META-ANALYSIS

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

Thomas J. Matthews[1,2]*, Kostas A. Triantis[1,3], François Rigal[2,4], Michael K. Borregaard[1,5], François Guilhaumon[2,6] and Robert J. Whittaker[1,5]

[1]*Conservation Biogeography and Macroecology Programme, School of Geography and the Environment, University of Oxford, Oxford OX1 3QY, UK,* [2]*CE3C – 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,* [3]*Department of Ecology and Taxonomy, Faculty of Biology, National and Kapodistrian University, Athens GR-15784, Greece,* [4]*Environment and Microbiology Team, Université de Pau et des Pays de l'Amour, Pau Cedex 64013, France,* [5]*Center for Macroecology, Evolution and Climate, Department of Biology, University of Copenhagen, Copenhagen Ø DK-2100, Denmark,* [6]*IRD UMR 9190 MARBEC, IRD-CNRS-IFREMER-UM, Université de Montpellier, Montpellier 34095, France*

*Correspondence: Thomas J. Matthews, Conservation Biogeography and Macroecology Programme, School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, UK.
E-mail: thomas.matthews@ouce.ox.ac.uk

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

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

Abbreviation	Full term	Definition
–	Species–area curve/relationship	Here used as general terms for the relationship between sample area and species richness/number
ISAR	Island species–area relationship	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
SAC	Species accumulation curve	Plots of increasing cumulative species number with increasing sampling effort/area
SL-SAC	Small–large SAC	The order of island incorporation into the SAC starts with the smallest island and increases up to the largest
Ran-SAC	Random SAC	The order of island accumulation when constructing the SAC is random
z	–	The slope parameter of the power species–area relationship model
zDif	–	The difference between the SAC z-value and the ISAR z-value
BRT	Boosted regression tree analysis	A regression method which combines a large number of tree models
Nestedness	–	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
Anti-nestedness	–	A dataset which is significantly less nested than expected by chance (Matthews *et al.*, 2015a)
NODF	Nestedness metric based on overlap and decreasing fill	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
ArcRes	Archipelagic residual	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)

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

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 ≥ 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 Akaike information criterion corrected for small sample size (AIC$_c$) as providing the best fit (Burnham & Anderson, 2002). AIC$_c$ 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, AIC$_c$ 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 ≥ 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 (zDif; Table 1) was used as our response variable. We also re-ran our analyses using the zDif 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 zDif, 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.

To determine if any dataset characteristics (above) could explain variation in ArcRes between datasets we repeated our BRT analyses (learning rate of 0.01) using ArcRes as the response variable. We only used ArcRes values from datasets in which the power (log–log) provided a significant fit (cf. Santos *et al.*, 2010). Nine outlying points had to be removed to induce normality, as there was a long left-hand tail in the distribution of ArcRes values. The optimal number of fixed trees was 2200. All analyses and simulations were conducted in R (v.3.1.1.; R Development Core Team, 2014). A significance level of 0.05 was employed in all 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 wAIC$_c$ (AIC$_c$ weights [wAIC$_c$; 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 AIC$_c$) 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 *z*Dif 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 *z*Dif 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 *z*Dif 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-
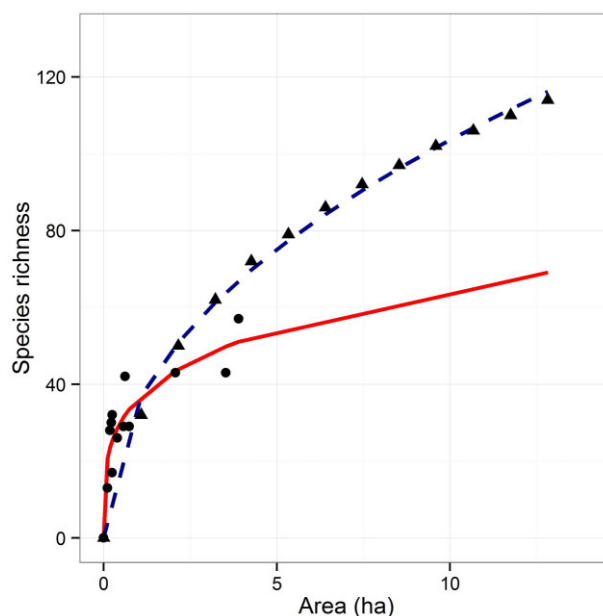

Figure 1 Power model fits constructed from the same habitat island dataset: Gavish *et al.* (2012) (invertebrates in a study system from Israel; 12 islands and 114 species). The fits of the power model (nonlinear) to the island species–area relationship (ISAR) data (solid circles; the solid line is the model fit) and the randomized species accumulation curve (Ran-SAC) data (solid triangles; the dashed line is the model fit) are shown. For the ISAR model fit, the power model has been extrapolated to the total cumulative area of all habitat islands in the system using the parameters derived from the model fitting process. The *z*-values of the models are 0.25 for the ISAR model and 0.46 for the SAC model.

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 was 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

Table 2 The relative contributions (%) of predictor variables for boosted regression tree models developed using cross-validation. The model was fitted using two different response variables: *z*Dif (number of datasets = 50; 7050 fixed trees) and ArcRes (number of datasets = 66; 2300 fixed trees). A set of dataset characteristics were used as the predictor variables (see Materials and Methods). The nestedness of the archipelago was measured using the NODF (nestedness metric based on overlap and decreasing fill) metric (maximally packed matrix). Island type was a categorical variable indicating whether a dataset was a forested or a non-forest island, and taxon was a categorical variable indicating whether a dataset was a plant, vertebrate or invertebrate dataset. Matrix was a variable indicating the permeability of the surrounding matrix type. Variables are ordered by their relative influence in each case.

