

A hierarchical model of whole assemblage island biogeography

Jesse R. Lasky^{1,2,3}, Timothy H. Keitt³, Brian C. Weeks^{2,4}, Evan P. Economo⁵

¹Department of Biology, Pennsylvania State University, University Park, PA 16802, USA

²Department of Ecology, Evolution and Environmental Biology, Columbia University, New York, NY 10025, USA

³Department of Integrative Biology, University of Texas at Austin, Austin, TX 78704, USA

⁴Department of Ornithology, American Museum of Natural History, New York, NY 10024, USA

⁵Okinawa Institute of Science and Technology Graduate University, Onna-son, Okinawa, 904-0495, Japan

Corresponding author: Evan P. Economo, Okinawa Institute of Science and Technology Graduate University, Onna-son, Okinawa, 904-0495, Japan. E-mail: evaneconomo@gmail.com

Decision date: 22-Jul-2016

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1111/ecog.02303].

ABSTRACT

Island systems have long played a central role in the development of ecology and evolutionary biology. However, while many empirical studies suggest species differ in vital biogeographic rates, such as dispersal abilities, quantitative methods have had difficulty incorporating such differences into analyses of whole-assemblages. In particular, differences in dispersal abilities among species can cause variation in the spatial clustering and localization of species distributions. Here, we develop a single, hierarchical Bayes, assemblage-wide model of 252 bird species distributions on the islands of Northern Melanesia and use it to investigate a) whether dispersal limitation structures bird assemblages across the archipelago, b) whether species differ in dispersal ability, and c) test the hypothesis that wing aspect ratio, a trait linked to flight efficiency, predicts differences inferred by the model. Consistent with island biogeographic theory, we found that individual species were more likely to occur on islands with greater area, and on islands near to other islands where the species also occurred. However, species showed wide variation in the importance and spatial scale of these clustering effects. The importance of clustering in distributions was greater for species with low wing aspect ratios, and the spatial scale of clustering was also smaller for low aspect ratio species.

These findings suggest that the spatial configuration of islands interacts with species dispersal ability to affect contemporary distributions, and that these species differences are detectable in occurrence patterns. More generally, our study demonstrates a quantitative, hierarchical approach that can be used to model the influence of dispersal heterogeneity in diverse assemblages and test hypotheses for how traits drive dispersal

Accepted Article

differences, providing a framework for deconstructing ecological assemblages and their drivers.

Keywords: island biogeography, hierarchical Bayes, dispersal, occupancy model, species-area curve, birds, Melanesia

INTRODUCTION

A central goal of both community ecology and biogeography is to understand the processes determining species distributions and how they contribute to assemblage-level patterns (MacArthur and Wilson 1963, Cracraft 1986, Leibold et al. 2004, Lomolino et al. 2010). Biologists have long recognized that the spatial context of habitat patches (*e.g.* islands) interacts with dispersal ability to influence the number and distribution of resident species (*e.g.* MacArthur and Wilson 1963, 1967). The role of patch area and isolation, in particular, is a theoretical cornerstone of island biogeography, metapopulation, and metacommunity ecology (MacArthur and Wilson 1967, Hanski 1994, Rosenzweig 1995, Hubbell 2001, Leibold et al. 2004). Species interact with patch area and isolation, and broader landscape structures, through their abilities to colonize and persist in habitat patches. In turn, numerous studies have documented morphological and behavioral traits contributing to dispersal ability (Lockwood et al. 1998, Dawideit et al. 2009), and it is well understood that species and higher taxa vary in their dispersal abilities with consequences for a wide range of ecological and evolutionary processes (Wilson 1961, Hanski 1994, Economo and Sarnat 2012, Jønsson et al. 2014, Weeks and Claramunt 2014).

Although dispersal limitation is a central feature of most quantitative models of island biogeography (*e.g.* Rosindell and Phillimore 2011, Triantis et al. 2015, Valente et al. 2015, Carvalho et al. 2015), these models generally do not incorporate variability in dispersal abilities among species in the community (but see Sukumaran et al. 2016). Furthermore, statistical approaches to island biogeography often focus on aggregate

Accepted Article

patterns such as species-area curves while losing information about individual species and their differences (Rosenzweig 1995, Kéry & Royle 2008, Triantis et al. 2012).

Similarly, multivariate methods in community ecology often operate on measures of aggregate community dissimilarity thereby eliminating information about individual species (Legendre and Legendre, 2012).

Here, we develop a novel statistical approach to modeling island biogeographic patterns that quantifies the effects of island area and spatial configuration (isolation) on community patterns, but deconstructs the biotic pattern based on species differences in how they respond to area and isolation. This allows us to address a series of questions about the roles of dispersal limitation, dispersal heterogeneity, and dispersal functional traits in driving island biogeography of birds in Northern Melanesia, a classic system in biogeography.

Two quantitative approaches frame our development of a parametric, spatially-explicit multi-species model of island occurrences: 1) spatially-explicit auto-regressive models of occurrence (*e.g.* Hanski 1994, Gelfand et al. 2006) and 2) multilevel statistical models (Clark 2005, Gelfand et al. 2006, Condit et al. 2013). Like MacArthur and Wilson's theory (1967), the spatially explicit metapopulation model of Hanski (1994) emphasized patch size and isolation as determinants of occurrence, but incorporated more realistic spatial structures than possible in mainland-island models. Hanski (1994) models the stochastic process of colonization and extinction in patches. In this model, the probability of a species colonizing a focal patch is a function of its presence in surrounding patches, the distance of those patches to the focal patch (with distance negatively related to occurrence) and the area of those patches (with area positively

related to occurrence, Hanski 1994, Hanski et al. 1996). As focal patch area increases, extinction probability decreases. Hanski explicitly modeled the influence of patch area via extinction. However, additional synergistic processes may operate: greater patch area can increase colonization rates due to larger target size (Lomolino 1990), greater immigration can reduce extinction rates (“rescue effect”) (Brown and Kodric-Brown 1977), and larger patches or islands often contain a wider range of environmental conditions, thus enhancing the chance of having habitat appropriate for a given species (Lack 1969).

