

## 1 [Supporting information](#)

### 2 `mobsim`: An R package for the simulation and measurement of biodiversity 3 across spatial scales

4

### 5 [Similar and complementary software](#)

6 There are several software packages that partly overlap, but are also extended by `mobsim`.

7         The R package `vegan` (Oksanen *et al.* 2017) provides rich functionality to analyse non-spatial  
8 community data. In contrast to `mobsim` there are no tools for simulations and for spatially-explicit  
9 analysis, such as species-accumulation and diversity-area curves.

10         The R package `sads` (Prado *et al.* 2016) offers functions to analyse and also to simulate (non-  
11 spatial) SADs. `mobsim` extends `sads` by the option to control for both the numbers of individuals and  
12 species at the same time.

13         The R package `spatstat` (Baddeley *et al.* 2015) includes functions for the simulation of  
14 random and aggregated point processes. The `mobsim` function `sim_thomas_coords` is based on  
15 the function `spatstat::rThomas`. However, `mobsim` uses a more efficient re-implementation in  
16 C++ and integrates the simulation of point processes smoothly with species-abundance distributions,  
17 whereas `rThomas` becomes rather slow for a large number of points.

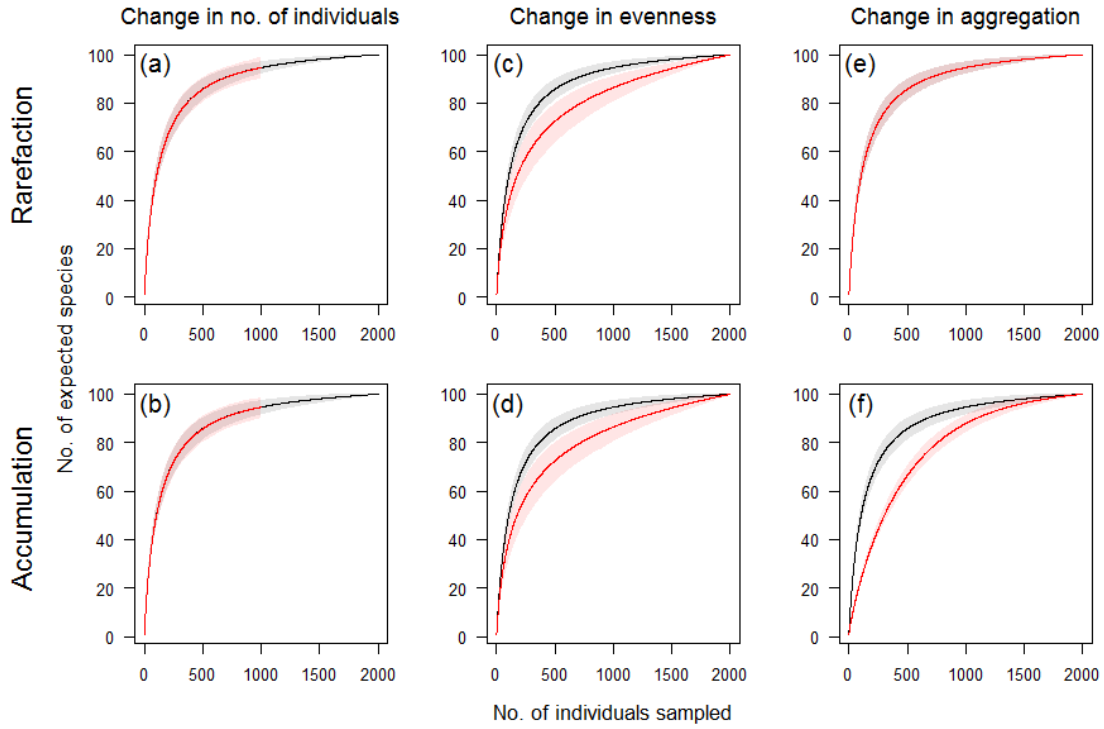
18         The R package `iNEXT` (Hsieh *et al.* 2016) offers rarefaction curves for Hill numbers, including  
19 species richness, ENS Shannon, and ENS Simpson. However, `iNEXT` only works with non-spatial  
20 abundance data and is not suited for the analysis of spatial data.

21           The R package `untb` (Hankin 2007) simulates SADs from a neutral model of community  
22 dynamics and the parameters of this specific model can be estimated from data. In contrast to  
23 `mobsim`, the simulations can represent temporal dynamics, but no spatial patterns and distributions.

24           The R package `VirtualCom` (Münkemüller & Gallien 2015) simulates phylogenetic and  
25 functional biodiversity patterns in communities, based on the processes of species competition and  
26 habitat filtering. In contrast to `mobsim`, the simulations do not consider space and focus on specific  
27 processes, rather than on general community properties, such as evenness and intraspecific  
28 aggregation.

29           The Python package `macroeco` and its graphical user interface `MacroecoDesktop` (Kitzes &  
30 Wilber 2016) provide tools for computing SADs, SARs, spatial species abundance distributions (Harte  
31 2011), spatially explicit O-ring and commonality statistics from empirical data. This software also  
32 provides a suite of tools for fitting models to these patterns and plotting the results, as well as, fitting  
33 the predicted SAR and EAR from the Maximum Entropy Theory of Ecology (Harte *et al.* 2008, 2009;  
34 Harte 2011) This project does not currently provide simulation based tools and instead is focused on  
35 the measurement and modelling of biodiversity patterns.

