

Supplementary material for: Hidden state models improve state-dependent diversification approaches, including biogeographical models

Daniel S. Caetano^{1,3}, Brian C. O'Meara², and Jeremy M. Beaulieu¹

¹Department of Biological Sciences, University of Arkansas, Fayetteville AR 72701.

²Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville TN 37996-1610

³Author for correspondence: Daniel S. Caetano, Email: dcaetano@uark.edu

TABLE OF CONTENTS

List of Tables

Table S1: Description of scenarios and parameter values used to simulate the data.

Table S2: Additional 17 models used in empirical study of conifers.

Table S3: Description of scenarios and parameter values used to simulate phylogenetic trees and range distributions under the GeoSSE+extirpation model.

List of Figures

Figure S1: Scheme of the transition rates between rate classes (RC0 to RC4) used for the simulation scenarios with multiple rate classes (Sims B, C and D).

Figure S2: Proportion of widespread lineages on trees simulated under scenarios E and F.

Figure S3: Summary of model support for simulated scenarios A to D.

Figure S4: Summary of model support for simulated scenarios E to H.

Figure S5: Accuracy of turnover and extinction fraction estimates for simulations scenarios A to D.

Figure S6: Accuracy of net diversification estimates for simulations scenarios A to D.

Figure S7: Results for relative net diversification rates and Akaike model weights (AICw) for simulation scenarios B and C.

Figure S8: Distribution of Akaike weights for the model set fitted to simulation scenarios ext_A to ext_D.

Figure S9: Distribution of parameter values across 100 simulation replicates for each of the scenarios ext_A to ext_D.

Section 1 - Testing the effect of heterogeneous rates on GeoSSE models

Section 2 - Extended simulation results

Section 3 - Performance of GeoSSE+extirpation models

Table S1: Description of scenarios and parameter values used to simulate the data. Scenarios A to E are instances of the original GeoSSE model and GeoHiSSE models with varying number of rate categories. Scenarios F to H are comprised of different models. Scenario F is a custom extension of the GeoSSE model allowing anagenetic transitions (i.e., jumps) between the endemic ranges. Scenario G has only two endemic areas A and B (see more information in Magnuson-Ford and Otto 2012). Scenario H is not a joint tree and trait model and follow similar procedures as used by Rabosky and Goldberg (2015).

Scenario	Model	Rate class	Parameters							Simulation stop criteria
			sAB	sA	sB	xA	xB	dA	dB	
A	GeoSSE	0	0.1	0.1	0.2	0.03	0.03	0.05	0.05	ntips=500
B	GeoHiSSE	0	0.1	0.1	0.1	0.03	0.03	0.05	0.05	ntips=500
		1	0.2	0.2	0.2	0.03	0.03	0.05	0.05	
		2	0.4	0.4	0.4	0.03	0.03	0.05	0.05	
C	GeoHiSSE	0	0.1	0.1	0.1	0.03	0.03	0.05	0.05	ntips=500
		1	0.2	0.2	0.2	0.03	0.03	0.05	0.05	
		2	0.3	0.3	0.3	0.03	0.03	0.05	0.05	
		3	0.4	0.4	0.4	0.03	0.03	0.05	0.05	
		4	0.5	0.5	0.5	0.03	0.03	0.05	0.05	
D	GeoHiSSE	0	0.1	0.2	0.1	0.03	0.03	0.05	0.05	ntips=500
		1	0.2	0.4	0.2	0.03	0.03	0.05	0.05	
		2	0.4	0.8	0.4	0.03	0.03	0.05	0.05	
E	GeoSSE	0	0.5	0.1	0.1	0.03	0.03	0.05	0.05	ntips=500
F	GeoSSE + jumps with $q_{AB} = q_{BA} = 0.05$	0	0.1	0.1	0.1	0.03	0.03	0.01	0.01	ntips=500
G	BiSSEness with $p_{0c}=p_{1c}=0.1$ and $p_{0a}=p_{1a}=1$	0	NA	0.1	0.1	0.03	0.03	0.01	0.01	ntips=500
H	BiSSE tree and Mk3 ranges	For generating the trees: $\lambda_{00} = 0.5$; $\lambda_{01} = 1$; $\mu_0 = \mu_1 = 0$; $q_{01} = q_{10} = 0.006$ For simulation of the areas: $q_{A_AB} = q_{AB_A} = 0.05$; $q_{B_AB} = q_{AB_B} = 0.05$; $q_{A_B} = 205$ $q_{B_A} = 0$ [Kept only sims with at least 50 species in the less frequent range.]								ntips=500

207 Table S2: Additional 17 models used in empirical study of conifers in addition to the 18 models described
 208 in Table 1 in the main text.

Model	Description	free parameters
19	CID - GeoHiSSE, 4 rate classes, null model, cladogenetic	11
20	CID - GeoHiSSE+extirpation, 4 rate classes, null model, cladogenetic	13
21	CID - GeoHiSSE, 4 rate classes, null model, anagenetic	13
22	CID - GeoSSE, $s_{0l}=0$ and $x_{0l}=0$, anagenetic	4
23	GeoSSE, $s_{0l}=0$ and $x_{0l}=0$, anagenetic	6
24	CID - GeoHiSSE, 3 rate classes, $s_{0li}=0$ and $x_{0li}=0$, anagenetic	9
25	GeoHiSSE, 2 rate classes, full model, $s_{0li}=0$ and $x_{0li}=0$, anagenetic	13
26	CID - GeoHiSSE, 5 rate classes, $s_{0li}=0$ and $x_{0li}=0$, anagenetic	13
27	CID - GeoHiSSE, 2 rate classes, $s_{0li}=0$ and $x_{0li}=0$, anagenetic	7
28	CID - GeoHiSSE, 4 rate classes, $s_{0li}=0$ and $x_{0li}=0$, anagenetic	11
29	CID - GeoSSE+extirpation, $s_{0l}=0$ and $x_{0l}=0$, anagenetic	6
30	GeoSSE+extirpation, $s_{0l}=0$ and $x_{0l}=0$, anagenetic	8
31	CID - GeoHiSSE+extirpation, 3 rate classes, $s_{0li}=0$ and $x_{0li}=0$, anagenetic	11
32	GeoHiSSE+extirpation, 2 rate classes, full model, $s_{0li}=0$ and $x_{0li}=0$, anagenetic	17
33	CID - GeoHiSSE+extirpation, 5 rate classes, $s_{0li}=0$ and $x_{0li}=0$, anagenetic	15
34	CID - GeoHiSSE+extirpation, 2 rate classes, $s_{0li}=0$ and $x_{0li}=0$, anagenetic	9
35	CID - GeoHiSSE+extirpation, 4 rate classes, $s_{0li}=0$ and $x_{0li}=0$, anagenetic	13