Second, the approach of incorporating multiple species into multilevel models allows one to share information across species and constrain parameter estimates for species for which few data exist (Gelfand et al. 2006, Ruger et al. 2012, Lasky et al. 2013, Condit et al. 2013). Models of occupancy are underpowered when variance in occurrence is low (*i.e.* when a species is nearly always present or always absent, Hanski 1994). Additionally, multi-species models permit inferences of differences among islands (via random effects) caused by processes not captured by measured covariates, inference that would not be possible for single species models. Here we develop a hierarchical Bayes approach to jointly model the occurrence of all species in an assemblage. We assume that species-specific parameters for response to island area and isolation arise from hyperdistributions of parameters for all species.

If species distributions are limited by dispersal, then we expect that the clustering of species occurrences will be related to interspecific differences in dispersal ability (Nekola and White 1999, Levine and Murrell 2003). For birds, dispersal is a complex process influenced by behavior, physiology and morphology. In particular, wing

morphology is a key determinant of the efficiency of long distance flight; wing aspect ratio has been correlated to dispersal distance and the length of migration (Lockwood et al. 1998, Dawideit et al. 2009), and range size. Claramunt et al. (2012) demonstrated that aspect ratio can be approximated using a measure called the hand-wing index (HWI) (Kipp 1959, see similar metrics for bats in Norberg and Rayner 1987), which can be obtained by measuring museum study skins. High values of HWI indicate species with longer and more pointed wings, which are expected to improve flight efficiency, resulting in superior dispersal ability. Here, we tested whether interspecific variation in HWI was associated with our estimates of interspecific variation in putative dispersal parameters.

We apply our model to the complete avifauna of the northern Melanesia archipelago, a system that has produced classic empirical studies of island biogeography (*e.g.* Diamond 1973, 1977, Diamond and Mayr 1976, Mayr and Diamond 1976, Diamond and Marshall 1976, Filardi and Moyle 2005). Historically, studies of the Melanesian avifauna have largely relied on estimating dispersal ability based on the relationship between species distributions and biogeographical breaks. These approaches require *a priori* knowledge of biogeographic breaks, which are poorly known in many systems. We ask what can be inferred about the role of dispersal limitation and dispersal heterogeneity using occurrence data for a whole assemblage.

Our analysis focuses on the following questions. First, we ask whether there is evidence of dispersal limitation *in general* (*i.e.* across the entire avifauna) structuring island occupancy. Second, we test the hypothesis that dispersal heterogeneity across species is important for driving island occupancy patterns, based on differences in the spatial distribution of each species, and in particular how each species distribution

Accepted Article
responds to the area and arrangement of patches. Third, we test the hypothesis that HWI is predictive of the species-specific dispersal parameters inferred by our model. This latter test provides a cross-validation of our modeling approach and a potential functional trait that determines bird dispersal.

METHODS

Species distributional data

We obtained distributional data from a recent monograph that included a compilation of bird species distributions across 75 islands in Northern Melanesia plus Australia and New Guinea (Mayr and Diamond 2001). There were 252 total species and 3,347 total species-island occurrences, with the median species occurring on 5 islands and the median island containing 41 species. For many of these species, Mayr and Diamond (2001) also delineated multiple subspecies. However, recent molecular work suggests that this classification underestimates the diversity of the system, some subspecies represent more significant evolutionary units (Moyle et al. 2005, Reddy and Moyle 2011, Andersen et al. 2013), and there is extensive paraphyly within species (*e.g.* Andersen et al. 2015). Thus we conducted two analyses: one in which we treated taxa with multiple subspecies as a single species ($N = 252$), and a second in which each subspecies was treated as a distinct species ($N = 524$). If subspecies delineated by Mayr and Diamond (2001) are younger lineages than species, there may have been less time for range expansion following lineage birth, limiting the amount of influence dispersal would have on distributions (de

Moraes Weber et al. 2014). Thus we expect to find weaker evidence for dispersal limitation of subspecies.

Geographic data

We used the GSHHG shoreline database (Wessel and Smith 1996) distributed by NOAA (www.noaa.gov) to find the areas and minimum distances between islands. These computations were performed in MATLAB.

Wing data

HWI was calculated using museum study skins at the American Museum of Natural History. We include HWI for all landbird species in the Solomon Archipelago, and for an additional set of taxa from outside of the Solomon Archipelago but within Northern Melanesia. In total we included 501 specimens representing 166 species. Following Claramunt et al. (2012), HWI was calculated for each specimen as:

$$\text{HWI} = 100 \times (\text{WL} - \text{SL})/\text{WL} \quad (\text{eqn. 1})$$

where WL is the standard measure of wing length, and SL is the distance from the carpal joint to the first secondary feather. For each species, 3 adult male specimens were measured – when available – and an average value for each species was used.

Statistical model

We developed a hierarchical Bayes model of island biogeography that accounts for the effects of island area, the influence of mainland and surrounding islands and interspecific variation in the effects of these factors. The model makes use of information

in the spatial distributions of species across the archipelago to infer aspects of area and distance dependency, which in turn should reflect species biology and landscape structure.

We modeled the probability p_{ki} that species k occurs on island i , given the area of the island and the presence or absence of species k on nearby islands. The latter is modeled as an auto-covariate such that the model is auto-regressive. We modeled all species simultaneously and estimated hyperparameters that describe the distribution of covariate effects across species (Gelfand et al. 2006, Ruger et al. 2012, Lasky et al. 2013, Condit et al. 2013).

