

1 A new dispersal-informed null model for community ecology shows strong performance

2

3 Eliot T. Miller<sup>1</sup>

4

5 <sup>1</sup> E. T. Miller (eliotm@uidaho.edu), Dept of Biological Sciences, Univ. of Idaho,

6 Moscow, Idaho, 83844, USA.

7

8 **Author contribution statement** ETM conceived, designed, and executed this study

9 and wrote the manuscript. No other person is entitled to authorship.

10

11 **Abstract** Null models in ecology have been developed that, by maintaining some aspects  
12 of observed communities and repeatedly randomizing others, allow researchers to test for  
13 the action of community assembly processes like habitat filtering and competitive  
14 exclusion. Such processes are often detected using phylogenetic community structure  
15 metrics. When biologically significant elements, such as the number of species per  
16 assemblage, break down during randomizations, it can lead to high error rates. Realistic  
17 dispersal probabilities are often neglected during randomization, and existing models  
18 make the oftentimes empirically unreasonable assumption that all species are equally  
19 probable of dispersing to a given site. When this assumption is unwarranted, null models  
20 need to incorporate dispersal probabilities. I do so here, and present a dispersal null  
21 model (DNM) that strictly maintains species richness, and approximately maintains  
22 species occurrence frequencies and total abundance. I tested its statistical performance  
23 when used with a wide breadth of phylogenetic community structure metrics across 3,000  
24 simulated communities assembled according to neutral, habitat filtering, and competitive  
25 exclusion processes. The DNM performed well, exhibiting low error rates (both type I  
26 and II). I also implemented it in a re-analysis of a large empirical dataset, an abundance  
27 matrix of 696 sites and 75 species of Australian Meliphagidae. Although the overall  
28 signal from that study remained unchanged, it showed that statistically significant  
29 phylogenetic clustering could have been an artifact of dispersal limitations.

30

31 **Keywords** Null model · Community ecology · Community assembly · Phylogenetic  
32 metrics · Dispersal probability

33

## 34 INTRODUCTION

35 Null models in ecology are often used to test whether an assemblage of co-  
36 occurring species differs from what would be expected of a random assortment of co-  
37 occurring species (Gotelli and McGill 2006). Of particular interest here has been  
38 determining whether a community shows evidence that competition structures its  
39 constituent member species. How to define random in the null model context has been a  
40 matter of great contention since the debates of the 1970s and 1980s (Connor and  
41 Simberloff 1979; Diamond and Gilpin 1982; Connor and Simberloff 1983). Ideally, null  
42 models shuffle elements of the observed data related to the null hypothesis and preserve  
43 unrelated aspects of the observed data (e.g., which vs. how many species co-occur). Null  
44 models often take the form of repeated randomizations of an observed community data  
45 matrix (CDM), which are then compared to the observed CDM to detect non-random  
46 patterns of community assembly. Matters of contention include what elements of the  
47 CDM should be maintained (e.g., row and/or column sums, Diamond and Gilpin 1982),  
48 the most efficient algorithms for shuffling matrices (Miklós and Podani 2004), what  
49 metrics should be used to calculate patterns in the matrices (Stone and Roberts 1990), and  
50 which null models provide the best statistical performance (Gotelli 2000). How to  
51 incorporate abundance instead of simply presence-absence into these null models is  
52 another important research focus that has to date only received some attention (Hardy  
53 2008; Ulrich and Gotelli 2010; Miller et al. 2016).

54 Further complicating the field has been the rise of neutral models (Bell 2000;  
55 Hubbell 2001). Here, I follow Gotelli and McGill (2006), in considering neutral models  
56 as best used for testing whether per capita demographic rates differ between species. As

57 the focus of this paper is on testing whether species interactions are important, and  
58 assumes that species differ in their per capita demographic rates, this paper is about a new  
59 null model. That said, neutral models derive from the theory of island biogeography,  
60 which is governed by the countervailing forces of dispersal and extinction. Because this  
61 paper introduces a null model that incorporates dispersal, it falls closer to what Gotelli  
62 and McGill consider a process-based model that “crosses the line”. This is similar to a  
63 recent method incorporating speciation into null expectations (Pigot and Etienne 2015).

