construct the SL (small–large) SAC (see Appendix S1), fitting the power model to the resulting curve. We used the SL-SAC in this analysis because constructing the Ran-SAC for this many matrices would have been computationally intensive. We then altered the presence–absence matrix (without changing the area and richness values) to change the degree of matrix nestedness by randomly shuffling the presences along the sites (i.e. matrix rows) using the ‘commsimulator’ function in the vegan R package and discarding any matrix which had already been simulated. The total number of species in the species pool (i.e. the total number of possible columns in the presence matrix) was set to 40, and we allowed gamma diversity to change between matrices, while the alpha diversity of each island remained constant. This permitted us to create matrices with substantially different levels of nestedness. However, as this meant that the number of columns (total number of species) varied between matrices we used the NODF-by-rows value as our measure of nestedness. For each accepted matrix, we constructed the SAC curve and fitted the power model. We started the random shuffling from different starting points (i.e. different initial presence–absence matrices), used a variety of community simulation algorithms (i.e. R0, R1 and R2) and repeated this process iteratively for 12,000 runs, to cover a wide range of NODF values. We then examined zDif values for each iteration, plotting zDif as a function of nestedness.

Determining the degree of deviation of the archipelagic point

To determine whether the archipelagic point (‘regional richness’) of a dataset deviated from the ISAR of the constituent islands, we followed the method and nomenclature outlined by Santos et al. (2010). We also use ‘archipelagic point’ to refer to the total area and richness of the corresponding set of habitat islands (i.e. the archipelago). For each dataset, we fitted the power (log–log, base 10) SAR model to the ISAR structure and derived the fitted values of the model for each island (‘SIpred’). The log–log version of the power model was used in this analysis as it was the model used by Santos et al. (2010) and thus allowed us to compare our results with theirs. Also following Santos et al. (2010), we added 0.01 to each species richness value to avoid zero values, although we acknowledge that other constants could also be used. We then calculated the archipelagic point as the cumulative total area of all habitat islands in the dataset plotted against the cumulative species richness total (‘SAobs’) and predicted the number of species in the archipelagic point (‘SApred’) using the ISAR model of the constituent islands. We followed Santos et al. (2010) and calculated the archipelagic residual (‘ArcRes’; Table 1) as the absolute difference between log(SAobs) and SApred, standardized by log(SAobs) (see Appendix S3 for details). We noted whether the ISAR over- or under-predicted richness in the archipelagic point for each dataset. This ArcRes methodology represents a simple metric with which to describe how well the ISAR predicts the archipelagic richness.
Differences in the $z$-value were deemed satisfactory for the three additional dataset acceptance criteria, 50 datasets of the ISAR datasets and 9 of the Ran-SAC datasets. Considering only the 50 satisfactory datasets, the Ran-SAC explained a larger amount of variance for the Ran-SAC (mean $z$-value was 0.97) than for the ISAR (mean $z$-value = 0.73), the archipelagic point $z$-value for 32 cases (e.g. see Fig. 1) and was smaller for the remaining 18. The power model was within the set of six best models (i.e. it was ranked as one of the top six models according to AICc) for convex model fits is 0.86, for linear fits 0.09 and for sigmoid fits 0.09; for Ran-SAC the equivalent results were 0.99, 0.00 and 0.02, respectively. The power model was within the set of six best models (i.e. it was ranked as one of the top six models according to AICc) for 67 of the ISAR datasets and 9 of the Ran-SAC datasets. According to our three additional dataset acceptance criteria, 50 datasets were deemed satisfactory for the zDif analyses.

**RESULTS**

We screened over 1000 published articles, of which 97 were deemed suitable for analysis (Appendix 1 and Table S2 in Appendix S2): 69 vertebrate, 20 invertebrate and 8 plant datasets. For the majority of both the ISAR and Ran-SAC datasets the observed best-fitting model shape was convex (for the ISARs the mean $\text{wAIC}_c\text{(AIC weights |waICc; the probability of each model being the best-fitting model given the set of models)}$) for convex model fits is 0.86, for linear fits 0.09 and for sigmoid fits 0.09; for Ran-SAC the equivalent results were 0.99, 0.00 and 0.02, respectively. The power model was within the set of six best models (i.e. it was ranked as one of the top six models according to AICc) for 67 of the ISAR datasets and 9 of the Ran-SAC datasets. According to our three additional dataset acceptance criteria, 50 datasets were deemed satisfactory for the zDif analyses.