The probability p_{ik} species k occurs on island i was calculated via a link function

$$p_{ik} = \frac{1}{1+e^{-\lambda_{ik}}} \quad (\text{eqn. 2})$$

where error in occurrence data y_{ik} follows a Bernoulli distribution, and λ_{ik} is given by

$$\lambda_{ik} = \beta_k + \beta_i + \beta_{area,k} \log(A_i) + M_{ik} \quad (\text{eqn. 3}).$$

The parameters β_k and β_i reflect random effects for species and island, respectively, the $\beta_{area,k}$ controls the strength of the area dependency, A_i is the area of the i^{th} island, and M_{ik} is the auto-covariate function reflecting the influence of dispersal. Focal patch area is considered by Hanski (1994) via an explicit model of extinction. Here we chose a more phenomenological formulation for area effects due to the multiple mechanisms that can link area and occurrence.

For M_{ik} , we drew upon the incidence auto-covariate function from the spatially explicit metapopulation model of Hanski (1994). This term, for M_{ik} of species k into focal island i , is

$$M_{ik} = \beta_{M,k} \sum_{j \neq i}^J y_{jk} A_j e^{-d_{ij}/\alpha_k} \quad (\text{eqn. 4})$$

where j indexes over the J islands, y_{jk} is a binary variable reflecting the occupancy of species k on island j (1 if present, 0 if absent), A_j is the area of the j^{th} island, and spatial separation between islands i and j is given by d_{ij} . The species-specific parameter α_k determines the distance decay in occupancy probability and thus should be informative about the scale of the distance effects, while $\beta_{M,k}$ reflects the importance of the distance effects overall in predicting the occupancy pattern. The influence of proximity to Australia and the island of New Guinea was included in the same manner as other neighboring islands. However, because it is unclear whether the scaling of area effects (A_j in eqn. 4) on colonization would be constant across such a wide range of island sizes from $\sim 1 \text{ km}^2$ to New Guinea at $\sim 786,000 \text{ km}^2$, the area effects of Australia and New Guinea were estimated as separate parameters.

Species intercept (β_k), species area ($\beta_{area,k}$), and island intercept (β_i) parameters were drawn from separate normal hyperdistributions with means of β_μ , β_{area} , and 0, respectively. Thus β_μ determined the grand mean of species occurrences and β_{area} determined the average species response to area. Species dispersal ($\beta_{M,k}$) and scale parameters (α_k) were drawn from separate gamma hyperdistributions, which constrained the effect of occurrence on neighboring islands j to have a positive effect on occurrence on island i and constrained this effect to decay with distance, respectively.

Parameters were estimated using Bayesian inference where parameters had proper and diffuse priors (see Supplemental Material for analysis and model code). Markov Chain Monte Carlo sampling of posterior distributions was conducted using JAGS 3.3 (<http://mcmc-jags.sourceforge.net/>) and R statistical software (CRAN Core Development Team). We assessed convergence using the potential scale reduction factor, ensuring that all hyperparameters had values <1.1 (Gelman and Rubin 1992). To assess the ability of our model to capture patterns in the data (as opposed to fitting patterns due to model bias) we sampled parameters when ignoring data (*i.e.* sampling from priors) and when modeling permuted data (Supplemental Material, Gotelli and Entsminger 2001).

Comparison of dispersal parameters with species traits

We compared the two species-specific parameters determining dispersal effects, $\beta_{M,k}$ and α_k , to HWI using non-parametric rank correlations. Mayr & Diamond (2001) also estimated species local abundance (categorized by Mayr & Diamond along a scale of 1-5), which may be related to the importance of dispersal on distributions, *e.g.* via mass effects (Shmida and Wilson 1985). We compared species abundance categories with the species-specific parameters determining putative dispersal effects and island area effects.

Posterior uncertainty in species parameters was propagated to correlations with HWI and abundance via repeated re-calculation of correlation from each posterior sample.

Because $\beta_{M,k}$ and HWI exhibited a triangular-shaped relationship, we tested quantile regression between HWI and the 90th percentile of log-transformed $\beta_{M,k}$, although results were consistent for other upper quantiles (Supplemental Material). We

also compared subspecies-specific parameters estimated from a model where subspecies were treated as a unique species, using the same HWI for each subspecies of a species.

RESULTS

Relationship between island area and occurrence

Island area had a significantly positive effect on occurrence for the distribution across species (Figure 1, all following parameter point estimates are medians of posterior samples, $\beta_{area} = 0.641$, 95% CI = 0.534, 0.750). Only three species showed a significantly negative effect of island area on occurrence, *Monarcha cinerascens* ($\beta_{area,k} = -0.550$, 95% CI = -0.795, -0.302), *Ducula pacifica* ($\beta_{area,k} = -0.365$, 95% CI = -0.594, -0.105), and *Aplonis feadensis* ($\beta_{area,k} = -0.317$, 95% CI = -0.522, -0.109).

Relationship between putative dispersal and occurrence

Species differed substantially in the effect of putative dispersal on occurrence (Figure 1), which was constrained to be non-negative. The mode of the hyperdistribution was near zero (0.005) although the variance was 0.069 (gamma hyperdistribution mean = 0.264, shape = 1.017, 95% CI = 0.757 1.371; rate = 3.845, 95% CI = 2.582, 5.701, Figure 1).

The species with the highest $\beta_{M,k}$ was *Myzomela pammelaena* and was restricted to the western islands. Examples of species distributions varying in $\beta_{M,k}$ are presented in Figure 2. Sampling from the prior distribution and posterior sampling using permuted data both gave very low estimates of $\beta_{M,k}$, with little interspecific variation, indicating that our

inferences on the real data were not driven simply by model constraint (Supplemental Material).

Species also differed extensively in the distance decay of putative dispersal effects, estimated by the α_k parameter (Figure 1). The mode of the hyperdistribution was zero although the variance was 0.228 (gamma hyperdistribution mean = 0.444, shape = 0.865, 95% CI = 0.652, 1.167; rate = 1.946, 95% CI = 1.244, 2.950, Figure 1). The effect of the distance decay parameter is partly conditional on the importance of dispersal ($\beta_{M,k}$). The species with the lowest distance decay (highest α_k) with at least moderate importance of dispersal (*i.e.* > 0.25, approximately the top quintile of species) was a trivial case, the omnipresent *Egretta sacra*. The next highest were *Chalcopsitta cardinalis*, which is largely restricted to the Solomon Islands despite large intervening distances between occupied islands, *Myzomela pammelaena*, which was completely restricted to the northwestern islands despite large distances among them, and *Monarcha cinerascens*, which is moderately clustered across northwestern and eastern islands with relatively large distances between them (Figure 2). Examples of species distributions varying in α_k are presented in Figure 3.