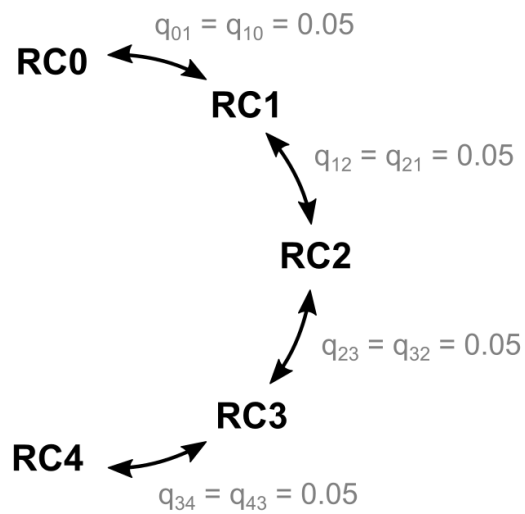


Figure S1: Scheme of the transition rates between rate classes (RC0 to RC4) used for the simulation scenarios with multiple rate classes (Sims B, C and D). Transitions between rate classes were modelled with the same rate (0.05) following a meristic Markov model. As a result, diversification rates vary following a gradient across the branches of the three.

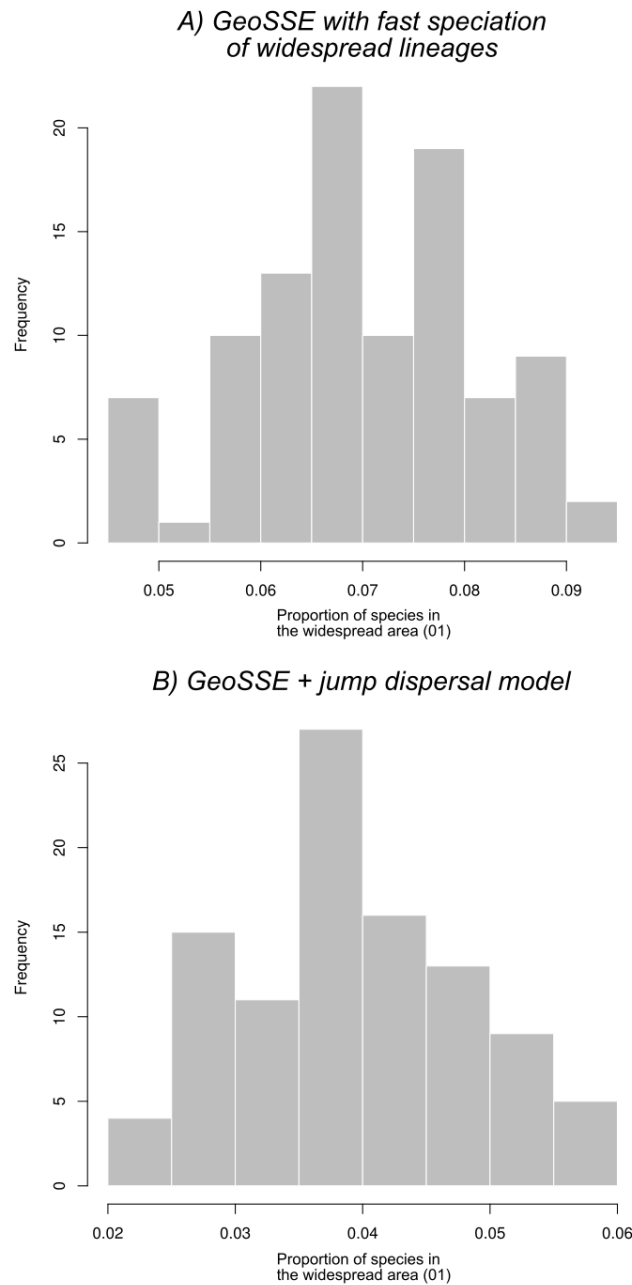


Figure S2: Proportion of widespread lineages on trees simulated under scenarios E and F. See Table S1 for description of model parameters. Top histogram (A) shows results from a standard GeoSSE model with speciation rates for widespread lineages five times faster than endemic lineages. Bottom histogram (B) are results from simulations using a modification of the GeoSSE model allowing jump dispersion events. Of course, scenario H has no extant or extinct lineages in the widespread range (see ‘Simulation study’ section of the main text and Table S1).