*z*Dif		ArcRes	
Variable	Relative influence	Variable	Relative influence
NODF	24.61	NODF	30.11
Min. sp.	17.94	No. sp.	16.62
No. Isl.	17.12	Min. sp.	13.60
Ar. ratio	13.15	Ar. ratio	8.00
Max. Ar.	9.80	Latitude	7.15
Longitude	4.53	Longitude	6.33
Latitude	4.52	Max. Ar.	5.31
No. sp.	2.78	Max Sp.	3.44
Max. sp.	2.14	No. Isl.	3.40
Matrix	1.43	Taxon	2.20
Min. Ar.	0.99	Min. Ar.	2.06
Isl. type	0.88	Isl. type	0.92
Taxon	0.11	Matrix	0.86

*z*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 the 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 Ran-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 SAR 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 SAobs, 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 *z*Dif 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
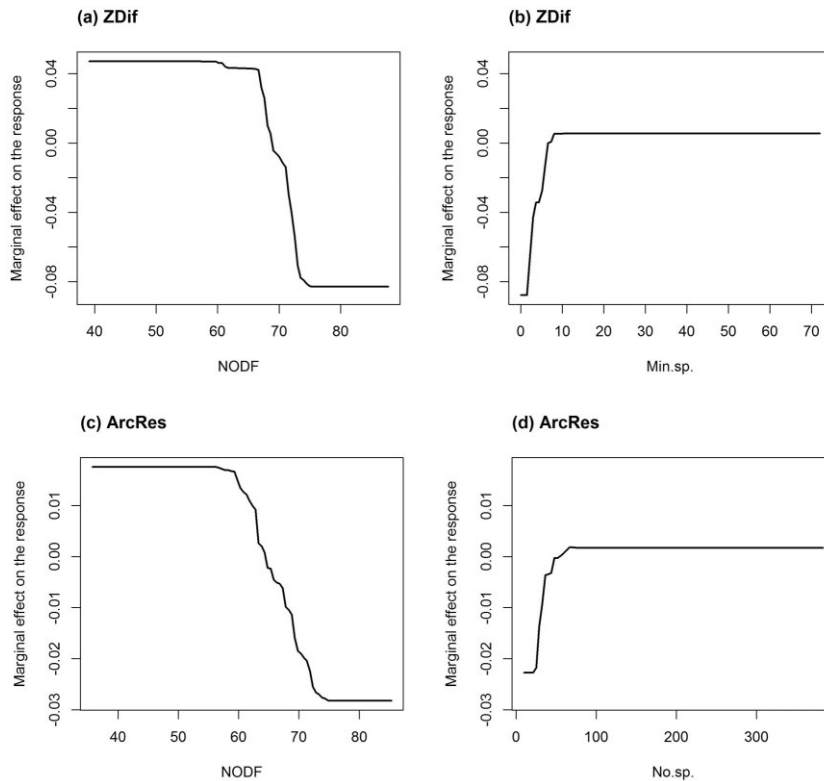
Figure 2 Partial dependence plots for two of the most influential variables in two boosted regression tree (BRT) models. The response variable was zDif (a, b) [the difference between the z-value of the species accumulation curve and the z-value of the island species–area relationship (ISAR) for each dataset], and in the second model the response was the the archipelagic residual (ArcRes) (c, d) (the standardized difference between the cumulative species richness total of the archipelago and the predicted the number of species in the archipelagic point using the ISAR model of the constituent islands). Only zDif values from satisfactory datasets were used ($n = 50$) in (a) and (b), and only ArcRes values from datasets in which the fit of the power (log–log) model was significant ($n = 73$, reduced to 64 after removing nine outliers) were used in (c) and (d). The two predictor variables in the first model are (a) nestedness (NODF; for the relative contributions of each variable see Table 2) and (b) the minimum number of species in a dataset (Min. sp.). The two predictor variables in the second model are (c) nestedness (NODF), and (d) the number of species in a dataset (No. sp.). The y-axis shows the effect of a predictor variable (x-axis) on the response variable after accounting for the effects of the other model predictors (i.e. the marginal effect of the predictor). The y-axis represents the effect of the predictor on the response, and is not an indicator of the value of the response at a given value of the predictor. A positive y-value indicates that at the given x-value the effect (based on the model) on the y-value (the response) is positive, and vice versa. The ends of the curves represent areas of the plotting space with fewer data points and are relatively uninformative.