Relationship between traits and putative dispersal effects on occurrence

The importance of putative dispersal ($\beta_{M,k}$) for a species occurrence was weakly related to HWI (median of Spearman's rank correlations, ρ_{med} , between HWI and $\beta_{M,k}$, $\rho_{\text{median}} = -0.057$, 95% CI = -0.169, 0.050; Figure 4). However, there were no species with high HWI and high $\beta_{M,k}$, indicating a triangular relationship between these variables (Figure 4). Thus we tested the relationship between HWI and high $\beta_{M,k}$, using quantile regression

Accepted Article

for the 90th percentile (*i.e.* percentile among species) of log-transformed $\beta_{M,k}$ and found a decrease in $\beta_{M,k}$ with increasing HWI (posterior median of quantile regression coefficients, \square_{med} , between HWI and $\log \beta_{M,k}$, $\square_{\text{med}} = -0.035$, 95% CI = -0.045, -0.019 log units $\beta_{M,k}$ / unit HWI, other quantiles yielded similar results, Supplemental Material). The distance decay of putative dispersal effects was less when the α_k parameter was high, which was positively related to HWI ($\square_{\text{med}} = 0.266$, 95% CI = 0.163, 0.365; Figure 4). We also found that abundance categories (as estimated on a 1-5 scale by Mayr & Diamond 2001) were positively associated with $\beta_{M,k}$ (posterior median of Spearman's rank correlations, $\square_{\text{median}} = 0.260$, 95% CI = 0.167, 0.350) but were not associated with α_k ($\square_{\text{median}} = -0.067$, 95% CI = -0.160, 0.350). However, abundance was negatively related to HWI (Spearman's $\square_{\text{median}} = -0.067$, 95% CI = -0.160, 0.350), with no species in the highest abundance category having HWI > 28.75 (33rd percentile).

Results were weaker when comparing HWI to parameters estimated when subspecies were the unit of analysis. The importance of putative dispersal ($\beta_{M,k}$) for a subspecies occurrence was essentially unrelated to HWI (posterior median Spearman's rank correlation (\square_{med}) between HWI and $\beta_{M,k}$, $\square_{\text{median}} = 0.066$, 95% CI = -0.007, 0.136; Figure S1). Similarly, quantile regression showed no relationship between HWI and the subspecies dispersal parameter (posterior median of 90th percentile regression (\square_{med}) between HWI and $\log \beta_{M,k}$, $\square_{\text{med}} = 0.001$, 95% CI = -0.011, 0.014). The distance decay of putative dispersal effects was less when the subspecies α_k parameter was high, which was significantly positively related to HWI ($\square_{\text{med}} = 0.177$, 95% CI = 0.092, 0.256; Figure S1).

DISCUSSION

Island systems provide important test cases for understanding how ecological and evolutionary processes generate biodiversity patterns across spatially complex landscapes. We present a new quantitative approach to the analysis of island biota, encompassing both traditional theoretical ideas about the roles of area and isolation and new hierarchical approaches to modeling species and site differences. Our method provides a generalizable approach to understanding the roles of dispersal limitation and dispersal heterogeneity in driving biotic patterns, using the spatial occupancy pattern of different species to infer dispersal parameters and responses to island area and isolation.

Our model used the distributions of 252 species distributed across 75 islands to infer the responses of species to area and isolation parameters. First, as expected based on island biogeography theory (MacArthur and Wilson 1967) and the commonness of species-area relationships (Rosenzweig 1995, Triantis et al. 2012), we found a strong effect of island area on occupancy probability for most species. Most species were more likely to occur on large islands compared with small islands, even after accounting for spatial clustering effects. Higher probability of occurrence on larger islands may represent lower stochastic extinction due to larger population sizes (MacArthur and Wilson 1967, Hanski 1994), consistent with our finding that more abundant species exhibited weaker associations with island area. Additionally, larger islands may harbor more diverse environments and may have a greater chance of having habitat appropriate for a given species (Lack 1969).

Variation in the area parameters are consistent with known biology of the species involved. For example, the presences of several montane specialists, such as

Gymnophaps albertisii, *Accipiter brachyurus*, and *Rhipidura dahlia*, are among those species with the strongest estimated sensitivity to area (found only on two of the three largest islands). This suggests the presence of diverse environments is at least partly responsible for the area effect, as island area is strongly correlated with maximum elevation (Spearman's $\rho = 0.78$). In contrast, only three species (*Monarcha cinerascens*, *Ducula pacifica*, and *Aplonis feadensis*) exhibited a significantly negative relationship with area, indicating they are more likely to occur on small islands. All three of these species are classified as “supertramps” by Diamond (1975) because they are widely distributed, but only occur on smaller islands, possibly owing to competition with species inhabiting larger islands, a preference for habitats associated with small islands, or idiosyncratic colonization-extinction history. However, we note that these species may occur on larger islands outside of the current study area (BirdLife International (2015) IUCN Red List for birds. Downloaded from <http://www.birdlife.org> on 10/12/2015.).

Second, we found strong overall support for the hypothesis that dispersal limitation structures island bird biotas in Northern Melanesia, with an overall community-wide effect for a distance-decay effect of occupancy probability. We inferred strong clustering effects for many species (Figure 1), highlighting the non-random nature of species distributions, and furthermore there was evidence for a strong influence of dispersal heterogeneity across species. The negative relationship between the estimated importance of clustering on distributions and hand-wing index (HWI) suggests dispersal limitation is a major driver of this clustering. Furthermore, the spatial scale of clustering was positively related to HWI, suggesting that the parameters inferred by our model have a biological basis. This finding supports the proposition that dispersal traits are relevant

to species distributions (Seidler and Plotkin 2006), and provides support for the use of HWI as an indicator of dispersal ability in birds. Although we found stronger clustering effects for species with greater local abundance, the cause of this relationship is unclear. Higher abundance was correlated with lower HWI, which may drive stronger clustering effects. Additionally, higher local abundances may be associated with greater numbers of emigrants, potentially increasing the importance of dispersal in distributions, *e.g.* via mass effects (Shmida and Wilson 1985).