64 In this paper I define a CDM as a matrix with sites (e.g., quadrats, samples, plots)  
65 as rows and species as columns. The CDM as defined here can and, because of the  
66 additional detail afforded, ideally does incorporate relative or absolute abundance. During  
67 the null model debates, the focus was on demonstrating matrix-wide departures from  
68 expectations. In other words, the question was whether the entire CDM showed evidence  
69 of non-randomness. Accordingly, the metrics used to document the significance, or lack  
70 thereof, in community structural patterns were calculated per matrix, and generally dealt  
71 with presence-absence matrices (Schluter 1984; Stone and Roberts 1990). With the more  
72 recent focus on phylogenetic community structure (Webb 2000), the focus has shifted  
73 away from CDM-level patterns to assemblage-level (i.e. row-level) structural patterns.  
74 With this has come the introduction of a host of new phylogenetic community structure  
75 metrics, many of which incorporate abundance (Faith 1992; Webb 2000; Helmus et al.  
76 2007; Cadotte et al. 2010; Kembel et al. 2010; Miller et al. 2013). These metrics quantify  
77 the relatedness of co-occurring species, with the assumption being that closely related co-  
78 occurring species provide evidence of habitat filtering, while distantly related co-  
79 occurring species provide evidence of competitive exclusion (Webb 2000). The null

80 model introduced in this paper is intended for use in assessing assemblage-level patterns,  
81 and its statistical performance was tested here in that context. Its relevance to matrix-  
82 level patterns of co-occurrence is not tested here.

83 Many null models shuffle species' presences or abundances within rows, allowing  
84 species to occur with equal probability in the randomized matrices. Various  
85 improvements have been developed, including models that maintain species' occurrence  
86 frequencies (Gotelli 2000), both species richness and occurrence frequency (Miklós and  
87 Podani 2004), and elements of species' abundance distributions (Hardy 2008). Such  
88 models have been shown to reduce type I errors (Gotelli 2000; Miller et al. 2016). Miller  
89 et al. (2016) showed that a further reduction in error rates can be achieved by creating a  
90 CDM *de novo* that mimics regional dispersal pressures on a "local" community (the  
91 CDM). However, in all of these null models quadrats are disassociated from their  
92 geographic realities. Current null models randomize sites but do not take into account  
93 dispersal probabilities between sites in the randomization process. Thus, a species from a  
94 distant site is just as likely to be placed in a simulated site as a species from a nearby site.  
95 The 3<sup>rd</sup> null model introduced by Hardy (2008) made strides towards addressing this issue,  
96 but it does not maintain quadrat species richness and requires transect-like sampling.

97 Dispersal, even over short distances like those across a forest plot, can greatly  
98 influence which species occur where (MacArthur and Wilson 1967; Laurance et al.  
99 2002). Over large distances, such as those between grid cells in the desert of inland  
100 Australia and the rainforests of the coast, it can be difficult to parse the influence of  
101 community assembly processes such as habitat filtering versus that of physical dispersal  
102 limitation. For example, if a certain clade within the study system has diversified within a

103 small region of the continent, significant phylogenetic clustering in that region is not  
104 necessarily attributable to habitat filtering, and might be due entirely to a failure of these  
105 species to disperse to other regions. A more flexible, dispersal-informed null model is a  
106 clear research priority that should prove useful to empirical researchers. Having a null  
107 model that respects quadrat-specific dispersal probabilities, while also maintaining  
108 quadrat-specific species richness and species' matrix-wide occurrence probabilities  
109 would assist with teasing apart such community assembly subtleties. In this paper, I  
110 develop such a null model and test its statistical behavior and performance.

111

## 112 METHODS

### 113 *Description of the dispersal null model*

114 The dispersal null model (DNM) takes as input the original CDM,  $\mathbf{C}_O$ , and a  
115 symmetrical matrix,  $\mathbf{D}$ , that describes the distances among quadrats. It provides as output  
116 a randomized CDM,  $\mathbf{C}_R$ , with the same dimensions as  $\mathbf{C}_O$ . Distances in  $\mathbf{D}$  can be  
117 geographic, climatic, or otherwise of the researcher's choice. I then define the  
118 randomization procedure as follows. Let  $i_O$  be a row (quadrat) from  $\mathbf{C}_O$ , and  $i_R$  be the  
119 corresponding row from  $\mathbf{C}_R$ . Let  $j$  be any other row from  $\mathbf{C}_O$  where  $i_O \neq j$ . Let  $SR(n)$  be  
120 the species richness of some row  $n$  from a CDM. Then, for each  $i_O$ , some  $j$  is sampled  
121 with a probability determined by the reciprocal of its distance from  $i_O$ . A species is then  
122 sampled from  $j$  with a probability proportional to its abundance in  $j$ , placed into  $i_R$ , and  
123 assigned the same abundance as in  $j$ . Then this process is repeated until  $SR(i_O) = SR(i_R)$ . If  
124 for any  $j$ , the species sampled has already been settled into  $i_R$ , then it is discarded and  
125 another  $j$  and corresponding species sampled. The model goes on to repeat the process for

126 all quadrats  $i$ , which results in a filled  $C_R$ . The need to propose and then potentially reject  
127 species necessitates the use of a serial loop, which causes the DNM to run more slowly  
128 than simpler randomization procedures. I refer to this form of model as DNM<sub>1</sub>.