number of species (Min.) on an island in a dataset being the second most influential variable explaining variation in zDif (Table 2). The partial dependence plot of this variable (Fig. 2b) indicates that low Min. values have a negative effect on zDif; in this case, in many datasets with low Min. values the ISAR is steeper than the Ran-SAC (i.e. zDif 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. zDif; see above) NODF had the largest effect on ArcRes (Table 2), while the ISAR more frequently underpredicted 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
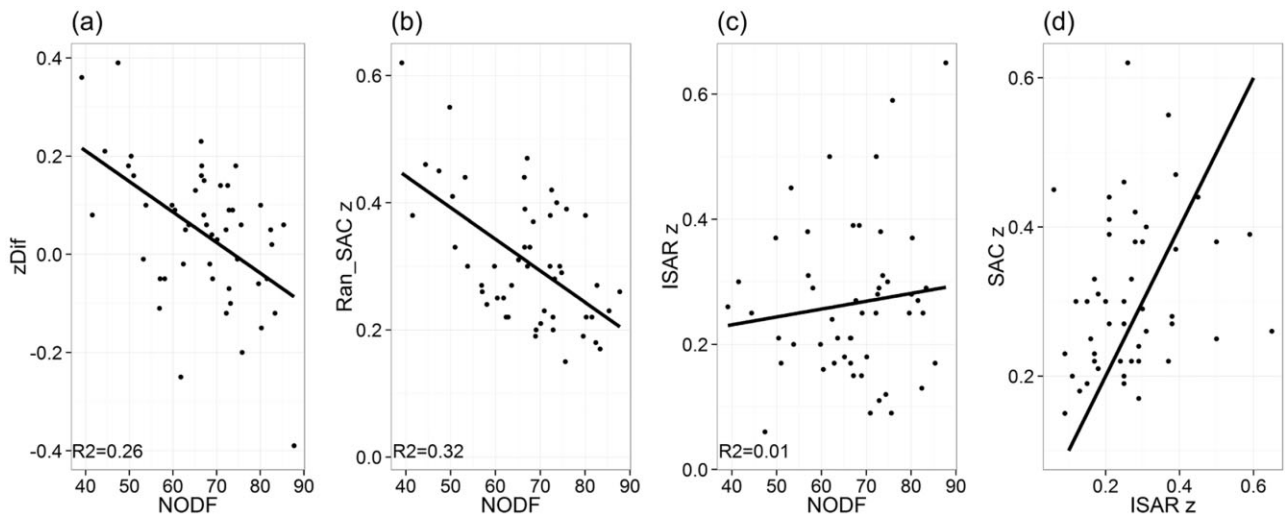
Figure 3 The relationship between nestedness (measured by the NODF metric) and (a) zDif [the difference between the randomized species accumulation curve (Ran-SAC) *z*-value and the island species–area relationship (ISAR) *z*-value], (b) the SAC *z*-value, and (c) the ISAR *z*-value. Only *z*-values from the satisfactory datasets ($n = 50$) were used to construct the plots. In each plot, the solid line represents the best line of a linear regression model, and the R^2 of this fit is given on each plot. The *z*-value was significant in each plot apart from (c), which is shown for illustrative purposes only. An increasing NODF value indicates an increasing degree of nestedness, according to this metric. (d) The *z*-value of the power model fitted using ISAR structured data plotted against *z*-values derived using randomly constructed SAC data. The solid line in (d) represents the 1:1 fit line; points below the line represent datasets in which the ISAR *z*-value was larger than the SAC *z*-value, and vice versa. The *z*-values in all plots relate to the nonlinear power SAR model.
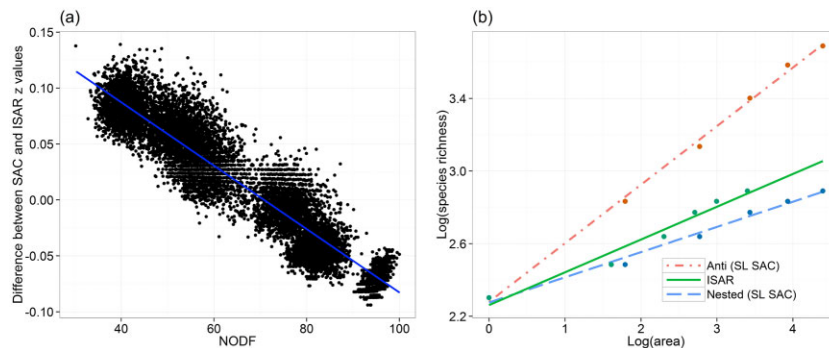

Figure 4 The relationship between nestedness and the variation in *z* values of ISARs and small-large (SL; the order of island incorporation into the SAC starts with the smallest island and increases up to the largest; see Appendix S1) constructed SACs. In (a) a perfectly nested presence–absence matrix for a set of six islands of varying area was simulated (area of each island = 1, 5, 10, 15, 20 and 30; species richness of each island = 10, 12, 14, 16, 17 and 18). First, the power model (nonlinear) was fitted to the data matrix in ISAR form, and the *z*-value recorded. The SL-SAC was then constructed using the same data matrix, and again the power model was fitted. The presence–absence matrix was then rearranged to change the level of nestedness, with the constraint that the species richness of each island was kept constant (i.e. the ISAR remained unchanged), although the overall number of species in the archipelago was allowed to vary; the power model was then fitted to both the ISAR and SAC constructed using this new data matrix. This process was repeated iteratively along a gradient of NODF (by rows, i.e. sites) values. A total of 12,000 matrix permutations were simulated. The best fit linear regression line (solid line) through these points (dots) is also shown. (b) The power (log–log) model, fitted to ISAR structured data (solid line, $z = 0.18$), and to SL-SAC structured data using the same simulated island data as for (a). The fits of SL-SAC curves to two matrix permutations are shown: a perfectly nested set of isolates (dashed line, $z = 0.14$), and a perfectly anti-nested set of isolates (dashed and dotted line, $z = 0.32$).