Our model represents an emerging approach to modeling island biogeography, and more generally, spatially structured biodiversity patterns. The accuracy of distribution models can be improved by incorporating dispersal components (Merow et al. 2011). A growing number of species distribution models employ hierarchical Bayesian spatial approaches (Gelfand et al. 2006). These approaches are flexible to allow accommodation of multiple sources of error (*e.g.* detection error, Kéry and Royle 2008, Dorazio 2014), prior ecological knowledge (*e.g.* response to an environmental gradient, Keil et al. 2014, Golding and Purse 2016), and allow a fuller characterization of uncertainty via posterior sampling (contrasted with parameter point estimates, Latimer et al. 2006). This uncertainty can be propagated into model predictions (see O'Hara et al. 2002 for a metapopulation example). An additional feature is that multi-species, or joint, distribution models provide a more process-oriented approach to modeling inter-site variation in species richness (Gelfand et al. 2006, Kéry and Royle 2008), although caution is warranted to avoid over-prediction (Clark et al. 2014).

Our current formulation has several limitations. First, we did not explicitly include effects of environmental gradients (abiotic and biotic) to explain species

distributions across islands (Gelfand et al. 2006, Condit et al. 2013, Clark et al. 2014).

While spatially-autocorrelated environmental factors likely influence species distributional clustering, our study system was chosen due to a lack of large-scale precipitation or temperature gradients that might cause major assemblage turnover. In systems with strong environmental gradients, such issues deserve greater consideration (Condit et al. 2013).

Additionally, our model does not account for imperfect detection of species presence by Mayr and Diamond (2001), or potential taxonomic uncertainty in species delineation. However, our use of different levels of taxa (species vs. subspecies) provides an extreme test of the influence of species limits, and there have been an extremely limited number of range expansions and discoveries of new species during re-surveys of Northern Melanesia since the Whitney South Sea expedition in the early 20th century. Although birds are one of the best sampled taxonomic groups, application of our method to more undersampled taxa, such as insects, would probably require formal consideration of the effect of undersampling or variation in sampling effort across the archipelago (*e.g.* Gelfand et al. 2006, Kéry & Royle 2008).

We modeled the potential inputs from neighboring islands as a function of neighbor area. However, populations and thus potential inputs of some species (*e.g.* coastal species) might not respond closely to island area. An alternative formulation could use an additional parameter to modify the dispersal inputs of neighboring islands. We opted not do so here in favor of limiting model complexity, though such a model may be more realistic.

Accepted Article

Finally, our model does not explicitly account for the effects of speciation and variation in spatial distribution driven by lineage age. For example, older species may have had more time to disperse and expand their range than younger species, even if dispersal abilities are similar (de Moraes Weber et al. 2014). We also found that putative dispersal limitation in our study system was weaker for subspecies delineated by Mayr and Diamond (2001), possibly because subspecies are younger lineages than species, leaving less time for range expansion and the accumulation of effects of dispersal limitation on contemporary distributions of assemblages. We emphasize that our model does not attempt to realistically capture all processes determining species distributions. Rather, our goal was to take initial steps toward assemblage-wide modeling of species-specific dispersal and area effects and cross-validate it with trait data.

Our study points to two main directions for further work. First, hierarchical modeling provides a route to disaggregating community patterns and evaluating the role of individual species differences, while sharing statistical information across the whole community (Lasky et al. 2013, Condit et al. 2013). In this way, species-specific responses to area and isolation can be characterized and hypotheses can be tested as to the causal morphological, physiological, or life history traits. In the example analysis presented, we had *a priori* expectations for the relationship between bird wing morphology and spatial distribution due to previous research on the topic. However, for other less-studied taxa, this modeling approach could be a tool for discovery and testing of novel ideas about what traits drive dispersal and area sensitivity.

The second general direction would be to add realism to the model by incorporating additional predictor variables likely to affect species distributions. For

Accepted Article

example, environmental, historical, and phylogenetic data may all provide information on the mechanisms driving species distributions (Leibold et al. 2010). Emerging approaches to modeling species co-occurrence allow direct modeling of species interaction effects on occurrence (Ovaskainen et al. 2010, Clark et al. 2014), though these approaches introduce many more free parameters and thus require larger numbers of sites than ours for the large number of species we studied (252).

Conclusions

MacArthur and Wilson's island biogeography theory is a foundational idea because it highlighted the roles of area and isolation and their effects on aggregate community patterns. Modern theory still gives prominence to area and isolation as drivers of biotic pattern, but countless empirical studies have highlighted the variation across species in dispersal ability and other factors important for the vital biogeographic rates of colonization, extinction, and speciation. Our study represents a statistical approach that assembles community patterns through variable and species-specific responses to area and isolation, providing a route to understanding how traits scale up to community patterns.

Acknowledgements

We thank Nichar Gregory, Erica Fischer, and Alyssa Klavans who assisted in collecting wing morphology data from collections at the American Museum of Natural History. Sarah Lucas assisted in digitizing data from Mayr & Diamond (2001). EPE was supported by subsidy funding to OIST. THK acknowledges the support of the U.S.

National Science Foundation (IOS-0922457, EF-1064901). BCW was supported by the
National Science Foundation Graduate Research Fellowship under grant no. (11-44155).