**Differences in the $z$-value of ISARs and Ran-SACs**

Considering only the 50 satisfactory datasets, the Ran-SAC $z$-value was greater than the ISAR $z$-value for 32 cases (e.g. see Fig. 1) and was smaller for the remaining 18. The power model explained a larger amount of variance for the Ran-SAC (mean $R^2 = 0.97$) than for the ISAR (mean $R^2 = 0.76$), although this result is easily explained by the smoothing process involved in calculating the Ran-SAC values. Considering all 97 datasets, the Ran-SAC $z$-value was larger in 75 cases. The $z$- and $c$-values are provided in Table S3 in Appendix S4, as are values from the power model fitted to the other SAC data structures (e.g. the SL-SAC).

**Boosted regression tree results**

When zDif values from satisfactory datasets were used as the response variable, NODF was the most important explanatory variable (relative influence value of 24.6; see Table 2). Both the minimum number of species and the number of islands also had relatively high influence values (Table 2, Fig. 2a, b). In terms of the overall predictive performance of the model, the mean coefficient of the correlation between the observed and predicted response values was 0.63 when based on the cross-validation data. Results were qualitatively similar when considering zDif values from all datasets (mean correlation coefficient 0.68).

As predicted, further analysis revealed that in highly nested systems the $z$-values for ISARs were greater than those for Ran-SACs, and the reverse for systems with little nesting (see Figs 2a & 3a). In addition, the $z$-value of the Ran-SAC was significantly related to the NODF value (Fig. 3b), whereas there was no relationship between the $z$-value of the ISAR and NODF (Fig. 3c).

**Nestedness simulations**

The simulation analyses revealed further evidence for the importance of nestedness in explaining the difference in $z$-value between ISARs and SACs. In the case of Sim1, when the system was anti-nested (low NODF values), the $z$-value of the SL-SAC was greater than that of the ISAR. As the degree of nestedness increased, the difference in $z$-values declined until the ISAR $z$-value exceeded that of the SL-SAC (Fig. 4a). Figure 4(b) illustrates this effect for two SL-SAC curves constructed from the most nested and anti-nested Sim1 iterations, respectively (we used the SL-SAC curve type as the area range of the ISAR and SAC are similar, making the plot easier to interpret).

**The fit of the archipelagic data point**

When considering only datasets in which the power (log–log) model provided a significant fit ($n = 73$), the archipelagic point deviated substantially (following the rule of thumb used by...
curves using a large compilation of habitat island datasets. The
analytical properties, we compared the analyses and syntheses, and to understand their differing emergent properties, we compared the z-values of ISAR and SACs using a large compilation of habitat island datasets. The

<table>
<thead>
<tr>
<th>Variable</th>
<th>Relative influence</th>
<th>Variable</th>
<th>Relative influence</th>
</tr>
</thead>
<tbody>
<tr>
<td>NODF</td>
<td>24.61</td>
<td>NODF</td>
<td>30.11</td>
</tr>
<tr>
<td>Min. sp.</td>
<td>17.94</td>
<td>No. sp.</td>
<td>16.62</td>
</tr>
<tr>
<td>No. Isl.</td>
<td>17.12</td>
<td>Min. sp.</td>
<td>13.60</td>
</tr>
<tr>
<td>Ar. ratio</td>
<td>13.15</td>
<td>Ar. ratio</td>
<td>8.00</td>
</tr>
<tr>
<td>Max. Ar.</td>
<td>9.80</td>
<td>Latitude</td>
<td>7.15</td>
</tr>
<tr>
<td>Longitude</td>
<td>4.53</td>
<td>Longitude</td>
<td>6.33</td>
</tr>
<tr>
<td>Latitude</td>
<td>4.52</td>
<td>Max. Ar.</td>
<td>5.31</td>
</tr>
<tr>
<td>No. sp.</td>
<td>2.78</td>
<td>Max. Sp.</td>
<td>3.44</td>
</tr>
<tr>
<td>Max. sp.</td>
<td>2.14</td>
<td>No. Isl.</td>
<td>3.40</td>
</tr>
<tr>
<td>Matrix</td>
<td>1.43</td>
<td>Taxon</td>
<td>2.20</td>
</tr>
<tr>
<td>Min. Ar.</td>
<td>0.99</td>
<td>Min. Ar.</td>
<td>2.06</td>
</tr>
<tr>
<td>Isl. type</td>
<td>0.88</td>
<td>Isl. type</td>
<td>0.92</td>
</tr>
<tr>
<td>Taxon</td>
<td>0.11</td>
<td>Matrix</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Dif, difference between the z-value of the species accumulation curve and the z-value of the island species–area relationship (ISAR), for each dataset; ArcRes, standardized difference between the cumulative species richness total of the archipelago and the predicted number of species in the archipelagic point using the ISAR model of the constituent islands; Min., minimum; Max., maximum; No., number; Ar., area; Isl., island, Sp., species.