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
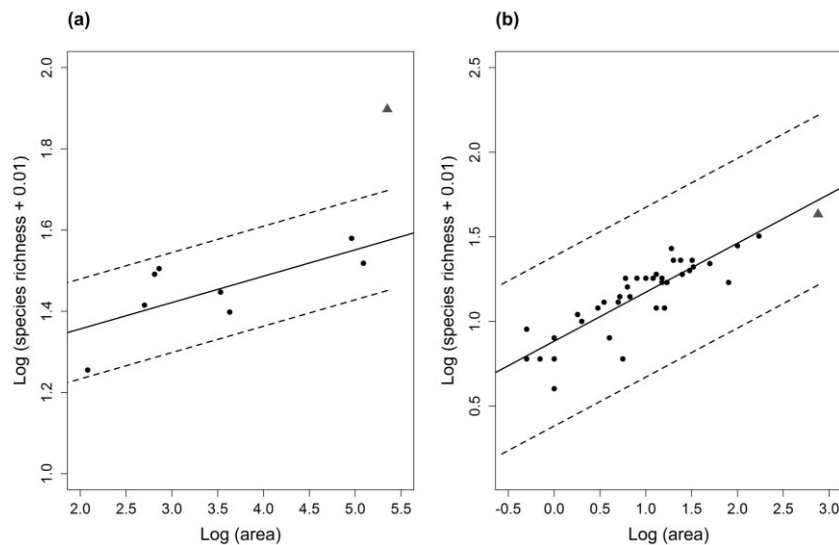
Figure 5 The island species–area relationship (ISAR; power log–log model) for two habitat island datasets: (a) Benedick *et al.* (2006) (invertebrates in Malaysia; number of islands = 8), and (b) this study (birds in the UK; number of islands = 37). The dots represent individual habitat islands and the fit of the power (log–log) model is indicated by the solid line. In both plots the archipelagic point, calculated as the total cumulative area and species richness of all habitat islands in the dataset, is illustrated as a triangle. The intervals defined by the maximum residual criterion are shown as the dashed line: this criterion relates only to the difference between the ISAR prediction and the archipelagic point, the dashed lines have then just been continued down the *y*-axis. If the archipelagic point lies outside these bounds (e.g. as in part (a)) it is said to deviate from the prediction of the ISAR calculated using the constituent habitat islands; and vice versa (e.g. as in part (b)); cf. Santos *et al.*, 2010).

construction: a common application of the ISAR (cf. Whittaker *et al.*, 2005; Ulrich & Buszko, 2007; Hui, 2008).

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 Harte *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

REFERENCES

Almeida-Neto, M., Guimarães, P., Guimarães, P.R., Loyola, R.D. & Ulrich, W. (2008) A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos*, **117**, 1227–1239.

Axelsen, J.B., Roll, U., Stone, L. & Solow, A. (2013) Species–area relationships always overestimate extinction rates from habitat loss: comment. *Ecology*, **94**, 761–763.

Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multi-model inference: a practical information-theoretic approach*, 2nd edn. Springer, New York.

Chase, J.M. & Knight, T.M. (2013) Scale-dependent effect sizes of ecological drivers on biodiversity: why standardised sampling is not enough. *Ecology Letters*, **16**, 17–26.

Crist, T.O. & Veech, J.A. (2006) Additive partitioning of rarefaction curves and species–area relationships: unifying α-, β- and γ-diversity with sample size and habitat area. *Ecology Letters*, **9**, 923–932.

Dengler, J. (2009) Which function describes the species–area relationship best? A review and empirical evaluation. *Journal of Biogeography*, **36**, 728–744.

Elith, J., Leathwick, J.R. & Hastie, T. (2008) A working guide to boosted regression trees. *Journal of Animal Ecology*, **77**, 802–813.

Fattorini, S. (2010) The use of cumulative area curves in biological conservation: A cautionary note. *Acta Oecologia*, **36**, 255–258.

Fischer, J. & Lindenmayer, D.B. (2002) Small patches can be valuable for biodiversity conservation: two case studies on birds in southeastern Australia. *Biological Conservation*, **106**, 129–136.

Galipaud, M., Gillingham, M.A.F., David, M. & Dechaume-Moncharmont, F.-X. (2014) Ecologists overestimate the importance of predictor variables in model averaging: a plea for cautious interpretations. *Methods in Ecology and Evolution*, **5**, 983–991.

Gavish, Y., Ziv, Y. & Rosenzweig, M.L. (2012) Decoupling fragmentation from habitat loss for spiders in patchy agricultural landscapes. *Conservation Biology*, **26**, 150–159.

Gray, J.S., Ugland, K.I. & Lambshead, J. (2004a) Species accumulation and species–area curves – a comment on Scheiner (2003). *Global Ecology and Biogeography*, **13**, 473–476.

Gray, J.S., Ugland, K.I. & Lambshead, J. (2004b) On species accumulation and species–area curves. *Global Ecology and Biogeography*, **13**, 567–568.

Guilhaumon, F., Mouillot, D. & Gimenez, O. (2010) mmSAR: an R-package for multimodel species–area relationship inference. *Ecography*, **33**, 420–424.

Halley, J.M., Sgardeli, V. & Monokrousos, N. (2013) Species–area relationships and extinction forecasts. *Annals of the New York Academy of Sciences*, **1286**, 50–61.

Harte, J., Kinzig, A. & Green, J. (1999) Self-similarity in the distribution and abundance of species. *Science*, **284**, 334–336.

He, F. & Hubbell, S.P. (2011) Species–area relationships always overestimate extinction rates from habitat loss. *Nature*, **473**, 368–371.

He, F. & Hubbell, S.P. (2013) Estimating extinction from species–area relationships: why the numbers do not add up. *Ecology*, **94**, 1905–1912.

Hui, C. (2008) On species–area and species accumulation curves: a comment on Chong and Stohlgren's index. *Ecological Indicators*, **8**, 327–329.