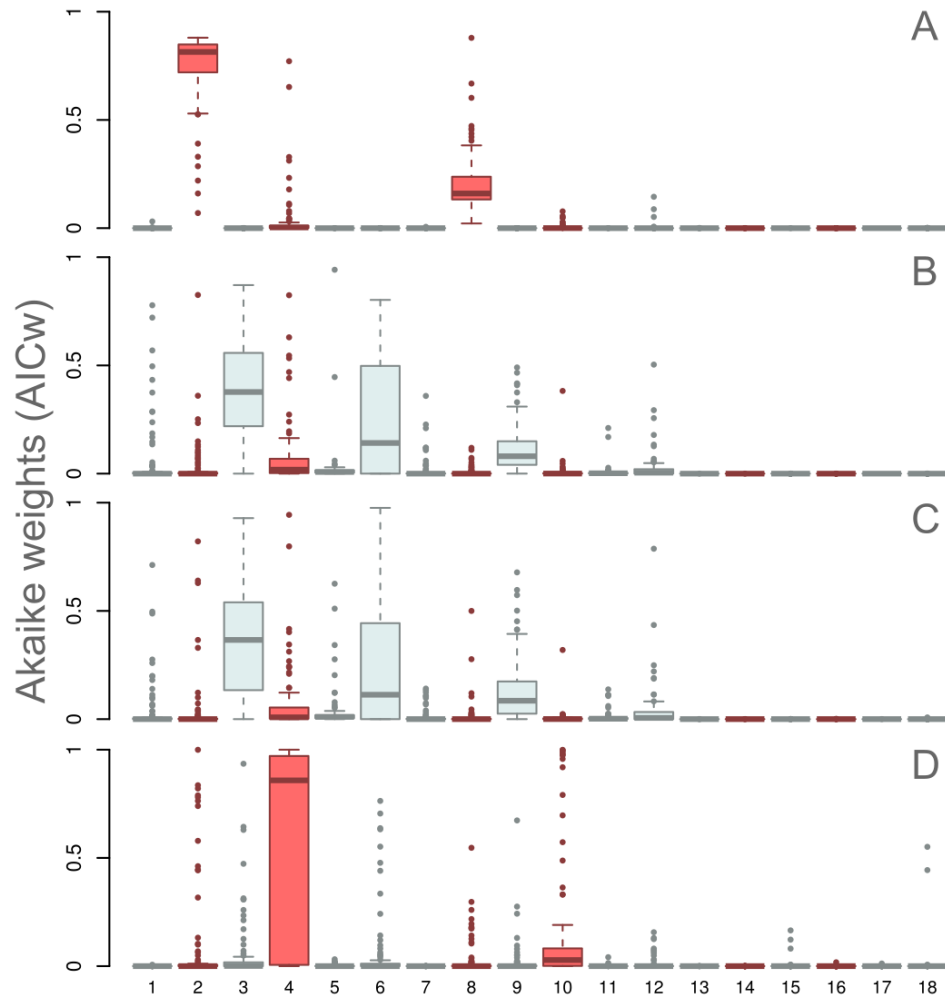


Figure S3: Summary of model support for simulated scenarios A to D. Plots show distribution of Akaike Information Criterion weights (AICw) for each model (columns) computed with 100 simulation replicates. Box-plots in red are area-dependent models and gray plots are area-independent models. A) Data simulated under the area-dependent GeoSSE model. B) and C) simulated phylogenies with three and five diversification shifts unrelated to geography, respectively. D) Simulation of area-dependent diversification GeoHiSSE model with two rate classes. Table 1 shows list and description of fitted models and Table S1 show details for each simulation scenario.

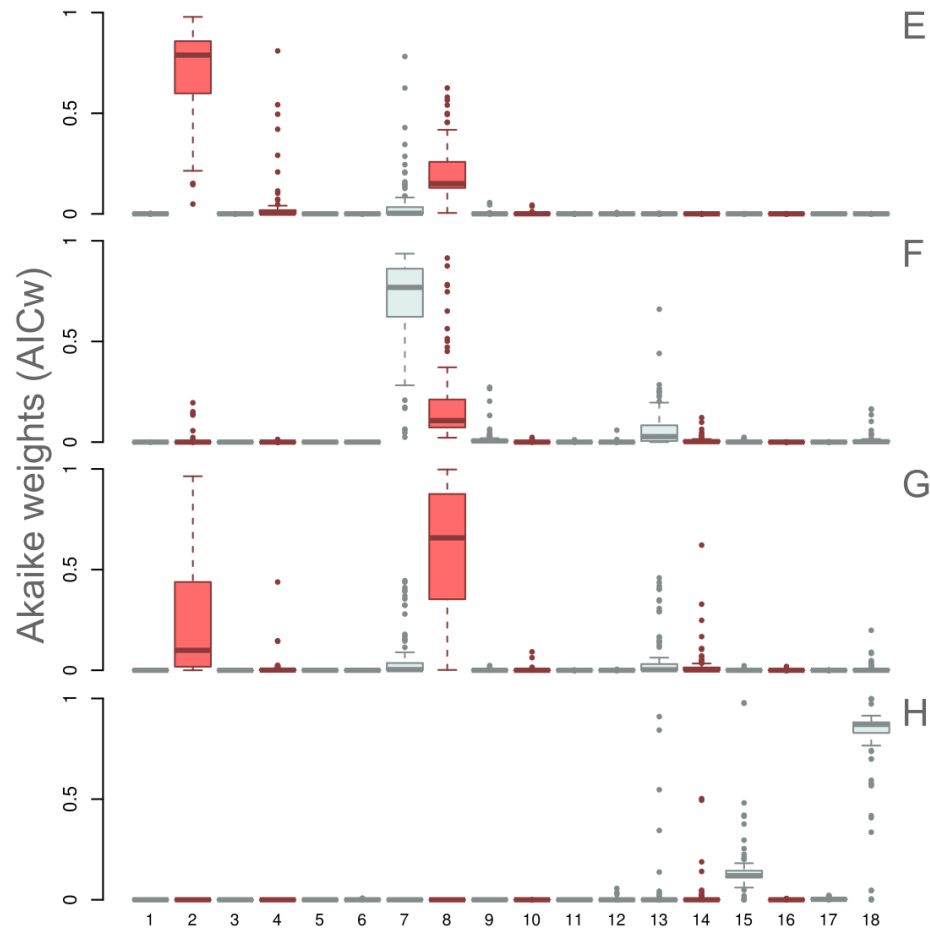


Figure S4: Summary of model support for simulated scenarios E to H. Plots show distribution of Akaike Information Criterion weights (AICw) for each model (columns) computed with 100 simulation replicates. Box-plots in red are area-dependent models and gray plots are area-independent models. E) Data simulated under a GeoSSE model with speciation rates associated to the widespread range 5x faster than endemic areas. F) Data simulated under a modified GeoSSE model allowing anagenetic transitions (i.e., jumps) between endemic ranges. G) Data generated using a BiSSEness model (Magnuson-Ford and Otto 2012) with two endemic regions. H) Data generated using an anagenetic Markov model (details on main text and Table S1). Table 1 shows list and description of fitted models and Table S1 show details for each simulation scenario.