129 A slight variation of the model, DNM<sub>2</sub>, generates a normal distribution (with a  
130 standard deviation of 1, rounded to whole numbers and values  $< 1$  rounded up to 1)  
131 centered on the abundance of the sampled species in  $j$ . Rather than directly assigning a  
132 species the same abundance it had in  $j$ , a value is sampled from this distribution. This  
133 causes a slight slowdown (~5% longer) in null model performance, but theoretically  
134 results in additional exploration of null biological space.

135 A considerable variant of the model (DNM<sub>3</sub>) does not incorporate species'  
136 abundances in  $j$  into the probability that they will be sampled. Instead, all species present  
137 in  $j$  have an equal probability of being sampled and placed into  $i_R$ . Thus a species'  
138 proximity to but not its abundance in  $j$  influences its probability of settling in  $i_R$ . The  
139 biological meaning here is changed from DNM<sub>1</sub>. With DNM<sub>3</sub>, any factors influencing  
140 individual species' abundances within a quadrat, such as competition, are less influential  
141 in the randomized matrices. While DNM<sub>1</sub> can help researchers test for non-random  
142 patterns of community assembly given realistic dispersal pressures, where both  
143 abundance and distance to a focal quadrat matters, DNM<sub>3</sub> would be more pertinent if the  
144 focus was on the influence of competition given distance dispersal limitations only. That  
145 is, if a researcher thought it possible that species might be rare in observed quadrats as a  
146 consequence of competition, then DNM<sub>3</sub> would randomize those structures in the  
147 observed data and allow that hypothesis to be tested. With DNM<sub>3</sub>, I recommend that

148 researchers assign species abundances by sampling from the vector of observed non-zero  
149 abundances in the original CDM.

150 The three forms of DNM are available in the R package *metricTester* (Miller et al.  
151 2016). Though the DNM runs more slowly than traditional matrix randomization,  
152 *metricTester* utilizes multicore processing and thus can manage a reasonable number of  
153 randomizations of an observed CDM.

154

155 *Statistical behavior and performance of the dispersal null model*

156 As explained above, the DNMs strictly maintain species richness. I was interested  
157 in how well they maintain species' occurrence frequencies and total abundances in the  
158 randomized matrices. To test this, I created a CDM with the `simulateComm` function in  
159 *metricTester*. The CDM contained 100 quadrats and species. Species richness varied  
160 from 10 to 34, with each value represented four times. Species were assigned abundances  
161 by drawing from a log-normal distribution with mean=2 and SD=1 (on a log scale). I then  
162 randomized the CDM 20 times with DNM<sub>1</sub>, calculating species' occurrence frequencies  
163 and total abundances after each randomization. I took the mean of these observations and  
164 compared it to observed values from the observed CDM. I performed the same procedure  
165 but randomized the CDM 20 times with DNM<sub>3</sub>, in this case setting  
166 `abundance.assigned` to "overall".

167 I used identical methodology as Miller et al. (2016) to test the performance of the  
168 DNMs. Appendix S3 of that paper provides schematic illustration of the methodology.  
169 Thus, I used the `multiLinker` function with the following parameters. I set `no.taxa` to  
170 100, `arena.length` to  $\sqrt{(10^5)}$ , `mean.log.individuals` to 3.5, `length.parameter`



171 to 1000, `sd.parameter` to 40, `max.distance` to 20, `proportion.killed` to 0.2,  
172 `competition.iterations` to 60, `no.quadrats` to 20, `quadrat.length` to  
173  $\sqrt{(1000)}$ , `concat.by` to “both”, and `randomizations` to 1000. The simulation and  
174 performance testing process is as follows. (1) Generate a phylogeny describing the  
175 relationships among the species that will be involved in the simulation. (2) Generate  
176 realistic spatial arenas of 2,000-4,000 individuals according to either random, habitat  
177 filtering or competitive exclusion community assembly rules. In the habitat filtering  
178 simulations, individuals are placed in the arena according to spatial preferences that  
179 exhibit Brownian motion evolution along the simulated phylogeny. The competitive  
180 exclusion simulations begin with the random simulation. Individuals in genetically  
181 clustered areas of the arena then compete, resulting in the removal of some of these most  
182 closely related individuals. Removed individuals are then replaced according to their  
183 initial arena-wide abundances (i.e. simulating regional dispersal pressures), and the entire  
184 process is repeated for 60 generations. (3) Create CDMs (one per spatial simulation) by  
185 placing 20 quadrats in each arena and determining which individuals fall within each  
186 quadrat. (4) Calculate a wide breadth of phylogenetic community structure metrics on the  
187 observed quadrats. These metrics were PAE (phylogenetic abundance evenness),  $H_{AED}$   
188 (community abundance-weighted evolutionary distinctiveness), IAC (imbalance of  
189 abundance),  $E_{AED}$  (equitability abundance-weighted evolutionary distinctiveness),  $H_{ED}$   
190 (community evolutionary distinctiveness),  $E_{ED}$  (equitability evolutionary distinctiveness),  
191 MNTD (mean nearest taxon distance), AW MNTD (abundance-weighted MNTD), PD  
192 (phylogenetic diversity), PD<sub>c</sub> (PD not including the root), MPD (mean pairwise  
193 phylogenetic distance), interspecific MPD (interspecific abundance-weighted MPD),