Santos et al., 2010) from the prediction of the constituent ISAR in 14 datasets (see Fig. 5 for an example of both scenarios). The ISAR under-predicted the archipelagic richness in 45 cases (Table S3 in Appendix S4). Considering 64 datasets (the 75 significant fits minus the 9 outliers), when ArcRes was used as a response variable in a BRT analysis, NODF was the variable with the highest relative influence value (Table 2). The number of species and the minimum number of species in a dataset were also important variables (Table 2). The mean coefficient of the correlation between the observed and predicted response values was 0.35.

**DISCUSSION**

To evaluate the impact of confounding ISAR and SACs in SAR analyses and syntheses, and to understand their differing emergent properties, we compared the z-values of ISAR and SAC curves using a large compilation of habitat island datasets. The

construction of SACs from island data has been attempted previously (e.g. Quinn & Harrison, 1988; Rosenzweig, 1995; Fischer & Lindenmayer, 2002; Fattorini, 2010), although these earlier studies have largely focused on small–large and large–small SACs. We decided to focus on a Ran-SAC as it represents an average of the various SAC construction permutations (see Appendix S1) and thus seems a sensible choice if only one type of SAC is to be used in comparative analyses. As noted by previous authors (e.g. Ulrich & Buszko, 2007), critical tests of how different types of SAC affect model shape and parameters are lacking. We found that the z-values and form of the ISAR and SAC curves varied considerably within datasets, and that this variation could primarily be explained as a function of the degree of nestedness in the archipelago. We also observed that the ISAR was often a poor predictor of SAbos, demonstrating the dangers of extrapolating the ISAR, for example, to predict the richness of an archipelago of fragments.

**Why are SACs generally steeper than ISARs?**

In accordance with our first hypothesis we found considerable differences between the z-values of ISARs and SACs (e.g. Figs 1 & S1). Hence, our results illustrate that different inferences might be drawn depending on which data structure is used in an analysis (and see Ulrich & Buszko, 2007, for an analysis based on a different SAC type). Consistent with our second hypothesis, the difference in the z-values of SACs and ISARs is best explained by variation in compositional nestedness, a pattern which is to be expected as a function of the procedures used in constructing the two types of curves (Gray et al., 2004a,b); a fact highlighted by both empirical and simulation analyses (Figs 2 & 4).

A number of factors have been argued to underpin the z-value of SACs, through an effect on species overlap and/or nestedness (e.g. Quinn & Harrison, 1988; Chase & Knight, 2013). Two of these factors are particularly relevant to habitat island systems. First, a large species pool means that, all else being equal, there is a smaller probability that the same species will succeed in occupying each island, and thus species overlap between islands is reduced. Furthermore, if islands within an archipelago (of real or habitat islands) draw species from different species pools the likelihood is that species overlap across all islands will be reduced (Whittaker & Fernández-Palacios, 2007). Second, habitat islands are generally disturbed systems, and disturbance usually has a disproportionate effect in smaller islands (Whittaker & Fernández-Palacios, 2007). This means that matrix and generalist species are relatively over-represented in smaller fragments (Matthews et al., 2014b), again resulting in reduced overlap. In previous work we have shown that the relatively high incidence of generalist species in smaller fragments reduces the slope of a multimodel ISAR curve for a number of habitat island datasets (Matthews et al., 2014b). This reduction in ISAR slope coupled with the reduced species overlap across islands provides one explanation for the many positive zDif values (i.e. Ran-SAC steeper than ISAR) observed.

Our analyses indicated that other variables are also important, independent of nestedness (Table 2), with the minimum
number of species (Min.) on an island in a dataset being the second most influential variable explaining variation in \( z \)Dif (Table 2). The partial dependence plot of this variable (Fig. 2b) indicates that low Min. values have a negative effect on \( z \)Dif; in this case, in many datasets with low Min. values the ISAR is steeper than the Ran-SAC (i.e. \( z \)Dif is negative).

Interestingly, the maximum number of species and the total number of species in the archipelago had low values of relative influence (Table 2). It is possible that this finding could be due to an indirect effect of additional variables not included in our analyses that are correlated with Min. Further research is needed to fully explore the implications of these results.

Landscape context variables (e.g. the habitat matrix and the island type) were relatively unimportant (Table 2). This is surprising as the habitat matrix is considered to be important in determining differences in \( z \)-value between ISARs and SACs due to the influence of matrix properties on species turnover (see Crist & Veech, 2006). The lack of an effect of these variables is probably due, at least in part, to the coarse matrix and island-type classifications used in our analyses.