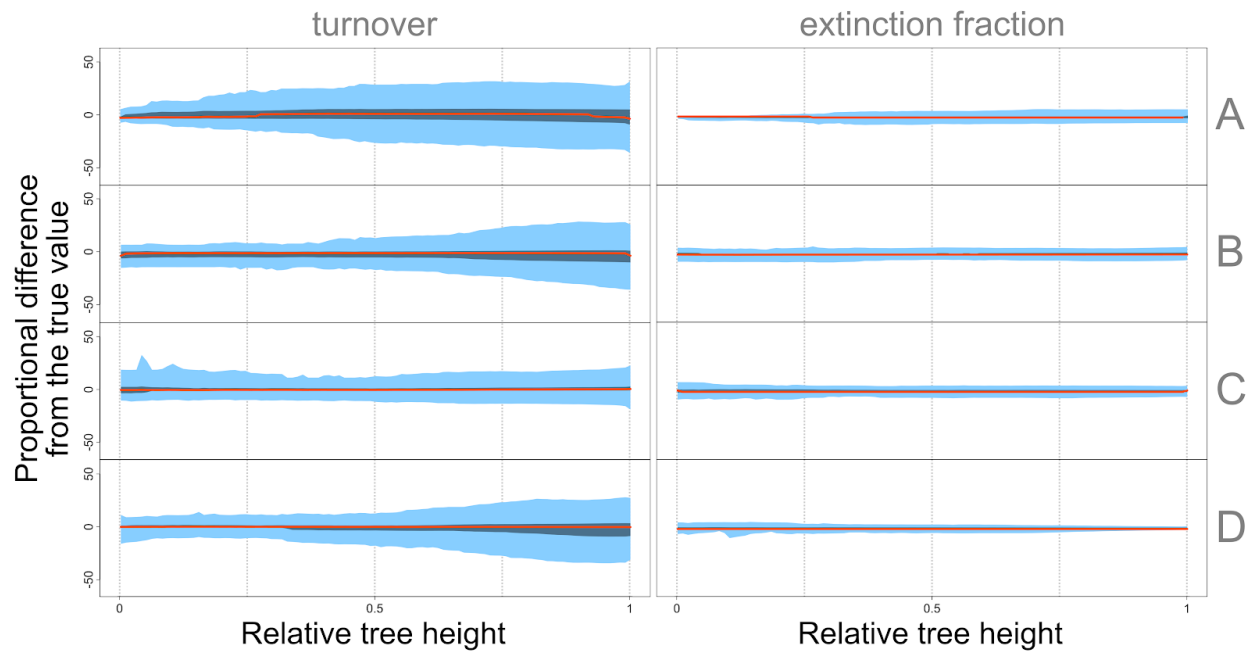


Figure S5: Accuracy of parameter estimates for simulations scenarios A to D. Plots show the proportional difference between estimated parameters and their true values computed for nodes and tips in the tree. Light blue shades represent the running 5% and 95% quantiles computed for all simulation replicates using 100 cumulative bins equally spaced from the root towards the tips of the tree. Dark blue shades (not always visible) show the limits between the running 25% and 75% quantiles. Red lines show the median of parameter estimates. See Figure S5 for estimates of net diversification.

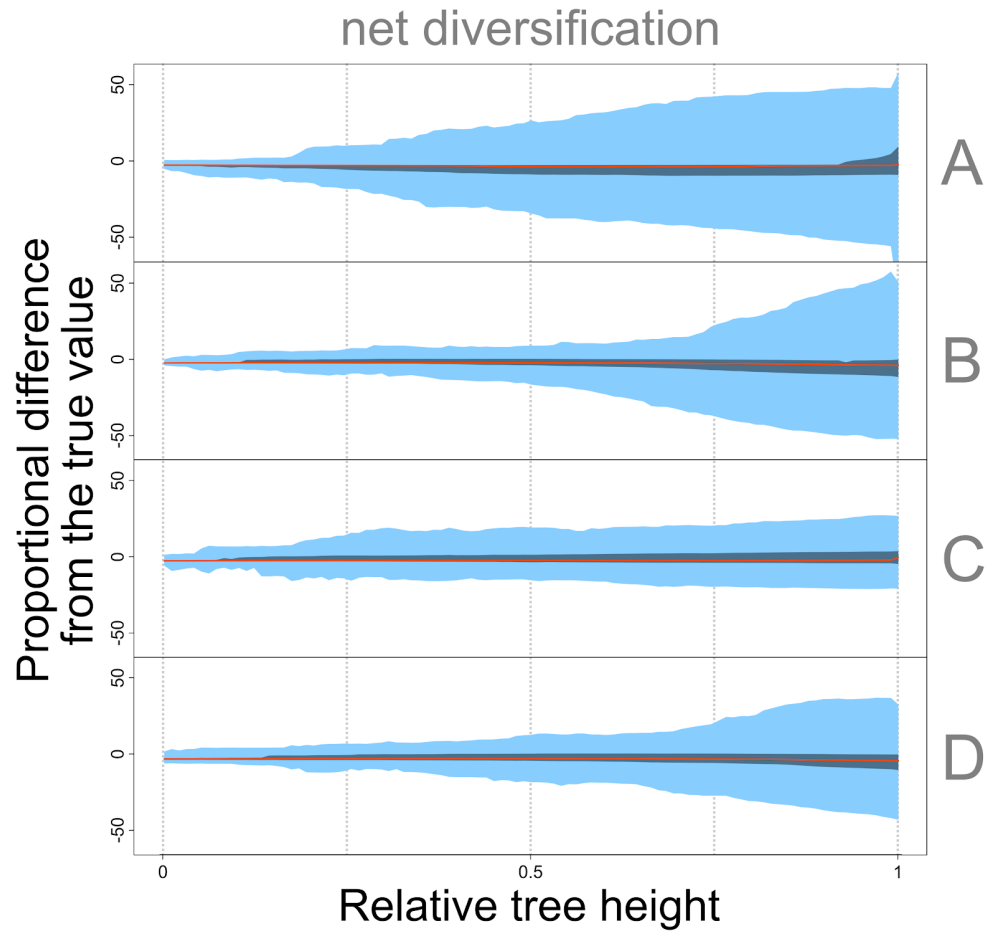


Figure S6: Accuracy of parameter estimates for simulations scenarios A to D. Plots show the proportional difference between estimated parameters and their true values computed for nodes and tips in the tree. Light blue shades represent the running 5% and 95% quantiles computed for all simulation replicates using 100 cumulative bins equally spaced from the root towards the tips of the tree. Dark blue shades (not always visible) show the limits between the running 25% and 75% quantiles. Red lines show the median of parameter estimates.

Section 1

Testing the effect of heterogeneous rates on GeoSSE models -- Earlier studies showed that the Binary State-dependent Speciation and Extinction model (BiSSE) shows an undesirable behavior when faced with rate heterogeneity in diversification across the tree that is not associated with character states (Rabosky and Goldberg, 2015). Rabosky and Goldberg (2015) show that BiSSE has an issue both with respect to the frequency in which the trait-dependent models are selected when no such process is present and with misleading parameter estimates for such models. The explanation for this behavior is general enough and may apply to every State-dependent Speciation and Extinction model (SSE). When rates of diversification are heterogeneous across the phylogenetic tree, the original SSE models have no means to accommodate the shifts in rates other than set speciation and/or speciation associated with different states to differ across the branches of the tree (Beaulieu and O'Meara, 2016). Thus, here we evaluate the behavior of the original GeoSSE model to area-independent shift in diversification rates, in order to show evidence of a similar pattern.

For this we used the same datasets generated for the simulation scenarios B and C (Table S1), but we restricted the set to include only the homogeneous GeoSSE models (see Figure 3, top panel). We chose to include representatives of our expanded GeoSSE models (i.e., GeoSSE+extirpation and anagenetic GeoSSE) because these might be prone to the same issues when no hidden rate classes are included in the set of models. The model set is comprised by one area-dependent and one area-independent configuration of the original GeoSSE, the GeoSSE+extirpation, and the anagenetic GeoSSE models (i.e., models 1, 2, 7, 8, 19, and 20 described on Table 1). We fit each model using Maximum Likelihood to obtain parameter estimates, computed their Akaike model weight (AIC_w) and performed model averaging.

Results show that the distribution of parameter estimates averaged across all models in the set and pooled for all simulation replicates is centered in the true value for the simulated datasets (Figure S6, top row). In other words, parameter estimates show no difference between rates of diversification associated with areas 0 or 1. Akaike weights across all simulations are not overwhelmingly biased towards area-dependent models (Figure S6, bottom row). For instance, in many of the simulation replicates there is substantial AIC_w for both area-independent and

area-dependent models, meaning that parameter estimates averaged across these models include contributions from both equal rates and area dependent rates. These results contrast with previous descriptions of the problem associated with BiSSE models.