194 intraspecific MPD (intraspecific abundance-weighted MPD), complete MPD (complete  
195 abundance-weighted MPD), and PSE (phylogenetic species evenness) (Faith 1992; Webb  
196 2000; Helmus et al. 2007; Cadotte et al. 2010; Kembel et al. 2010; Miller et al. 2013). (5)  
197 Use the Euclidean distances between quadrat centroids to randomize each CDM 1,000  
198 times with DNM<sub>1</sub>, and 1,000 times with DNM<sub>3</sub>. Note that this happens across all spatial  
199 simulations, such that after the three spatial simulations, then 1,000 randomizations of  
200 each observed CDM with each of DNM<sub>1</sub> and DNM<sub>3</sub>, the result is a collection of 6,000  
201 randomly assembled CDMs and three observed CDMs. (6) After each randomization,  
202 calculate all phylogenetic community structure metrics across the assembled CDM and  
203 retain. (7) Concatenate the randomized metrics by the quadrat from which they come, and  
204 derive a standardized effect score (SES) per observed quadrat as the difference between  
205 the observed and the mean of the randomized scores (at that quadrat) divided by the SD  
206 of the randomized scores (at that quadrat). (8) Per CDM, use a Wilcoxon signed-rank test  
207 to compare the distribution of SES scores to zero. Record a type I error if the distribution  
208 differs significantly from zero for the random community assembly, and a type II error if  
209 the distribution was not significantly less or greater than zero in the habitat filtering or  
210 competitive exclusion simulation, respectively. (9) Repeat the entire process 1,000 times  
211 (i.e. run 1,000 each of the neutral, habitat filtering, and competitive exclusion spatial  
212 simulations, with the resulting CDMs randomized 1,000 times each). In this way, per  
213 metric per null model, I calculated an average type I error rate as the proportion of 1,000  
214 random communities whose SES distribution differed significantly from zero, and an  
215 average type II error rate as the mean of the proportion of communities from either the

216 1,000 habitat filtering or the 1,000 competitive exclusion simulations whose SES  
217 distributions did not differ as expected from zero.

218

219 *Testing the dispersal null model on an empirical dataset*

220 I was also interested in whether DNM<sub>1</sub> could be readily applied to an empirical  
221 dataset, and what influence it would have on the interpretation of results. Miller et al.  
222 (2013) found strong evidence that phylogenetic niche conservatism shapes which  
223 lineages of Australian honeyeaters (Meliphagidae) occur where. Meliphagidae  
224 assemblages showed increasing phylogenetic clustering along a gradient of decreasing  
225 precipitation away from the ancestral state, with statistically significant clustering  
226 observed in the driest sites. Miller et al. (2013) controlled for spatial auto-correlation and  
227 found that the significance of the overall relationship was unaffected by such auto-  
228 correlation, but the null models they used allowed any species to occur in any quadrat.

229 To better account for the influence of dispersal limitations, I re-calculated the  
230 significance of the observed phylogenetic community structure metrics in Meliphagidae  
231 assemblages as compared with expectations from DNM<sub>1</sub> (both MPD and interspecific  
232 MPD). For the matrix describing distances between quadrats, I used: 1) great circle  
233 distances calculated with the Haversine formula; 2) the Euclidean distances after a  
234 principal components analysis of the 19 bioclim ecological variables (Hijmans et al.  
235 2005), scaled and centered; 3) the product of these two distance matrices. I calculated the  
236 significance of the observed metrics against 1,500 randomly assembled matrices with  
237 each of the distance matrices. Because results were qualitatively identical with any of the

238 distance matrices, I chose to use the product of the geographic and climate distance  
239 matrices and re-ran the analysis with 30,000 matrix randomizations.