MacArthur, R.H. & Wilson, E.O. (1967) *The theory of island biogeography*. Princeton University Press, Princeton, NJ.

Matthews, T.J. (2015) Analysing and modelling the impact of habitat fragmentation on species diversity: a macroecological perspective. *Frontiers of Biogeography*, **7**, 60–68.

Matthews, T.J., Steinbauer, M.J., Tzirkalli, E., Triantis, K.A. & Whittaker, R.J. (2014a) Thresholds and the species–area relationship: a synthetic analysis of habitat island datasets. *Journal of Biogeography*, **41**, 1018–1028.

Matthews, T.J., Cottee-Jones, H.E.W. & Whittaker, R.J. (2014b) Habitat fragmentation and the species–area relationship: a focus on total species richness obscures the impact of habitat loss on habitat specialists. *Diversity and Distributions*, **20**, 1136–1146.

Matthews, T.J., Cottee-Jones, H.E.W. & Whittaker, R.J. (2015a) Quantifying and interpreting nestedness in habitat islands: a synthetic analysis of multiple datasets. *Diversity and Distributions*, **21**, 392–404.

Matthews, T.J., Guilhaumon, F., Triantis, K.A., Borregaard, M.K. & Whittaker, R.J. (2015b) On the form of species–area relationships in habitat islands and true islands. *Global Ecology and Biogeography*, doi:10.1111/geb.12269.

Oksanen, J.F., Blanchet, G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B. & Simpson, G.L. (2013) *Package 'vegan'. R package version 2.0-9*. Available at: http://cran.r-project.org/

Quinn, J. & Harrison, S. (1988) Effects of habitat fragmentation and isolation on species richness: evidence from biogeographic patterns. *Oecologia*, **75**, 132–140.

R Development Core Team (2014) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.r-project.org/ (accessed 1 August 2014).

Rosenzweig, M.L. (1995) *Species diversity in space and time*. Cambridge University Press, Cambridge.

Rosenzweig, M.L. (2003) Reconciliation ecology and the future of species diversity. *Oryx*, **37**, 194–205.

Santos, A.M.C., Whittaker, R.J., Triantis, K.A., Borges, P.A.V., Jones, O.R., Quicke, D.L.J. & Hortal, J. (2010) Are species–area relationships from entire archipelagos congruent with those of their constituent islands? *Global Ecology and Biogeography*, **19**, 527–540.

Scheiner, S.M. (2003) Six types of species–area curves. *Global Ecology and Biogeography*, **12**, 441–447.

Scheiner, S.M. (2004) A mélange of curves – further dialogue about species–area relationships. *Global Ecology and Biogeography*, **13**, 479–484.

Tjørve, E. & Tjørve, K.M.C. (2008) The species–area relationship, self-similarity, and the true meaning of the *z*-value. *Ecology*, **89**, 3528–3533.

Triantis, K.A., Guilhaumon, F. & Whittaker, R.J. (2012) The island species–area relationship: biology and statistics. *Journal of Biogeography*, **39**, 215–231.

Ugland, K.I., Gray, J.S. & Ellingsen, K.E. (2003) The species–accumulation curve and estimation of species richness. *Journal of Animal Ecology*, **72**, 888–897.

Ulrich, W. & Buszko, J. (2007) Sampling design and the shape of species–area curves on the regional scale. *Acta Oecologia*, **31**, 54–59.

Watson, J.E.M., Watson, A.W.T., Fischer, J., Ingram, J.C. & Whittaker, R.J. (2009) Using nestedness and species–accumulation analyses to strengthen a conservation plan for littoral forest birds in south-eastern Madagascar. *International Journal of Biodiversity and Conservation*, **1**, 67–80.

Whittaker, R.J. & Fernández-Palacios, J.M. (2007) *Island biogeography: ecology, evolution, and conservation*, 2nd edn. Oxford University Press, Oxford.

Whittaker, R.J. & Matthews, T.J. (2014) The varied form of species–area relationships. *Journal of Biogeography*, **41**, 209–210.

Whittaker, R.J., Araújo, M.B., Jepson, P., Ladle, R.J., Watson, J.E.M. & Willis, K.J. (2005) Conservation biogeography: assessment and prospect. *Diversity and Distributions*, **11**, 3–23.

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., F.R. and F.G. All authors discussed the results and commented on the manuscript.

Editor: Adriana Ruggiero

APPENDIX 1 DATA SOURCES

dos Anjos, L. & Boçon, R. (1999) Bird communities in natural forest patches in southern Brazil. *The Wilson Bulletin*, **111**, 397–414.

Báldi, A. & Kisbenedek, T. (1999) Orthopterans in small steppe patches: an investigation for the best-fit model of the species–area curve and evidences for their non-random distribution in the patches. *Acta Oecologica*, **20**, 125–132.

Behle, W.H. (1978) Avian biogeography of the Great Basin and intermountain region. *Great Basin Naturalist Memoirs*, **2**, 55–80.

Benedick, S., Hill, J.K., Mustaffa, N. et al. (2006) Impacts of rain forest fragmentation on butterflies in northern Borneo: species richness, turnover and the value of small fragments. *Journal of Applied Ecology*, **43**, 967–977.

Blake, J.G. & Karr, J.R. (1984) Species composition of bird communities and the conservation benefit of large versus small forests. *Biological Conservation*, **30**, 173–187.

Brotons, L. & Herrando, S. (2001) Reduced bird occurrence in pine forest fragments associated with road proximity in a Mediterranean agricultural area. *Landscape and Urban Planning*, **57**, 77–89.