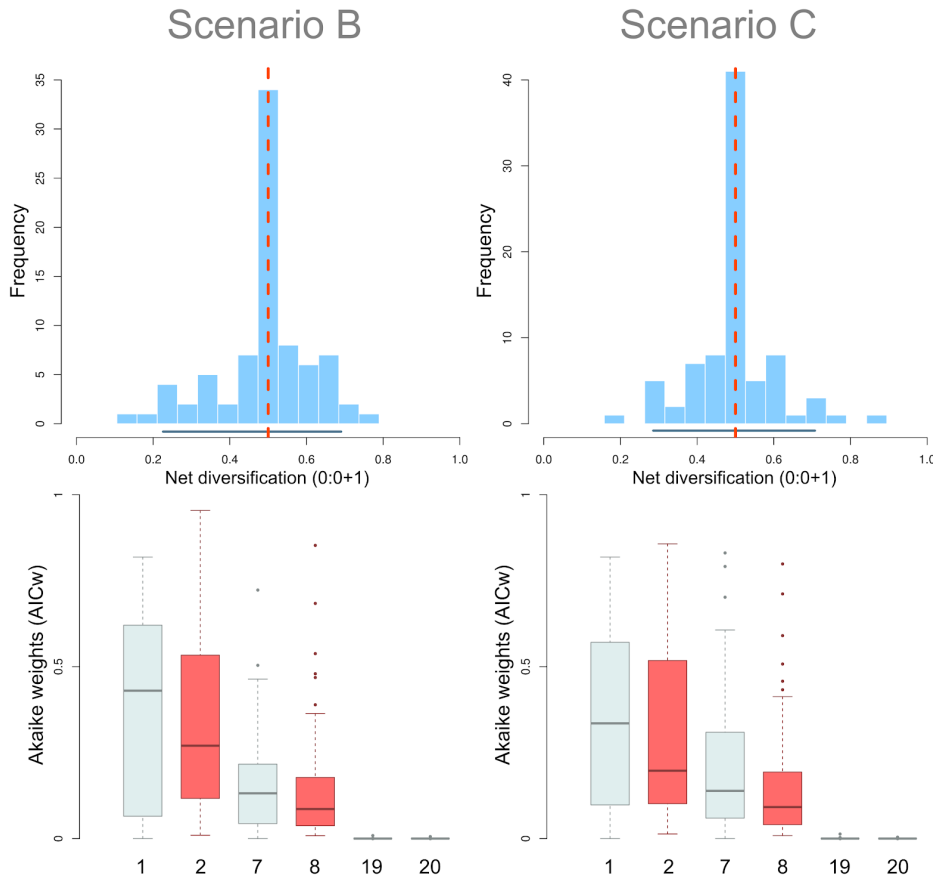


Figure S7: Results for relative net diversification rates and Akaike model weights (AICw) for simulation scenarios B and C. Both scenarios show heterogeneous diversification rates not associated with the areas (Table S1). Top plots show the distribution of ratios between net diversification rates for areas 0 and 0+1 computed for each simulation replicate. Red dashed vertical lines represent the true value for the ratios whereas horizontal blue lines show the empirical 95% CI. Bottom plots show distribution of weights across all simulation replicates for each of the models in the set, see Table 1 (in the main text) for description of the models.

When we evaluate the same results using model selection based on AIC and keeping all models within 2 AIC units of the best model, the results are very different. About one third of the replicates show strong support for area-dependent models of diversification (30% for Scen B and 28% for Scen C), approximately 40% showed decisive support for null models (37% for Scen B

and 44% for Scen C), whereas the rest included both area-dependent and independent models among the best models. In other words, one of each three simulation replicates showed a problem of model misspecification or returned inconclusive results.

Our results show that the original implementation of the GeoSSE model, without the inclusion of hidden states, shows model adequacy issues that are similar to BiSSE (Rabosky and Goldberg 2015; Beaulieu and O’Meara 2016). However, by focusing on the parameter estimates averaged across the models, instead of relying solely on a model choice framework, one is much more likely to arrive in the conclusion that there is no difference in the rates of diversification associated with each of the geographical areas or that this difference is reduced even when using only the simplest set of models. It is important to note, however, that proper usage of the GeoHiSSE (as well as HiSSE) models require that there is an trait-independent model (CID) with a comparable number of free diversification parameters for each of the area-dependent models in the set. Notwithstanding, our results here show that model-averaging over the parameters estimates for area-independent and area-dependent GeoSSE models can provide accurate parameter estimates.

Section 2

Extended simulation results -- Here we performed two sets of simulations to test the behavior of the GeoSSE model including hidden states. The first set of simulations is composed of four different scenarios that test area-independent and area-dependent diversification with varying degrees of heterogeneity in rates of diversification (Scenarios A to D). The second set has another four simulation scenarios that explore the behavior of the model under extreme cases. The first three scenarios (E to G) simulate cases of reduced frequency (or complete absence) of widespread lineages whereas the last scenario (H) tests the case in which ranges have evolved only due to anagenetic changes (no cladogenetic events). Table 1 shows the parameter values used to simulate trees and geographic range distributions observed at the tips for 100 replicates for each of the scenarios. In all simulations we used 500 lineages in the tree as the stopping criteria. For the simulations with multiple rate classes (scenarios B, C and D), we used a meristic Markov model to control the transitions between rate classes such that each transition would represent a gradual change from the fastest rate class to the slowest by passing through the intermediary ones (see Figure S1). We fitted the 18 models shown in the Table 1 (available on the main text of this study) to each of the replicates for each of the simulation scenarios using maximum likelihood. Then we performed model-averaging using Akaike weights. Figures S3 and S4 show the distribution of Akaike weights of models pooled for all replicates for each simulation scenario for simulation sets A to D and E to G, respectively.

Area-dependent and area-independent simulations -- Results with scenarios A to D show that our GeoHiSSE models, in overall, are adequate to study rates of diversification dependent or not on geographical ranges. Figure S3 shows a summary of the results. In many cases multiple models with congruent diversification histories showed high Akaike weights. For example, in simulation scenario A, model 2 is an original implementation of the area-dependent GeoSSE model without hidden states whereas model 8, which also showed high Akaike weight in part of the simulations, is an area-dependent GeoSSE+extirpation model.