240

## 241 RESULTS

### 242 *Statistical behavior and performance of the dispersal null model*

243 With DNM<sub>1</sub>, species' mean occurrence frequencies across 20 randomized  
244 matrices were correlated with their occurrence frequency in the observed CDM ( $r^2 =$   
245 0.45,  $p < 0.001$ ). The same was true for their total abundance in the randomized CDMs  
246 ( $r^2 = 0.79$ ,  $p < 0.001$ ). With DNM<sub>3</sub>, species' randomized occurrence frequencies were  
247 correlated with observed frequencies ( $r^2 = 0.84$ ,  $p < 0.001$ ), but species' total randomized  
248 abundances were only weakly correlated with observed values ( $r^2 = 0.18$ ,  $p < 0.001$ ).

249 Averaging across its performance with all tested metrics, DNM<sub>1</sub> exhibited a  
250 22.8% error rate (15.1% type I, 30.5% type II). However, the bulk of these errors can be  
251 attributed to a few metrics (Table 1), namely IAC,  $H_{AED}$  and, to a lesser extent, PAE and  
252  $E_{AED}$ . With the best performing metric, MPD, the type I and II error rates were 0.1 and  
253 13.8%, respectively. In terms of other metrics that performed reasonably well with  
254 DNM<sub>1</sub>, MPD, PD, interspecific MPD, MNTD and AW MNTD all exhibited strong power  
255 to detect habitat filtering, but much less power to detect competitive exclusion. The  
256 opposite was true of intraspecific MPD and  $E_{AED}$ .

257 DNM<sub>3</sub> also performed favorably (Table 2). As compared with DNM<sub>1</sub>, it showed  
258 increased power to detect simulated community assembly processes, particularly  
259 competitive exclusion, with only a slight increase in type I error rates (across all metrics  
260 16.7% type I, 23.2% type II). MPD, again the best performing metric, exhibited an

261 overall error rate of only 2.2%. Both forms of PD also performed well, though they  
262 showed less power to detect the effects of competitive exclusion.  $H_{ED}$  and  $E_{ED}$ , which did  
263 not perform well in Miller et al. (2016), performed better than all other tested metrics  
264 except MPD and PD.

265

### 266 *Testing the dispersal null model on an empirical dataset*

267 As in Miller et al. (2013), there was strong signal of increasing phylogenetic  
268 clustering along a gradient of decreasing precipitation. For instance, using the product of  
269 the climatic and geographic distance matrices, SES MPD and SES interspecific MPD  
270 were positively correlated with  $\log_{10}$  mean annual precipitation ( $r^2 = 0.14$  and  $p < 0.001$ ,  
271 and  $r^2 = 0.41$  and  $p < 0.001$ , respectively). However, the significance of individual  
272 quadrat deviations beyond expectations can and did change with  $DNM_1$ . Only one and  
273 two of the quadrats were considered significantly phylogenetically clustered with MPD  
274 and interspecific MPD, respectively. Similarly, two and zero quadrats were considered  
275 significantly phylogenetically overdispersed with MPD and interspecific MPD,  
276 respectively. Results were qualitatively identical with any of the distance matrices (i.e.,  
277 few individual quadrats deviated beyond statistical expectations). Thus, when compared  
278 against a null model that simulates realistic dispersal probabilities, the overall pattern of  
279 increasing phylogenetic clustering along a gradient of decreasing precipitation did not  
280 change, but few if any individual sites deviated beyond null expectations.

281

282 DISCUSSION

283 Null models in ecology have been contentious for well over 40 years. Many  
284 technical improvements have been developed over this time, including models that  
285 account for species-specific patterns of spatial distributions (Roxburgh and Chesson  
286 1998; Roxburgh and Matsuki 1999). A great deal of sound reasoning and guidance has  
287 also been offered (Gotelli and Graves 1996; Gotelli 2000; Gotelli and Entsminger 2001;  
288 Ulrich and Gotelli 2010). But to my knowledge, no null model that maintains realistic,  
289 quadrat-specific dispersal pressures has yet been developed. In this paper I introduced  
290 and tested such a dispersal null model (DNM). This is similar to recent efforts to include  
291 other biologically important processes in the null model. For instance, Pigot and Etienne  
292 (2015) showed that incorporating allopatric speciation into null models erases signatures  
293 in phylogenetic community structure that were previously considered to represent  
294 competitive exclusion. Indeed, those authors and other recent reviews (Gotelli and Ulrich  
295 2012) highlighted the need for a DNM.

296 To date, the conceptual link remains weak between neutral models for community  
297 assembly and null models for phylogenetic community structure. The DNM and other  
298 recent null models provide the foundation for a bridge to link the ideas, but that bridge  
299 remains to be built. Future researchers will need to merge ideas of ecological sorting with  
300 those of evolutionary processes, e.g. competitive exclusion versus character displacement  
301 (or allopatric speciation). Ultimately, a model linking dispersal, speciation and extinction  
302 may allow researchers to untangle the influences of these processes in community  
303 assembly.