Brown, J.H. (1971) Mammals on mountaintops: nonequilibrium insular biogeography. *The American Naturalist*, **105**, 467–478.

Brown, J.H. (1978) The theory of island biogeography and the distribution of boreal birds and mammals. *Great Basin Naturalist Memoirs*, **2**, 209–227.

Cabrera-Guzmán, E. & Reynoso, V. (2012) Amphibian and reptile communities of rainforest fragments: minimum patch size to support high richness and abundance. *Biodiversity and Conservation*, **21**, 3243–3265.

Castelletta, M., Thiollay, J.-M. & Sodhi, N.S. (2005) The effects of extreme forest fragmentation on the bird community of Singapore Island. *Biological Conservation*, **121**, 135–155.

Charles, J. & Ang, B. (2010) Non-volant small mammal community responses to fragmentation of kerangas forests in Brunei Darussalam. *Biodiversity and Conservation*, **19**, 543–561.

Cieślak, M. & Dombrowski, A. (1993) The effect of forest size on breeding bird communities. *Acta Ornithologica*, **27**, 97–111.

Crooks, K.R. (2002) Relative sensitivities of mammalian carnivores to habitat fragmentation. *Conservation Biology*, **16**, 488–502.

Crowe, T.M. (1979) Lots of weeds: insular phytogeography of vacant urban lots. *Journal of Biogeography*, **6**, 169–181.

Daily, G.C. & Ehrlich, P.R. (1995) Preservation of biodiversity in small rainforest patches: rapid evaluations using butterfly trapping. *Biodiversity and Conservation*, **4**, 35–55.

Dalecky, A., Chauvet, S., Ringuet, S. et al. (2002) Large mammals on small islands: short term effects of forest fragmentation on the large mammal fauna in French Guiana. *Revue d'écologie*, **57**, 145–164.

Darlington, J.P.E.C., Kaib, M. & Brand, R. (2001) Termites (Isoptera) in forest remnants and forest islands in the Shimba Hills National Reserve, coastal province of Kenya. *Sociobiology*, **37**, 527–538.

Davies, R.G., Hernández, L.M., Eggleton, P., Didham, R.K., Fagan, L.L. & Winchester, N.N. (2003) Environmental and spatial influences upon species composition of a termite assemblage across Neotropical forest islands. *Journal of Tropical Ecology*, **19**, 509–524.

Davis, R., Dunford, C. & Lomolino, M.V. (1988) Montane mammals of the American Southwest: the possible influence of post-Pleistocene colonization. *Journal of Biogeography*, **15**, 841–848.

Dickman, C.R. (1987) Habitat fragmentation and vertebrate species richness in an urban environment. *Journal of Applied Ecology*, **24**, 337–351.

Dinesen, L., Lehmberg, T., Rahner, M.C. & Fjeldså, J. (2001) Conservation priorities for the forests of the Udzungwa Mountains, Tanzania, based on primates, duikers and birds. *Biological Conservation*, **99**, 223–236.

Edenius, L. & Sjöberg, K. (1997) Distribution of birds in natural landscape mosaics of old-growth forests in northern Sweden: relations to habitat area and landscape context. *Ecography*, **20**, 425–431.

Essl, F. & Dirnböck, T. (2012) What determines Orthoptera species distribution and richness in temperate semi-natural dry grassland remnants? *Biodiversity and Conservation*, **21**, 2525–2537.

Feeley, K. (2003) Analysis of avian communities in Lake Guri, Venezuela, using multiple assembly rule models. *Oecologia*, **137**, 104–113.

Fernández-Juricic, E. (2000) Bird community composition patterns in urban parks of Madrid: the role of age, size and isolation. *Ecological Research*, **15**, 373–383.

Filgueiras, B.K.C., Iannuzzi, L. & Leal, I.R. (2011) Habitat fragmentation alters the structure of dung beetle communities in the Atlantic Forest. *Biological Conservation*, **144**, 362–369.

Flaspohler, D.J., Giardina, C.P., Asner, G.P., Hart, P., Price, J., Lyons, C.K. & Castaneda, X. (2010) Long-term effects of fragmentation and fragment properties on bird species richness in Hawaiian forests. *Biological Conservation*, **143**, 280–288.

Ford, H.A. (1987) Bird communities on habitat islands in England. *Bird Study*, **34**, 205–218.

Galle, R. (2008) The effect of a naturally fragmented landscape on the spider assemblages. *North-Western Journal of Zoology*, **4**, 61–71.

Galli, A.E., Leck, C.F. & Forman, R.T.T. (1976) Avian distribution patterns in forest islands of different sizes in central New Jersey. *The Auk*, **93**, 356–364.

Ganzhorn, J.U., Fietz, J., Rakotovao, E., Schwab, D. & Zinner, D. (1999) Lemurs and the regeneration of dry deciduous forest in Madagascar. *Conservation Biology*, **13**, 794–804.

Gaublomme, E., Hendrickx, F., Dhuyvetter, H. & Desender, K. (2008) The effects of forest patch size and matrix type on changes in carabid beetle assemblages in an urbanized landscape. *Biological Conservation*, **141**, 2585–2596.

Gavish, Y., Ziv, Y. & Rosenzweig, M.L. (2012) Decoupling fragmentation from habitat loss for spiders in patchy agricultural landscapes. *Conservation Biology*, **26**, 150–159.

Gillespie, T.W. & Walter, H. (2001) Distribution of bird species richness at a regional scale in tropical dry forest of Central America. *Journal of Biogeography*, **28**, 651–662.