In the case of the area-independent simulation scenarios B and C, there are multiple models with high Akaike weight. However, every model showed in gray in Figure S3 (and

Figure S4) are area-independent models. Simulation scenarios B and C show replicates with high Akaike weights for the models 3, 6 and 9 (Table 1, main text). Models 3 and 6 are instances of the area-independent model with different number of rate categories; model 3 has three hidden states and assumes that all transitions (including dispersions and transitions between rate classes) are constrained to be symmetrical and model 6 has two hidden states, but transition rates are estimated from the data. Model 9 is another area-independent model with symmetrical transition rates, but allows for rates of local extinction and range reduction to be estimated separately.

Results for simulation scenarios A to C are examples of the utility of applying model-averaging to estimate parameters taking into account the uncertainty in model choice. Here different models show high Akaike weights on the simulations, but all these models are congruent with the generated data. The fact that there is uncertainty associated with which model show high Akaike weights has to do with the signal in the simulated data. Tree shape, frequency of observed ranges across lineages, distribution of branch lengths and etc, all vary among the simulation replicates within each scenario. Thus, it is natural to expect some level of model uncertainty, especially when fitting more realistic models that take into account heterogeneous rates of diversification associated or not with the observed species ranges, such as in our study. It is important to note, however, that model uncertainty is different from process uncertainty. Note that the diversity of models observed for each simulation scenario do not vary in function of whether the process is area-dependent or area-independent diversification. Although there is variance on the support for each model across replicates, the conclusion of whether the process is dependent or independent of range does not change. Applying model-averaging across the set of models allows to estimate parameter estimates incorporating this uncertainty and one can evaluate the certainty about the process by evaluating the distribution of Akaike weight across the models. [Not very distant from the interpretation of model averaging when performing Bayesian model averaging - See *Similarities between BAMM and model-averaging using AICw.*]

The last simulation scenario (D) is an area-dependent scenario with three rate classes. Results are congruent with simulation scenario A, but with more variance in Akaike weights among models. It seems that model uncertainty is somewhat associated with rate heterogeneity across the tree, which is not surprising, given that this is the main reason for the inadequacy of

the original implementation of the SSE models. Again, the uncertainty among models does not bias our conclusions about the biological process behind the observed data. Most of the model weight across the simulation replicates is associated with area-dependent models.

Model performance under extreme scenarios -- The simulated scenarios from A to D followed a joint tree and geography model of evolution, where diversification rates were tied to geographical areas and range evolution occurred through cladogenesis (i.e., through speciation of widespread lineages), or along the branches of the tree (i.e., due to dispersion and extirpation). However, our knowledge of the processes that led us to observe lineages in particular geographic areas is often incomplete and empirical data can behave in ways not expected by our models. In other words, simulations based solely on SSE models, including hidden states or not, as generating models are naive surrogates with respect to empirical data sets. Here, we study two extreme cases of geographic range evolution with the objective of identifying odd behaviors when simulating data sets 1) where widespread ranges are rare or absent in extant species, and 2) where the evolution of areas are not tied to cladogenetic events.

Transitions between endemic areas in GeoSSE and GeoHiSSE are modelled as a two-step process. First, an endemic lineage disperses and becomes widespread and then it can either undergo cladogenesis, which generates two endemic sister lineages (one in area 0 and another in 1), or a local extinction in one of the areas reduces the range to endemic again. If extant widespread lineages are rare or absent, the information to infer cladogenetic and dispersion events between endemic ranges become limited, possibly leading to issues with parameter estimates and model adequacy. We first simulated datasets with widespread lineages as being rarely observed at the tips by generating data and trees under a GeoSSE model with speciation rates of widespread lineages five times faster than endemics (see scenario E in Table S1). This produced data sets with an average of ~7%, out of 500, extant lineages occupying widespread areas (Figure S2A). However, parameter estimates across all simulation replicates showed that the low frequency of widespread extant lineages does not prevent our set of models from reaching meaningful estimates using model-averaging (Figure 4E).

Alternatively, range expansion could have been rare throughout the history of the group whereas jump dispersal events (i.e., direct transitions between endemic distributions) have played an important role. To simulate such a scenario we used a GeoSSE model, but we allowed lineages to disperse between endemic areas without becoming widespread first (scenario F). This scenario resulted in an average of only 4%, out of 500, extant species occupying widespread areas (Figure S2B). However, there is no evidence for a significant bias in parameter estimates for both area-dependent diversification rates or between-area speciation rates (Figure 4F). On the whole, our approach of model-averaging across a large set of candidate models does not appear sensitive to rare extant widespread areas.

Finally, we explored the extreme possibility that the widespread range is completely absent both in extant distributions and in the evolutionary history of the group. In this scenario, changes in geographical distribution are the result of a) jump dispersal events between endemic areas or b) speciation events in one of the endemic ranges leading to one sister lineage occurring in the other area (see Magnuson-Ford and Otto, 2012). We relied on BiSSE-ness, the cladogenetic model for binary states (Magnuson-Ford and Otto, 2012), in order to simulate data sets that result in only two endemic areas observed at the tips (scenario G). When fit to our model set, the absence of widespread areas among the extant species produces estimates of the rates of between-area speciation (s_{AB}) that are highly uncertain (Figure 4G). The 95% density interval for the model-averaged estimate of s_{AB} across nodes spans the extreme wide interval between 4 and 58 units. These estimates are orders of magnitude higher than the rates of speciation associated with each of the endemic regions. In contrast, estimates for the relative difference in within-area net diversification rates associated with each endemic area did not show a strong bias (Figure 4G), suggesting that poor estimates for between-area speciation does not strongly bias our conclusions about range-dependent diversification rates.

All previous scenarios assumed that cladogenetic events were important in the evolutionary history of the lineages. This is a very plausible element of the model; since the data is expected to describe geographical ranges. However, here we also considered the performance of the model when this is not the case, perhaps because the coarse subdivision of ranges required by GeoSSE is grossly inadequate for the study system. For this, we generated datasets with

transitions between areas restricted only to anagenetic dispersal events along the branches of the tree. We simulated trait-independent phylogenetic trees with two rates of diversification following Rabosky and Goldberg (2015). We then generated datasets using only a simple transition-based Markov model by restricting transitions between endemic areas to always pass through the widespread area (see the anagenetic GeoSSE model in Figure 3 of the main text, middle panel). The difference in within-area rates of diversification is larger than observed in any other simulation scenario (Figure 4H). Moreover, the absence of cladogenetic events makes estimates for between-area speciation (s_{AB}) uncertain, although raw values for the parameter are within the same order of magnitude of the true rates of diversification across the tree (grey lines in Figure 4H).