304 In this paper, the overall error rates of DNM<sub>1</sub> and DNM<sub>3</sub> were 22.8% and 19.95%,  
305 respectively. In a previous test of null model performance (Miller et al. 2016), across all

306 metrics, the regional model showed the lowest overall error rates (8%), followed by the  
307 3x, 2x, trial swap, independent swap, and frequency concatenated by richness models,  
308 which all had overall error rates of approximately 19%. The 1s, richness and frequency  
309 concatenated by quadrat models showed error rates of 25-27%. This would suggest the  
310 DNM was outperformed by a number of other null models. However, error rates  
311 calculated across all metrics are misleading in this case, in that some metrics performed  
312 quite well with the DNM, while others performed very poorly. For instance, as compared  
313 with the regional null model in Miller et al. (2016), where MPD showed 6.2 and 3.2%  
314 type I and II error rates, respectively, with DNM<sub>1</sub> these rates were 0.1 and 13.8% (Table  
315 1), while with DNM<sub>3</sub> they were 3.2 and 1.2% (Table 2).

316         As compared to other null models, the decrease in type I error rates for the DNM  
317 is attributable to the fact that simulated quadrats closely resemble observed quadrats in  
318 species richness and composition. To deviate beyond expectations, observed quadrats  
319 need to show strong signals in terms of co-occurrence and/or, for abundance-weighted  
320 metrics, the relative abundances of co-occurring species. As compared with DNM<sub>1</sub>, the  
321 increased power of DNM<sub>3</sub> to detect competitive exclusion seems to be because the latter  
322 does not incorporate a species' relative abundance into its probability of being settled in  
323 simulated quadrats. So, if a species is rare in a given quadrat as a function of competition  
324 with co-occurring species, this element is randomized in the simulated CDMs, allowing  
325 appropriate rejection of the null hypothesis. Conversely, there was a slight overall  
326 decrease in power with DNM<sub>3</sub> to detect habitat filtering. On the surface it would seem  
327 this is because if a species' abundance attenuates away from the center of its distribution,  
328 the DNM<sub>3</sub> may occasionally settle the species at high abundances towards the periphery

329 of its range. In practice, however, most abundance-weighted metrics actually showed  
330 increased power to detect habitat filtering with DNM<sub>3</sub>, and the overall decrease in power  
331 in this respect as compared with DNM<sub>1</sub> can be ascribed to  $E_{AED}$  and  $H_{AED}$  showing  
332 striking decreases in power to detect habitat filtering.

333 Notably, the non-abundance-weighted metrics  $H_{ED}$  and  $E_{ED}$  performed reasonably  
334 well with both null models tested here, particularly DNM<sub>3</sub>. These metrics had previously  
335 shown poor statistical performance (Miller et al. 2016). Overall, other non-abundance-  
336 weighted metrics (MNTD, PD, MPD) also outperformed abundance-weighted forms with  
337 DNM<sub>3</sub>. As noted above, abundance-weighted metrics were better than non-abundance-  
338 weighted metrics at detecting habitat filtering with DNM<sub>3</sub>, so this overall decrease in  
339 abundance-weighted metric performance with DNM<sub>3</sub> is attributed entirely to a stark  
340 increase in their type I error rates (Table 2). Since DNM<sub>3</sub> maintains species' occurrence  
341 frequencies but not abundances, this increase in type I error rates is to be expected.  
342 Abundance-weighted metrics are particularly driven by changes in the abundance of  
343 distantly related species. So, for instance, a monotypic genus that occurred regularly but  
344 at low abundance in an observed CDM might occur regularly but at high abundance in  
345 simulated CDMs, thereby triggering false positives.

346 The choice of which null model to use cannot be informed by statistical  
347 performance alone (Gotelli 2000). The null hypothesis to be tested must inform the  
348 choice as well. When used on an empirical dataset where dispersal limitations almost  
349 certainly influence probability of arrival at a site (Miller et al. 2013), DNM<sub>1</sub> did not  
350 change previous results that co-occurring Meliphagidae species are more closely related  
351 in arid areas. However, the number of quadrats considered significantly phylogenetically



352 clustered was much reduced. Thus, while the overall relationship remains unchanged, the  
353 strong signal of Meliphagidae phylogenetic clustering in arid regions would not have  
354 been detected had that study focused on single sites and their statistical significance,  
355 rather than the slope of unstandardized MPD across climate gradients.

356 As programmed here, the DNM runs more slowly than previously defined null  
357 models, since it requires the use of an indeterminate loop to create each cell in the  
358 random matrix, rather than using simple matrix shuffling. Fortunately, the model is  
359 available in a multithreaded version, which permits parallel matrix randomizations and a  
360 corresponding decrease in total computing time. For instance, constructing 30,000, 696  
361 quadrat by 75 species matrices against which to compare the observed Meliphagidae  
362 CDM took ~6 hrs on a MacBook Pro with a 2.5 GHz processor and 16 GB RAM.