Haila, Y., Hanski, I.K. & Raivio, S. (1993) Turnover of breeding birds in small forest fragments: the 'sampling' colonization hypothesis corroborated. *Ecology*, **74**, 714–725.

Hatt, R. (1948) *Island life: a study of land vertebrates of the islands of eastern Lake Michigan*. Cranbrook Institute of Science, Bloomfield Hills, MI.

Hattori, T. & Ishida, H. (2000) Relationship between species diversity, species composition and forest area of fragmented lucidophyllous forests in central Miyazaki Prefecture. *Japanese Journal of Ecology*, **50**, 221–234.

Holbech, L.H. (2005) The implications of selective logging and forest fragmentation for the conservation of avian diversity in evergreen forests of south-west Ghana. *Bird Conservation International*, **15**, 27–52.

Hu, G., Wu, J., Feeley, K.J., Xu, G. & Yu, M. (2012) The effects of landscape variables on the species-area relationship during late-stage habitat fragmentation. *PLoS ONE*, **7**, e43894.

Ishida, H., Hattori, T., Takeda, Y. & Kodate, S. (1998) Relationship between species richness or species composition and area of fragmented lucidophyllous forests in southeastern Hyogo Prefecture. *Japanese Journal of Ecology*, **48**, 1–16.

Johnson, N.K. (1975) Controls of number of bird species on montane islands in the Great Basin. *Evolution*, **29**, 545–567.

Kelt, D.A. (2000) Small mammal communities in rainforest fragments in central southern Chile. *Biological Conservation*, **92**, 345–358.

Kitchener, D.J., Chapman, A., Dell, J., Muir, B.G. & Palmer, M. (1980a) Lizard assemblage and reserve size and structure in the Western Australian wheatbelt – some implications for conservation. *Biological Conservation*, **17**, 25–62.

Kitchener, D.J., Chapman, A., Muir, B.G. & Palmer, M. (1980b) The conservation value for mammals of reserves in the Western Australian wheatbelt. *Biological Conservation*, **18**, 179–207.

Kratter, A.W. (1992) Montane avian biogeography in southern California and Baja California. *Journal of Biogeography*, **19**, 269–283.

Langrand, O. (1995) *The effects of forest fragmentation on bird species in Madagascar: a case study from Ambohitantely Forest Reserve on the central high plateau*. University of Natal, Pietermaritzburg.

Lomolino, M.V. & Davis, R. (1997) Biogeographic scale and biodiversity of mountain forest mammals of western North America. *Global Ecology and Biogeography Letters*, **6**, 57–76.

Lomolino, M.V. & Perault, D.R. (2001) Island biogeography and landscape ecology of mammals inhabiting fragmented, temperate rain forests. *Global Ecology and Biogeography*, **10**, 113–132.

Lomolino, M.V., Brown, J.H. & Davis, R. (1989) Island biogeography of montane forest mammals in the American Southwest. *Ecology*, **70**, 180–194.

Lumaret, R., Guillerm, J.-L., Maillet, J. & Verlaque, R. (1997) Plant species diversity and polyploidy in islands of natural vegetation isolated in extensive cultivated lands. *Biodiversity and Conservation*, **6**, 591–613.

McCollin, D. (1993) Avian distribution patterns in a fragmented wooded landscape (north Humberside, U.K.): the role of between-patch and within-patch structure. *Global Ecology and Biogeography Letters*, **3**, 48–62.

Maldonado-Coelho, M. & Marini, M.Â. (2000) Effects of forest fragment size and successional stage on mixed-species bird flocks in southeastern Brazil. *The Condor*, **102**, 585–594.

Marini, M.Â. (2001) Effects of forest fragmentation on birds of the Cerrado region, Brazil. *Bird Conservation International*, **11**, 13–25.

Matthews, T.J., Cottee-Jones, H.E. & Whittaker, R.J. (2014) Habitat fragmentation and the species–area relationship: a focus on total species richness obscures the impact of habitat loss on habitat specialists. *Diversity and Distributions*, **20**, 1136–1146.

Matthiae, P.E. & Stearns, F. (1981) Mammals in forests islands in southeastern Wisconsin. *Forest island dynamics in man-dominated landscapes* (ed. by R.L. Burgess and D.M. Sharpe), pp. 55–66. Springer-Verlag, New York.

Meynard, C.N. & Quinn, J.F. (2008) Bird metacommunities in temperate South American forest: vegetation structure, area, and climate effects. *Ecology*, **89**, 981–990.

Miyashita, T., Shinkai, A. & Chida, T. (1998) The effects of forest fragmentation on web spider communities in urban areas. *Biological Conservation*, **86**, 357–364.

Mohd-Azlan, J. & Lawes, M.J. (2011) The effect of the surrounding landscape matrix on mangrove bird community assembly in north Australia. *Biological Conservation*, **144**, 2134–2141.

Newmark, W.D. (1991) Tropical forest fragmentation and the local extinction of understory birds in the eastern Usambara Mountains, Tanzania. *Conservation Biology*, **5**, 67–78.

Nores, M. (1995) Insular biogeography of birds on mountain-tops in north western Argentina. *Journal of Biogeography*, **22**, 61–70.

Nufio, C., McClenahan, J. & Deane Bowers, M. (2011) Grasshopper response to reductions in habitat area as mediated by subfamily classification and life history traits. *Journal of Insect Conservation*, **15**, 409–419.