Literature Cited

- Andersen, M. J. et al. 2013. Phylogeography of the variable dwarf-kingfisher *Ceyx lepidus* (Aves: Alcedinidae) inferred from mitochondrial and nuclear DNA sequences. - *The Auk* 130: 118–131.
- Andersen, M.J. et al. 2015. Rapid diversification and secondary sympatry in Australo-Pacific kingfishers (Aves: Alcedinidae: *Todiramphus*). -*R. Soc. Open Sci.* 2: 140375.
- Brown, J. H. and Kodric-Brown, A. 1977. Turnover rates in insular biogeography: effect of immigration on extinction. - *Ecology* 58: 445–449.
- Carvalho, J.C. et al. 2015. Modeling directional spatio-temporal processes in island biogeography. -*Ecology and Evolution* 5: 4671-4682
- Claramunt, S. et al. 2012. High dispersal ability inhibits speciation in a continental radiation of passerine birds. - *Proc. R. Soc. B Biol. Sci.* 279: 1567-1574.
- Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. - *Ecol. Lett.* 8: 2–14.
- Clark, J. S. et al. 2014. More than the sum of the parts: forest climate response from joint species distribution models. - *Ecol. Appl.* 24: 990–999.
- Condit, R. et al. 2013. Species distributions in response to individual soil nutrients and seasonal drought across a community of tropical trees. - *Proc. Natl. Acad. Sci.* 110: 5064–5068.
- Cracraft, J. 1986. Origin and evolution of continental biotas: speciation and historical congruence within the Australian avifauna. - *Evolution*: 977–996.

- Dawideit, B. A. et al. 2009. Ecomorphological predictors of natal dispersal distances in birds. - *J. Anim. Ecol.* 78: 388–395.
- de Moraes Weber, M. et al. 2014. Have old species reached most environmentally suitable areas? A case study with South American phyllostomid bats. - *Glob. Ecol. Biogeogr.* 23: 1177–1185.
- Diamond, J. M. 1973. Distributional ecology of New Guinea birds: Recent ecological and biogeographical theories can be tested on the bird communities of New Guinea. - *Science* 179: 759–769.
- Diamond, J. M. 1977. Continental and insular speciation in Pacific land birds. - *Syst. Biol.* 26: 263–268.
- Diamond, J. M. and Mayr, E. 1976. Species-area relation for birds of the Solomon Archipelago. - *Proc. Natl. Acad. Sci.* 73: 262–266.
- Diamond, J. and Marshall, A. 1976. Origin of the New Hebridean avifauna. - *Emu* 76: 187–200.
- Dorazio, R. M. 2014. Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. - *Glob. Ecol. Biogeogr.* 23: 1472–1484.
- Economo, E.P. and Sarnat, E.M. 2012. Revisiting the ants of Melanesia and the taxon cycle: historical and human-mediated invasions of a tropical archipelago. - *Am. Nat.* 180: E1-E16.
- Filardi, C. E. and Moyle, R. G. 2005. Single origin of a pan-Pacific bird group and upstream colonization of Australasia. - *Nature* 438: 216–219.
- Gelfand, A. E. et al. 2006. Explaining species distribution patterns through hierarchical modeling. - *Bayesian Anal.* 1: 1–35.

- Gelman, A. and Rubin, D. B. 1992. Inference from iterative simulation using multiple sequences. - *Stat. Sci.*: 457–472.
- Golding, N. and Purse, B. V. 2016. Fast and flexible Bayesian species distribution modelling using Gaussian processes. - *Methods Ecol. Evol.* 7: 598–608.
- Gotelli, N. and Entsminger, G. 2001. Swap and fill algorithms in null model analysis: rethinking the knight's tour. - *Oecologia* 129: 281–291.
- Hanski, I. 1994. A practical model of metapopulation dynamics. - *J. Anim. Ecol.* 63: 151–162.
- Jønsson, K.A. et al. 2014. Evidence of taxon cycles in an Indo-Pacific passerine bird radiation (Aves: Pachycephala). - *Proc. R. Soc. B.* 281: 20131727.
- Keil, P. et al. 2014. Uncertainty, priors, autocorrelation and disparate data in downscaling of species distributions. - *Divers. Distrib.* 20: 797–812.
- Kéry, M. and Royle, J. A. 2008. Hierarchical Bayes estimation of species richness and occupancy in spatially replicated surveys. - *J. Appl. Ecol.* 45: 589–598.
- Kipp, F. A. 1959. Der Handflügel-Index als flugbiologisches Maß. - *Vogelwarte* 20: 77–86.
- Lack, D. 1969. The numbers of bird species on islands. - *Bird Study*, 16:4, 193-209.
- Lasky, J. R. et al. 2013. Trait-mediated effects of environmental filtering on tree community dynamics. - *J. Ecol.* 101: 722–733.
- Latimer, A. M. et al. 2006. Building statistical models to analyze species distributions. - *Ecol. Appl.* 16: 33–50.
- Legendre, P., and Legendre, L. F. J. 2012. *Numerical ecology.* -Elsevier.

- Leibold, M. A. et al. 2004. The metacommunity concept: a framework for multi-scale community ecology. - *Ecol. Lett.* 7: 601–613.
- Leibold, M. A. et al. 2010. Metacomunity phylogenetics: separating the roles of environmental filters and historical biogeography. - *Ecol. Lett.* 13: 1290-1299.
- Lockwood, R. et al. 1998. Avian wingtip shape reconsidered: wingtip shape indices and morphological adaptations to migration. - *J. Avian Biol.*: 273–292.
- Lomolino, M. V. 1990. The target area hypothesis: The influence of island area on immigration rates of non-volant mammals. - *Oikos* 57: 297–300.
- Lomolino, M. V. et al. 2010. *Biogeography*. - Sinauer Associates, Inc.
- MacArthur, R. H. and Wilson, E. O. 1963. An equilibrium theory of insular zoogeography. - *Evolution* 17: 373–387.
- MacArthur, R. H. and Wilson, E. O. 1967. *The theory of island biogeography*. - Princeton University Press.
- Mayr, E. and Diamond, J. M. 1976. Birds on islands in the sky: Origin of the montane avifauna of Northern Melanesia. - *Proc. Natl. Acad. Sci.* 73: 1765–1769.
- Merow, C. et al. 2011. Developing dynamic mechanistic species distribution models: predicting bird-mediated spread of invasive plants across northeastern North America. - *Am. Nat.* 178: 30–43.
- Moyle, R. G. et al. 2005. Molecular phylogenetic analysis of the white-crowned forktail *Enicurus leschenaulti* in Borneo. - *J. Avian Biol.* 36: 96–101.
- Norberg, U. M. and Rayner, J. M. V. 1987. Ecological morphology and flight in bats (Mammalia; Chiroptera): wing adaptations, flight performance, foraging strategy and echolocation. - *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 316: 335–427.