363 The DNM should be applicable to a wide breadth of research questions. There is  
364 one situation, however, under which the DNM will fail to run. If any quadrats within an  
365 observed community data matrix (CDM) contain more species than are available in the  
366 sum of other quadrats, then the model cannot run to completion. For example, if  $SR(i_O) =$   
367 15, and the remaining quadrats contain only 12 unique species in total, then the model  
368 would loop indefinitely trying to find 15 species to fill  $i_R$ . The DNM function has an  
369 internal check for this problem, and returns an error message if it is manifest in a CDM.  
370 Aside from this hopefully unusual empirical situation, if users are able to generate a  
371 distance matrix summarizing dispersal probabilities between quadrats (e.g., geographic or  
372 climatic distances between sampling sites), DNM will almost certainly be easy to  
373 implement. It strictly maintains species richness, approximately maintains species

374 occurrence frequencies, overall abundances, and realistic dispersal probabilities, and  
375 showed suitable performance when compared against simulated community processes.

376

#### 377 ACKNOWLEDGEMENTS

378 Funding for this study was provided by the National Science Foundation (DBI-1402506).

379 Simulations were run on the Grethor cluster at the University of Missouri, St. Louis, and

380 the Lewis cluster at the University of Missouri, Columbia. Cody Hinchliff, Matthew

381 Pennell, Luke Harmon and three anonymous reviewers provided valuable input during

382 the writing of the manuscript.

383

#### 384 REFERENCES

385 Bell G (2000) The distribution of abundance in neutral communities. *Am Nat* 155:606–  
386 617.

387 Cadotte MW, Jonathan Davies T, Regetz J, et al (2010) Phylogenetic diversity metrics for  
388 ecological communities: integrating species richness, abundance and evolutionary  
389 history. *Ecol Lett* 13:96–105.

390 Connor EF, Simberloff D (1979) The assembly of species communities: chance or  
391 competition? *Ecology* 60:1132–1140.

392 Connor EF, Simberloff D (1983) Interspecific competition and species co-occurrence  
393 patterns on islands: null models and the evaluation of evidence. *Oikos* 41:455–  
394 465.

- 395 Diamond JM, Gilpin ME (1982) Examination of the “Null” Model of Connor and  
396 Simberloff for Species Co-Occurrences on Islands. *Oecologia* 52:64–74.
- 397 Faith DP (1992) Conservation evaluation and phylogenetic diversity. *Biol Conserv* 61:1–  
398 10.
- 399 Gotelli NJ (2000) Null model analysis of species co-occurrence patterns. *Ecology*  
400 81:2606–2621.
- 401 Gotelli NJ, Entsminger GL (2001) Swap and fill algorithms in null model analysis:  
402 rethinking the knight’s tour. *Oecologia* 129:281–291.
- 403 Gotelli NJ, Graves GR (1996) Null models in ecology. Smithsonian Institution Press,  
404 Washington
- 405 Gotelli NJ, McGill BJ (2006) Null versus neutral models: what’s the difference?  
406 *Ecography* 29:793–800.
- 407 Gotelli NJ, Ulrich W (2012) Statistical challenges in null model analysis. *Oikos* 121:171–  
408 180.
- 409 Hardy OJ (2008) Testing the spatial phylogenetic structure of local communities:  
410 statistical performances of different null models and test statistics on a locally  
411 neutral community. *J Ecol* 96:914–926.
- 412 Helmus MR, Bland TJ, Williams CK, Ives AR (2007) Phylogenetic measures of  
413 biodiversity. *Am Nat* 169:E68–E83.

- 414 Hijmans RJ, Cameron SE, Parra JL, et al (2005) Very high resolution interpolated climate  
415 surfaces for global land areas. *Int J Climatol* 25:1965–1978.
- 416 Hubbell SP (2001) *The unified neutral theory of biodiversity and biogeography*.  
417 Princeton University Press, Princeton, NJ
- 418 Kembel SW, Cowan PD, Helmus MR, et al (2010) Picante: R tools for integrating  
419 phylogenies and ecology. *Bioinformatics* 26:1463–1464.
- 420 Laurance WF, Lovejoy TE, Vasconcelos HL, et al (2002) Ecosystem decay of  
421 Amazonian forest fragments: a 22-year investigation. *Conserv Biol* 16:605–618.
- 422 MacArthur RH, Wilson EO (1967) *The theory of island biogeography*. Princeton  
423 University Press, Princeton, NJ
- 424 Miklós I, Podani J (2004) Randomization of presence-absence matrices: comments and  
425 new algorithms. *Ecology* 85:86–92.
- 426 Miller ET, Farine DR, Trisos CH (2016) Phylogenetic community structure metrics and  
427 null models: a review with new methods and software. *Ecography* In press. doi:  
428 10.1111/ecog.02070
- 429 Miller ET, Zanne AE, Ricklefs RE (2013) Niche conservatism constrains Australian  
430 honeyeater assemblages in stressful environments. *Ecol Lett* 16:1186–1194.
- 431 Pigot AL, Etienne RS (2015) A new dynamic null model for phylogenetic community  
432 structure. *Ecol Lett* 18:153–163.