Nyeko, P. (2009) Dung beetle assemblages and seasonality in primary forest and forest fragments on agricultural landscapes in Budongo, Uganda. *Biotropica*, **41**, 476–484.

Peltzer, P.M., Lajmanovich, R.C. & Beltzer, A.H. (2003) The effects of habitat fragmentation on amphibian species richness in the floodplain of the Middle Parana River, Argentina. *Herpetological Journal*, **13**, 95–98.

Pineda, E. & Halffter, G. (2004) Species diversity and habitat fragmentation: frogs in a tropical montane landscape in Mexico. *Biological Conservation*, **117**, 499–508.

Ramanamanjato, J.-B. (2000) Fragmentation effects on reptile and amphibian diversity in the littoral forest of southeastern Madagascar. *Isolated vertebrate communities in the tropics* (ed. by G. Rheinwald), pp. 297–308. Zoologisches Forschungsinstitut und Museum A. Koenig, Bonn.

Ribas, C.R., Sobrinho, T.G., Schoereder, J.H., Sperber, C.F., Lopes-Andrade, C. & Soares, S.M. (2005) How large is large enough for insects? Forest fragmentation effects at three spatial scales. *Acta Oecologica*, **27**, 31–41.

Rosenblatt, D.L., Heske, E.J., Nelson, S.L., Barber, D.M., Miller, M.A. & MacAllister, B. (1999) Forest fragments in east-central Illinois: islands or habitat patches for mammals? *The American Midland Naturalist*, **141**, 115–123.

Ruiz-Gutiérrez, V., Gavin, T.A. & Dhondt, A.A. (2008) Habitat fragmentation lowers survival of a tropical forest bird. *Ecological Applications*, **18**, 838–846.

dos Santos, K., Kinoshita, L.S. & Santos, F.A.M.D. (2007) Tree species composition and similarity in semideciduous forest fragments of southeastern Brazil. *Biological Conservation*, **135**, 268–277.

Shreeve, T.G. & Mason, C.F. (1980) The number of butterfly species in woodlands. *Oecologia*, **45**, 414–418.

Silva, M. (2001) Abundance, diversity, and community structure of small mammals in forest fragments in Prince Edward Island National Park, Canada. *Canadian Journal of Zoology*, **79**, 2063–2071.

Silva, M. & Pôrto, K. (2009) Effect of fragmentation on the community structure of epixylic bryophytes in Atlantic Forest remnants in the northeast of Brazil. *Biodiversity and Conservation*, **18**, 317–337.

Simberloff, D. & Martin, J.L. (1991) Nestedness of insular avifaunas: simple summary statistics masking complex species patterns. *Ornis Fennica*, **68**, 178–192.

Smith, G.T., Arnold, G.W., Sarre, S., Abensperg-Traun, M. & Steven, D.E. (1996) The effect of habitat fragmentation and livestock grazing on animal communities in remnants of gimlet *Eucalyptus salubris* woodland in the Western Australian wheatbelt. II. Lizards. *Journal of Applied Ecology*, **33**, 1302–1310.

Suarez, A.V., Bolger, D.T. & Case, T.J. (1998) Effects of fragmentation and invasion on native ant communities in coastal southern California. *Ecology*, **79**, 2041–2056.

Summerville, K.S., Veech, J.A. & Crist, T.O. (2002) Does variation in patch use among butterfly species contribute to nestedness at fine spatial scales? *Oikos*, **97**, 195–204.

Tonn, W.M. & Magnuson, J.J. (1982) Patterns in the species composition and richness of fish assemblages in northern Wisconsin lakes. *Ecology*, **63**, 1149–1166.

Usher, M.B. & Keiller, S.W.J. (1998) The macrolepidoptera of farm woodlands: determinants of diversity and community structure. *Biodiversity and Conservation*, **7**, 725–748.

Vallan, D. (2000) Influence of forest fragmentation on amphibian diversity in the nature reserve of Ambohitantely, highland Madagascar. *Biological Conservation*, **96**, 31–43.

Viveiros de Castro, E.B. & Fernandez, F.A.S. (2004) Determinants of differential extinction vulnerabilities of small mammals in Atlantic forest fragments in Brazil. *Biological Conservation*, **119**, 73–80.

Wang, Y., Bao, Y., Yu, M., Xu, G. & Ding, P. (2010) Nestedness for different reasons: the distributions of birds, lizards and small mammals on islands of an inundated lake. *Diversity and Distributions*, **16**, 862–873.

Watson, D.M. (2003) Long-term consequences of habitat fragmentation – highland birds in Oaxaca, Mexico. *Biological Conservation*, **111**, 283–303.

Weaver, M. & Kellman, M. (1981) The effects of forest fragmentation on woodlot tree biotas in southern Ontario. *Journal of Biogeography*, **8**, 199–210.

Willson, M.F., De Santo, T.L., Sabag, C. & Armesto, J.J. (1994) Avian communities of fragmented south-temperate rainforests in Chile. *Conservation Biology*, **8**, 508–520.

Yong, D.L., Qie, L., Sodhi, N.S., Koh, L.P., Peh, K.S.-H., Lee, T.M., Lim, H.C. & Lim, S.L.-H. (2011) Do insectivorous bird communities decline on land-bridge forest islands in Peninsular Malaysia? *Journal of Tropical Ecology*, **27**, 1–14.

Zimmerman, B.L. & Bierregaard, R.O. (1986) Relevance of the equilibrium–theory of island biogeography and species area relations to conservation with a case from Amazonia. *Journal of Biogeography*, **13**, 133–143.