**Interpreting variation in ArcRes**

Using the simple descriptive metric adopted by Santos et al. (2010) we found that the archipelagic point deviated substantially from the ISAR prediction in 19% of datasets with a significant ISAR, compared with 12% in Santos et al.’s analyses of true island datasets. Consistent with our findings for slope differences (i.e. \( z \)Dif; see above) NODF had the largest effect on ArcRes (Table 2), while the ISAR more frequently under-predicted than over-predicted the richness of the archipelagic point (Fig. 5a). A role for nestedness in under-prediction of system richness may reflect the fact that habitat islands often
contain a large number of singletons (i.e. species that are only sampled in one habitat island in a dataset and whose presence reduces the nestedness of the full data matrix) as a result of factors such as source–sink dynamics and transient species that may be using suitable, but unsampled, patches of habitat within the study area, thus enabling their persistence despite low frequencies in the dataset. A recent meta-analysis using many of the same datasets has shown that, contrary to earlier work, the majority of habitat island systems described in the literature have low levels of nestedness, and indeed that significant anti-nestedness is more common than significant nestedness (Matthews et al., 2015a). Taken together, these findings bring into question the extrapolation of ISARs (see also Fig. 1b) and certain other species–area curves of similar or composite

Implications and conclusions

The choice of how to construct the species–area curve and the method of analysis are important considerations when using the SAR for applied purposes (Whittaker et al., 2005; Ulrich & Buszko, 2007; Halley et al., 2013; Matthews et al., 2015b), but conservation biogeographical studies are often unclear about the type of SAR employed and, indeed, terminology is inconsistent and disputed across the literature (see e.g. Scheiner, 2003, 2004; Gray et al., 2004a,b; Whittaker & Fernández-Palacios, 2007). Our analyses have provided empirical proof that ISAR and SAC are not equivalent and may differ substantially in fitted parameters (contrary to assumptions in, e.g., MacArthur & Wilson, 1967, p10; Rosenzweig, 1995, p. 19). This distinction has not always been recognized, and greater attention should therefore be paid to explicitly identifying the type of species–area curve and method of fit used in future ecological analyses (see also Scheiner, 2003; Ulrich & Buszko, 2007). The present analysis does not argue in favour of one specific species–area function being universally preferable; rather, the choice depends on the aim of study and the data available. For example, SACs are arguably more useful than ISARs for assessing the contribution of different sized islands to a protected area network (Fischer & Lindenmayer, 2002; Watson et al., 2009).

Based on the self-similar scaling properties of the power model, Tjørve & Tjørve (2008) (see also Hart et al., 1999) used a form of SAC to show mathematically that the proportional overlap between two areas of the same size can be plotted as a function of z: as z increases, the degree of species overlap decreases. For example, a z-value of 0.58 equates to 50% of the species being shared between the two areas (Tjørve & Tjørve, 2008). Our results provide an empirical illustration of this point (e.g. Fig. 3b) and indicate that: (1) the SAC is often steeper than the ISAR when the curves are constructed using habitat island data, and (2) the z-value of the ISAR only becomes consistently greater than the z-value of the SAC when the data are highly nested, i.e. the NODF value is approximately 70 or above (e.g. Figs 3a & 4b).

It is also evident that for a number of habitat island systems the ISAR is a poor predictor of the overall number of species in an archipelago (e.g. Fig. 5a). Thus caution should be employed when using the ISAR for extrapolation purposes in fragmented systems, particularly when species overlap is thought to be low (Crist & Veech, 2006). The SAC is likely to provide more accurate results in such contexts as it incorporates information on the degree of nestedness/overlap in the region (Quinn & Harrison, 1988; Hui, 2008). It is already acknowledged that the choice of sampling design is an important consideration in SAR studies, and our results provide more evidence for those who have recently argued for a more consistent SAR theory applicable to general SAR sampling designs (e.g. Whittaker et al., 2005; He & Hubbell, 2013).

ACKNOWLEDGEMENTS

Joe Veech, the journal reviewers and editors, provided valuable comments and criticisms, for which we are grateful: any errors and misconceptions are ours to claim. M.K.B. was supported by...
a personal Sapere Aude grant from the Danish Councils for Independent Research. This work was partially supported by the Fundação para a Ciência e a Tecnologia (FCT) project ‘Biodiversity on oceanic islands: towards a unified theory’ (PTDC/BIA-BIC/119255/2010).

REFERENCES


APPENDIX 1 DATA SOURCES


SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** The different types of species–area relationship.

**Appendix S2** Source paper information.

**Appendix S3** Expanded study methods.

**Appendix S4** Supplementary results.

**BIOSKETCHES**

**Tom Matthews** is a research associate at the University of Oxford. He is interested in the study of macroecological patterns in habitat islands, and the application of island theory to fragmented landscapes.

Author contributions: T.J.M., K.A.T. and R.J.W. conceived the ideas; T.J.M. collected the data; T.J.M. ran the analyses; and T.J.M. and R.J.W. wrote the paper with a significant contribution from K.A.T., M.K.B., E.F. and F.G. All authors discussed the results and commented on the manuscript.

Editor: Adriana Ruggiero

Global Ecology and Biogeography, © 2016 John Wiley & Sons Ltd