- 433 Roxburgh SH, Chesson P (1998) A new method for detecting species associations with  
434 spatially autocorrelated data. *Ecology* 79:2180–2192.
- 435 Roxburgh SH, Matsuki M (1999) The statistical validation of null models used in spatial  
436 association analyses. *Oikos* 85:68–78.
- 437 Schluter D (1984) A variance test for detecting species associations, with some example  
438 applications. *Ecology* 65:998–1005.
- 439 Stone L, Roberts A (1990) The checkerboard score and species distributions. *Oecologia*  
440 85:74–79.
- 441 Ulrich W, Gotelli NJ (2010) Null model analysis of species associations using abundance  
442 data. *Ecology* 91:3384–3397.
- 443 Webb CO (2000) Exploring the phylogenetic structure of ecological communities: an  
444 example for rain forest trees. *Am Nat* 156:145–155.
- 445
- 446

447 Table 1. Error rates of the dispersal null model (DNM<sub>1</sub>) with the tested phylogenetic  
 448 community structure metrics. Type I errors were defined as a randomly assembled  
 449 community deviating beyond statistical expectations. For example, 0.1% of 1,000 total  
 450 randomly assembled communities deviated beyond expectations. Type II error rates were  
 451 defined as communities assembled according to either habitat filtering or competitive  
 452 exclusion not being considered significantly phylogenetically structured. Metrics are  
 453 ordered from best- to worst-performing according to the average of the type II and type I  
 454 error rates.

Metric	Type I error rate	Type II error rate – habitat filtering	Type II error rate – competitive exclusion	Overall error rate
MPD	0.10	5.65	21.90	6.94
Complete MPD	0.59	13.18	19.03	8.35
PSE	0.59	13.18	19.03	8.35
PD	0.20	0.00	37.26	9.42
PD <sub>c</sub>	0.20	0.00	37.26	9.42
Intraspecific				
MPD	6.24	21.21	5.65	9.84
$E_{ED}$	0.99	11.99	39.84	13.45
$H_{ED}$	0.99	11.99	39.84	13.45
Interspecific				
MPD	0.89	8.23	47.67	14.42
MNTD	1.09	0.00	58.08	15.06

---

AW MNTD	1.98	0.30	73.74	19.50
$E_{\text{AED}}$	8.72	65.71	14.87	24.50
PAE	3.17	51.04	98.12	38.88
$H_{\text{AED}}$	100.00	0.00	100.00	75.00
IAC	100.00	100.00	0.89	75.22

---

455

456

457 Table 2. Error rates of the dispersal null model (DNM<sub>3</sub>) with the tested phylogenetic  
 458 community structure metrics. Type I errors were defined as a randomly assembled  
 459 community deviating beyond statistical expectations. For example, 0.1% of 1,000 total  
 460 randomly assembled communities deviated beyond expectations. Type II error rates were  
 461 defined as communities assembled according to either habitat filtering or competitive  
 462 exclusion not being considered significantly phylogenetically structured. Metrics are  
 463 ordered from best- to worst-performing according to the average of the type II and type I  
 464 error rates.

Metric	Type I error rate	Type II error rate – habitat filtering	Type II error rate – competitive exclusion	Overall error rate
MPD	3.15	2.34	0.00	2.16
PD	1.32	0.00	15.26	4.48
PD <sub>c</sub>	1.32	0.00	15.46	4.53
MNTD	2.03	0.00	40.39	11.11
<i>H</i> <sub>ED</sub>	2.14	9.77	33.47	11.88
<i>E</i> <sub>ED</sub>	2.14	9.87	33.47	11.90
AW MNTD	24.62	1.32	26.45	19.25
Intraspecific				
MPD	34.08	10.99	0.00	19.79
Complete MPD	35.61	10.27	0.00	20.37
PSE	35.61	10.27	0.00	20.37
Interspecific	37.64	6.82	0.00	20.52



---

MPD				
IAC	11.19	96.64	6.41	31.36
$E_{\text{AED}}$	20.45	81.59	17.90	35.10
PAE	36.93	56.66	34.18	41.17
$H_{\text{AED}}$	2.75	94.40	78.54	44.61

---