- O'Hara, R. B. et al. 2002. Bayesian analysis of metapopulation data. - *Ecology* 83: 2408–2415.
- Ovaskainen, O. et al. 2010. Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. - *Ecology* 91: 2514–2521.
- Reddy, S. and Moyle, R. G. 2011. Systematics of the scimitar babblers (*Pomatorhinus*: *Timaliidae*): phylogeny, biogeography, and species-limits of four species complexes. - *Biol. J. Linn. Soc.* 102: 846–869.
- Rosenzweig, M. 1995. *Species diversity in space and time*. - Cambridge University Press.
- Rosindell, J. and Phillimore, A.B. 2011. A unified model of island biogeography sheds light on the zone of radiation. - *Ecol. Lett.* 14: 552-560.
- Rüger, N. et al. 2012. Functional traits explain light and size response of growth rates in tropical tree species. - *Ecology* 93: 2626–2636.
- Seidler, T. G. and Plotkin, J. B. 2006. Seed dispersal and spatial pattern in tropical trees. - *PLoS Biol.* 4: e344.
- Shmida, A. and Wilson, M. V. 1985. Biological determinants of species diversity. - *J. Biogeogr.* 12: 1–20.
- Sukumaran, J. et al. 2016. Machine learning biogeographic processes from biotic pattern: a trait-driven dispersal and diversification model with model-choice by simulation-trained discriminant analysis of principal components classification. - *Syst. Biol.* 65: 525-545.
- Triantis, K. A. et al. 2012. The island species-area relationship: biology and statistics. - *J. Biogeogr.* 39: 215-231.

- Accepted Article
- Triantis, K. A. et. al. 2015. Diversity regulation at macro-scales: species richness on oceanic archipelagos. - *Global Ecol. Biogeogr.* 24: 594-605.
- Valente, L. M. et. al. 2015. Equilibrium and non-equilibrium dynamics simultaneously operate in the Galápagos islands. - *Ecol. Lett.* 18: 844- 852.
- Webb, C. O. 2000. Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. - *Am. Nat.* 156: 145–155.
- Weeks, B. C. and Claramunt, S. 2014. Dispersal has inhibited avian diversification in Australasian archipelagos. - *Proc. R. Soc. B Biol. Sci.* 281: 20141257
- Wessel, P. and Smith, W. H. F. 1996. A global, self-consistent, hierarchical, high-resolution shoreline database. - *J. Geophys. Res. Solid Earth* 101: 8741–8743.
- Wilson, E. O. 1961. The nature of the taxon cycle in the melanesian ant fauna. - *Am. Nat.* 95: 169-193.

Note: The following data files and analysis code will be uploaded to dryad upon acceptance of the manuscript. If the reviewers would like to see these files for review purposes, a zipped archive can be downloaded from here:

https://dl.dropboxusercontent.com/u/10872221/dryad_submission.zip

Figure Legends

Figure 1

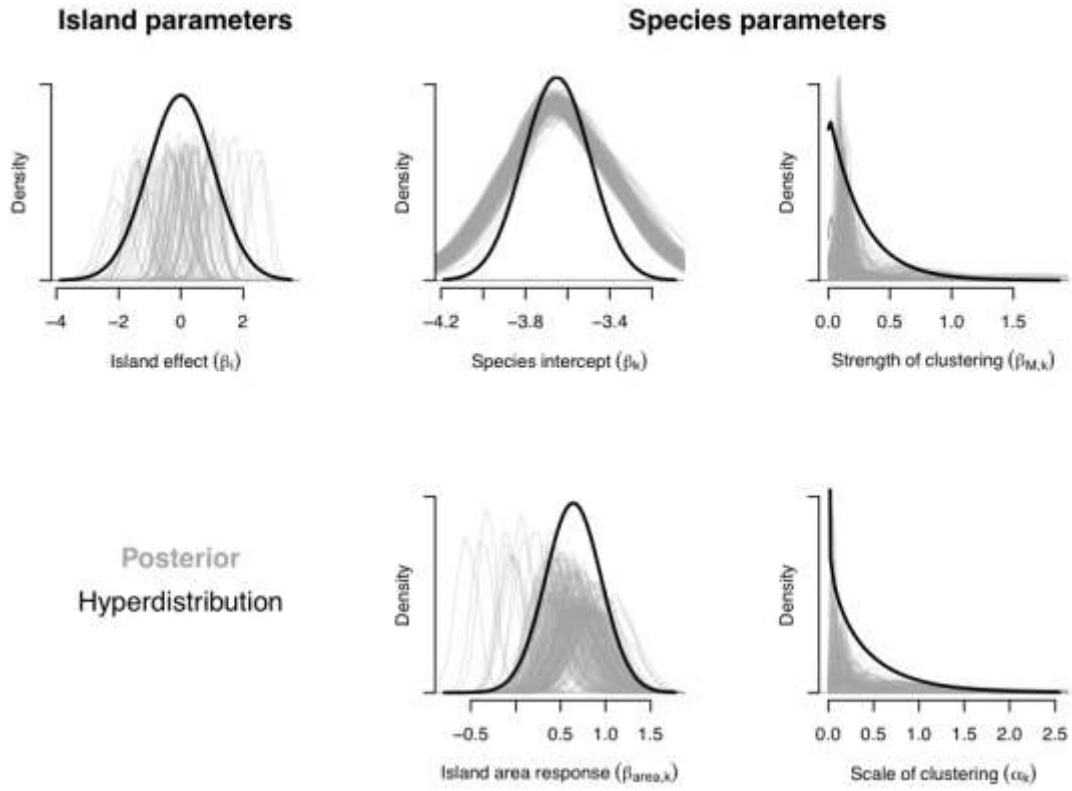


Figure 2

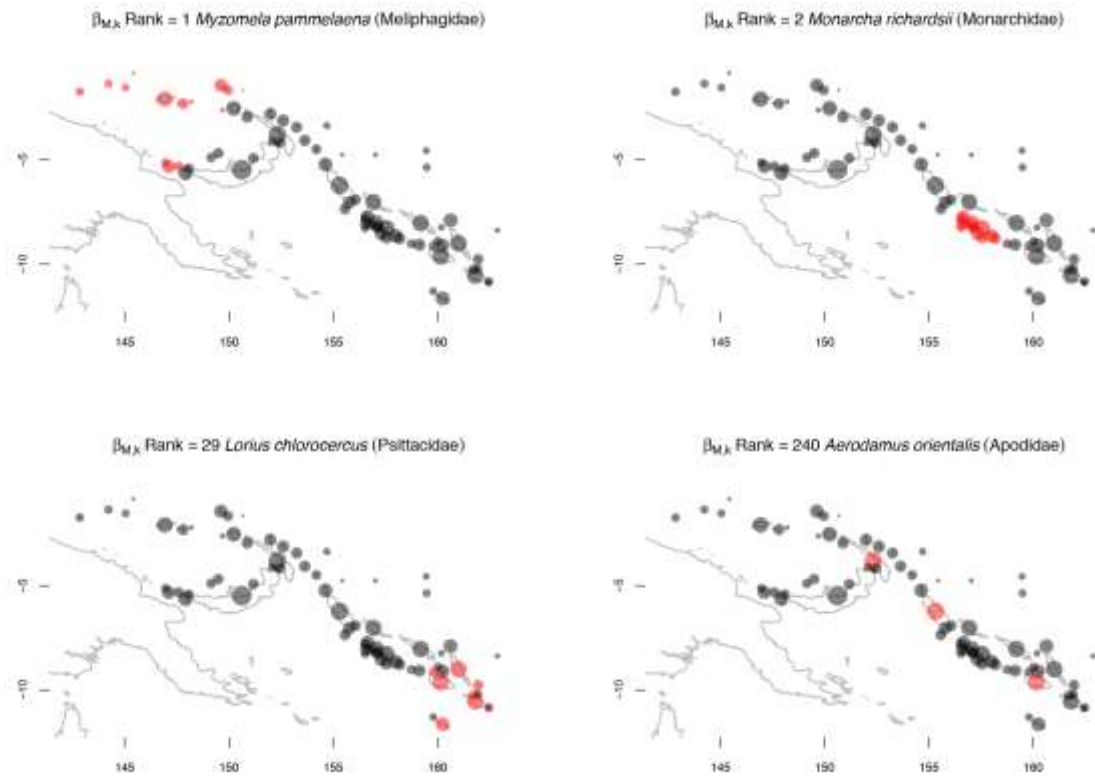


Figure 3

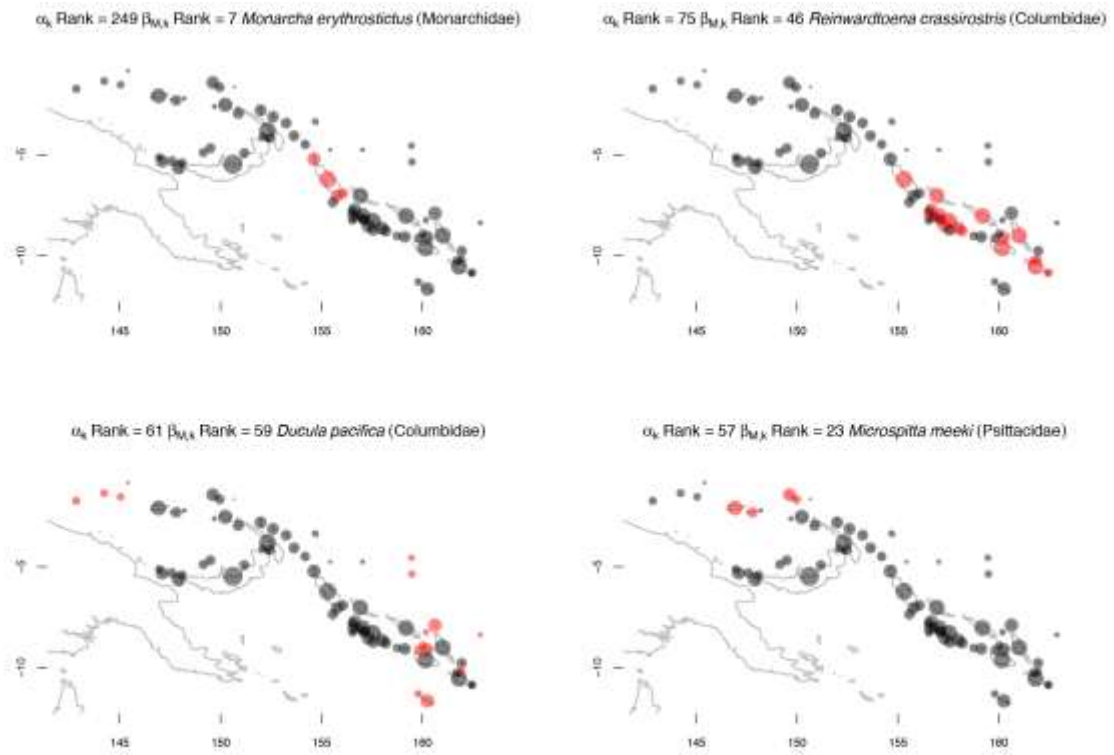


Figure 4