Section 3

Performance of GeoSSE+extirpation models -- The original GeoSSE model does not make a distinction between events of range reduction and extinction of endemic lineages. Range reduction happens when a widespread lineage ($0I$) becomes locally extinct in one of the areas, leading to an endemic distribution area 0 (when extirpated from I) or area I (when extirpated from 0). In contrast, the extinction of endemics result in the complete extinction of the lineage. The original parameterization of the GeoSSE model maps both events to a single extirpation rate parameter associated with the endemic area 0 (x_0) and I (x_I). Goldberg and colleagues (2011) do consider the expansion of GeoSSE into different, often more parameter rich, variants, but no work so far had explored the effect of separating rates of range reduction from the extinction of endemics. We performed a series of simulations to test whether we can properly estimate separate rates for range reduction ($d_{AB \rightarrow A}$ and $d_{AB \rightarrow B}$) and extinction of endemic lineages (x_A and x_B) using our expanded GeoSSE+extirpation model. We compared the parameter estimates between the original GeoSSE and the GeoSSE+extirpation model in the absence of hidden states.

In order to estimate separate rates for range reduction and extinction of endemic lineages the model needs to be expanded to include one more rates for each endemic area. The GeoSSE+extirpation model has a rate of range reduction for each area (parameters x_0 and x_I) and a separate rate of extinction of endemics (parameters x^*_0 and x^*_I). The GeoSSE+extirpation model can be extended to include hidden states, which we refer as the GeoHiSSE+extirpation set of models. Like all other GeoHiSSE models, the rates of range reduction as well as extinction of endemics are associated with the hidden states. Thus, for a GeoHiSSE+extirpation model with 2 hidden states we would have the parameters: x_{0A} , x_{0B} , x_{IA} , x_{IB} for range reduction and x^*_{0A} , x^*_{0B} , x^*_{IA} , x^*_{IB} for extinction of endemics. [Please see notes about model complexity and the need for more species in the phylogeny in the main text.]

Here we test if it is possible to differentiate between rates of range reduction from the extinction of lineages in endemic areas using the GeoSSE+extirpation model. We simulated phylogenetic trees and data under four distinct scenarios (Scen ext_A to ext_D, see Table S3). For the first scenario we set range contraction and dispersion to be more frequent than extinction of endemic lineages (Scen ext_A). This scenario models the case in which events of dispersal

and range contraction occur at the same rate, but the extinction of endemic lineage are relatively rare. In other words, recent dispersers face a higher chance of being extirpated from the area than established lineages. For the second scenario we kept the same generating values used for ext_A, but we increased the rate of extinction of endemic lineages in area 0 (Scen ext_B). Diversification rates associated with area 1 are higher than in 0, due to an increase in extinction in area 0. The first and second scenario test the performance of our estimates for separate rates of range reduction and extinction of endemics as well as if such processes can carry a signal of area-dependent diversification.

Table S3: Description of scenarios and parameter values used to simulate phylogenetic trees and range distributions under the GeoSSE+extirpation model. All simulations used models with a single rate category. Parameters in bold were modified from the original GeoSSE model. Here x_0 and x_1 are the rates of range reduction and x^*_0 and x^*_1 are the rates of extinction of endemic lineages associated with each of the areas.

Scen	Parameters									Simulation stop criteria
	s_{01}	s_0	s_1	x_0	x_1	d_0	d_1	x^*_0	x^*_1	
ext_A	0.1	0.1	0.1	0.05	0.05	0.05	0.05	0.01	0.01	ntips=500
ext_B	0.1	0.1	0.1	0.05	0.05	0.05	0.05	0.03	0.01	ntips=500
ext_C	0.1	0.1	0.1	0.05	0.05	0.10	0.05	0.01	0.01	ntips=500
ext_D	0.1	0.1	0.1	0.01	0.01	0.05	0.05	0.05	0.05	ntips=500

In the third and fourth simulation scenarios we changed the processes described for simulations ext_A and ext_B. Scenario ext_C repeats the same generating values as simulation ext_A, but we increased the rate of dispersion from area 0 to the widespread region 01. In this case, extinction of endemics is still rarer than dispersion, but dispersion events from area 0 are now twice as frequent as from area 1. With this we can explore if there are important confounding factors among the anagenetic events (i.e., dispersion, range reduction, and extinction of endemics). Finally, scenario ext_D flips the relationship between range reduction and extinction of endemics assumed in the previous simulations. Now extinction rates of endemic lineages are higher than the rate with which widespread lineages become endemic. Scenarios ext_A to ext_C assume that recent dispersers are likely to lose part of their range

591 whereas scenario ext_D explores the case that lineages are much more prone to extinction when
592 they become restrict to endemic distributions than to have their range contracted.

593 We fitted a reduced model set with four models using Maximum Likelihood Estimate and
594 estimated parameters by performing model averages using the Akaike weights for each of the
595 models. Since the aim of these tests are to show the performance of the GeoSSE+extirpation
596 models with respect to the original GeoSSE models, we decided not to include any instance of
597 the GeoHiSSE or anagenetic versions of the GeoSSE model. Here we use a collection of
598 area-independent and area-dependent models with and without separating range reduction from
599 extinction of endemics (see Table 1 in main text): Model 1 (CID original GeoSSE), Model 2
600 (CID original GeoSSE), Model 7 (CID GeoSSE+extirpation), and Model 8 (area-dependent
601 GeoSSE+extirpation).

602 In terms of model weight, results were similar across simulation scenarios ext_A to
603 ext_D. Both area-independent models 1 and 3 showed higher Akaike weight across all
604 simulation replicates (Figure S7). The different rates of extinction (Scen ext_B) or dispersion
605 (Scen ext_C) associated with areas simulated in the data show no reflection in model weight
606 when compared to other simulation scenarios (Scen ext_A or Scen ext_D). However, when we
607 look to the parameter estimates for each of the models across the 100 replicates there is strong
608 evidence that GeoSSE+extirpation models are able to correctly recover the generating parameters
609 for the different scenarios (Figure S7). Parameter estimates across all simulations (Scen ext_A to
610 Scen ext_D) show that we can adequately distinguish between rates of range reduction and rates
611 of extinction of endemic lineages. However, these results are conditioned on a phylogeny with
612 500 species and we strongly recommend performing similar tests if planning to use smaller trees.

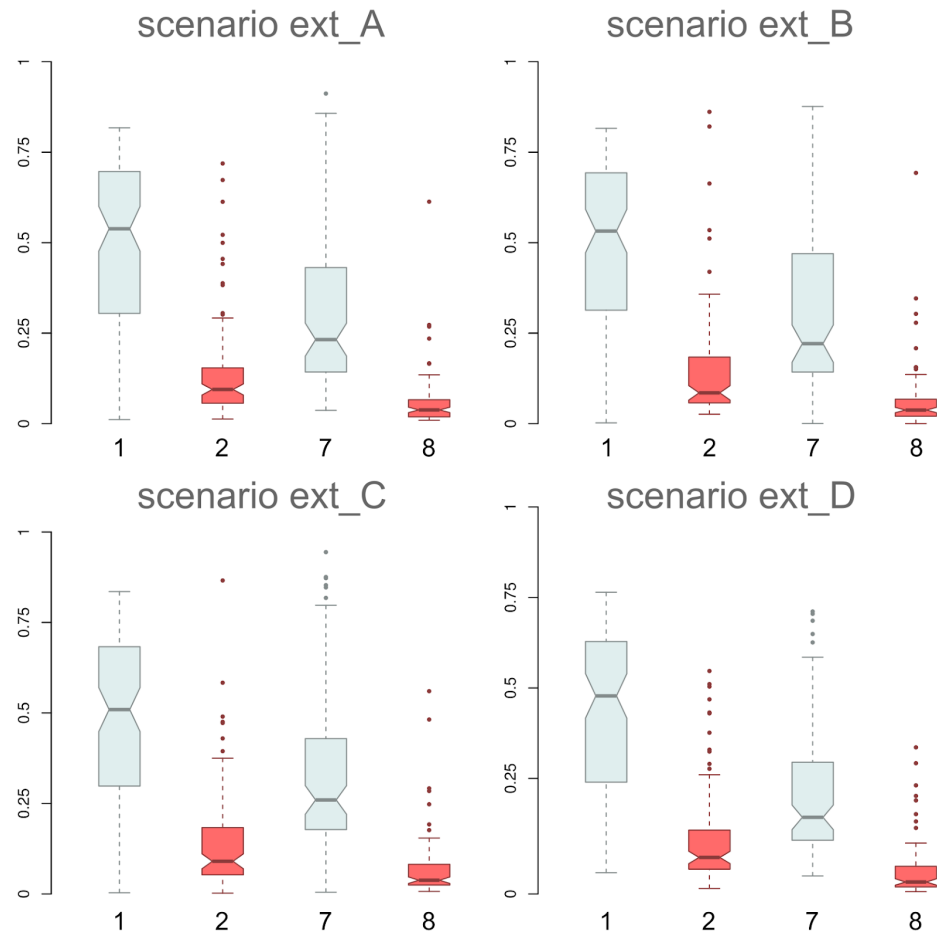


Figure S8: Distribution of Akaike weights for the model set fitted to simulation scenarios ext_A to ext_D. Box plots in grey are area-independent models and in red are area-dependent models. See Table 1 (main text) for description of the models and Table S3 for parameter values used for the simulations.

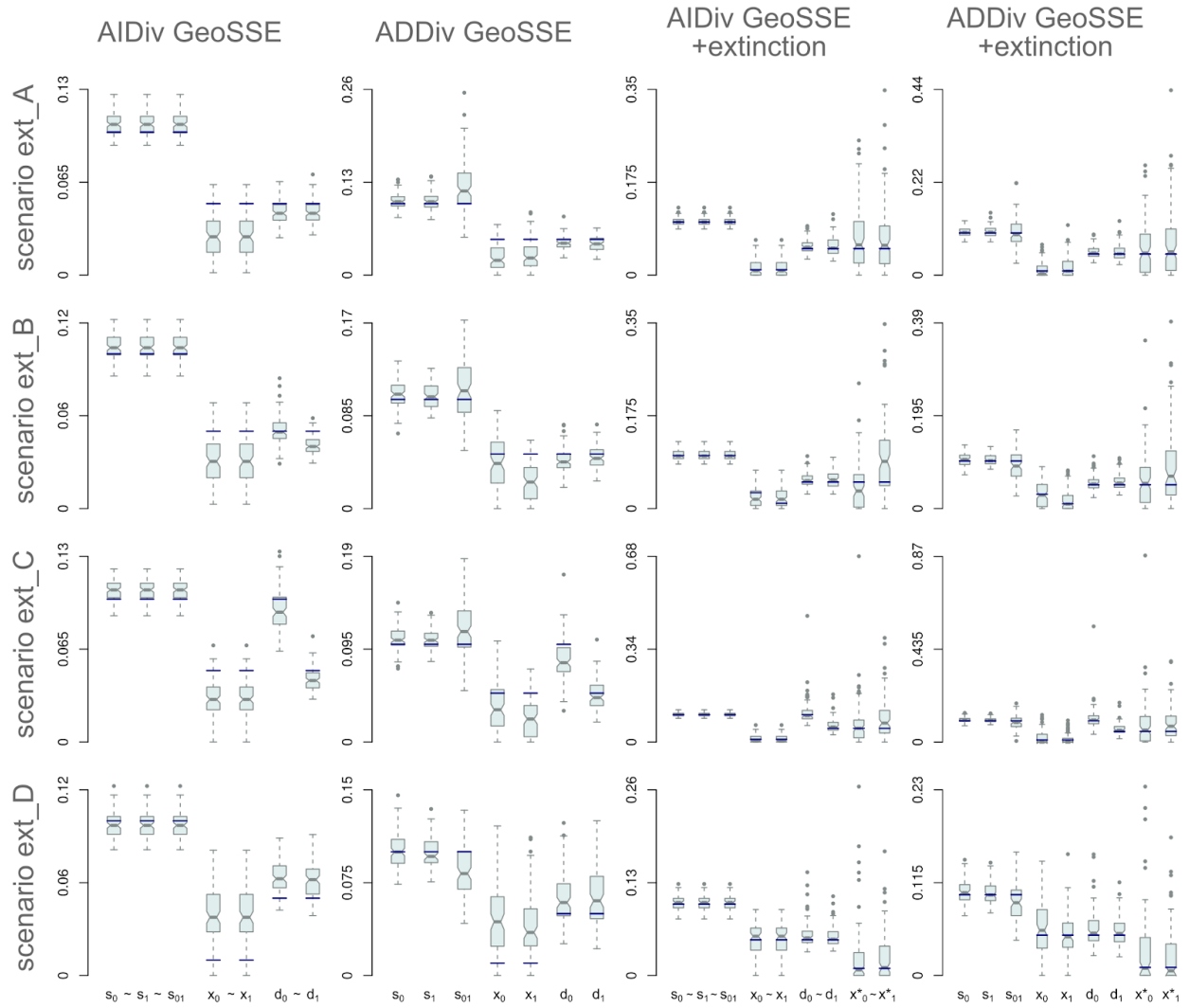


Figure S9: Distribution of parameter values across 100 simulation replicates for each of the scenarios ext_A to ext_D. Here ‘AIDiv’ denotes area-independent models and ‘ADDiv’ denotes area-dependent models. Rows represent simulation scenarios whereas columns are different models. Columns 1 and 2 show original GeoSSE models (7 parameters) and columns 3 and 4 are GeoSSE+extirpation models (9 parameters). Parameters linked by ‘~’ were constrained to the same value during Maximum Likelihood estimation. The blue horizontal lines show the values of the parameters used to generated the data for each scenario (note that scale in y axes vary).