36



37

38 **Figure S1:** Same as Fig. 1 in the main text, but using fixed species richness in the simulated community.

39 This is implemented by using the argument `fix_s_sim = T` in the functions `sim_sad`,

40 `sim_poisson_community`, or `sim_thomas_community`.

41

42

43

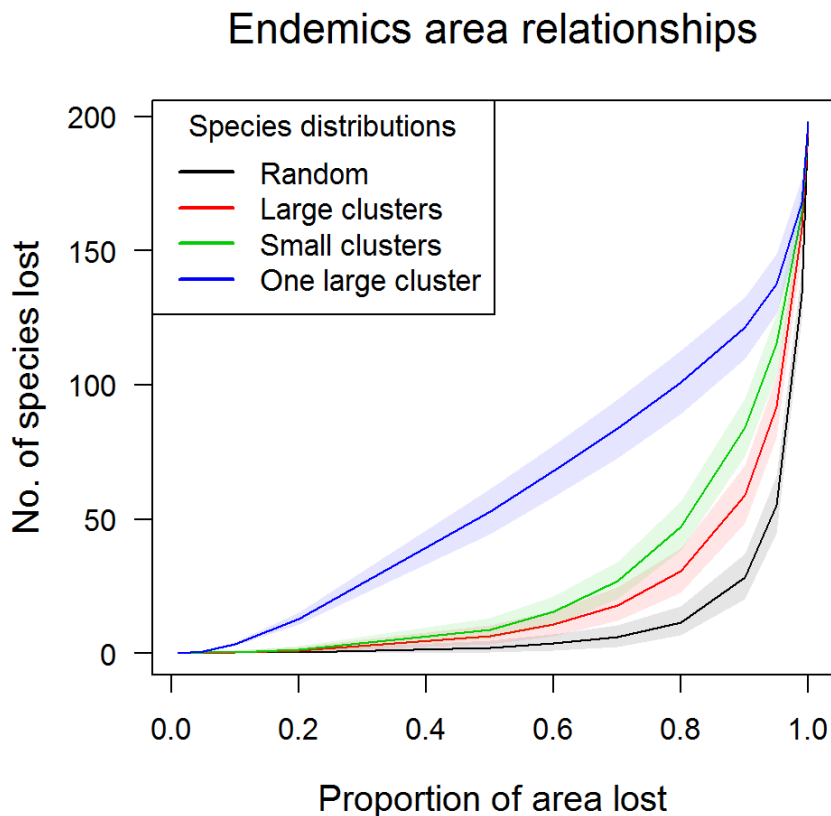
44 [Application example: Extinctions from habitat loss](#)

45 Land use changes and the resulting loss of natural habitat is one of the main drivers of biodiversity loss  
46 (Sala *et al.* 2000; Pereira *et al.* 2010), however, there has been an intense debate on how to predict  
47 extinctions due to habitat loss (e.g. He & Hubbell 2011; Keil *et al.* 2015). A key property in this debate  
48 is the number of species that are endemic to a certain area, because these species will go immediately  
49 extinct, when all individuals in this area die as consequence of land-use change (Harte & Kinzig 1997).  
50 Accordingly, the endemics-area relationship (EAR) has been suggested as important tool to predict  
51 extinctions due to habitat loss (Kinzig & Harte 2000). However, consideration of spatial aggregation  
52 among species is important. For example, He & Hubbell (2011) applied an EAR model, which assumes  
53 random spatial distributions of species and individuals and tested this model using data from tropical  
54 forests (see He & Hubbell 2011 Supplementary Information D). In contrast, Green & Ostling (2003)  
55 showed that the EAR is influenced by intraspecific aggregation using spatial simulations and they  
56 concluded that knowledge on spatial aggregation is important to predict extinctions due to habitat  
57 loss.

58 We can use `mobsim` to easily reproduce the analysis of Green & Ostling (2003). Here, we  
59 simulated four different spatial distributions that varied in intraspecific aggregation, but used the same  
60 SAD for all scenarios. The reference scenario was a random (Poisson) distribution. Then, we simulated  
61 large and small clusters of conspecific individuals using the Thomas process, but with several clusters  
62 per species. In the last scenario there was just one cluster of each species, i.e. this cluster represented  
63 the complete range of a species.

64 In agreement with Green & Ostling (2003), we found that the number of endemics, and thus  
65 the expected extinctions for a given habitat loss increases with intraspecific aggregation (Fig. 2). With  
66 a random distribution, there were none or few endemics at all but the largest areas. The numbers of  
67 endemics increased with decreasing cluster size (as specified by parameter `sigma` in the function  
68 `sim_thomas_coords`) and was by far highest in the scenario with just one cluster for every species.  
69 We conclude that the number of endemics is sensitive to intraspecific aggregation and that the finding

70 of He & Hubbell (2011) for tropical forest and birds cannot be generalized to cases where aggregation  
 71 is strong.



72

73 **Figure S2:** Endemics-area relationships for communities with the same species-abundance distribution  
 74 and the same numbers of individuals, but different spatial distributions of species. The lines and  
 75 intervals show mean and 95% confidence intervals based on 1,000 replicate simulations. The following  
 76 parameter values were used for the simulations: species pool richness: `s_pool = 200`; number of  
 77 individuals: `n_sim = 10,000`; log-normal SAD with `meanlog = 3` and `sdlog = 1`; large cluster  
 78 size: `sigma = 0.05`, small cluster size: `sigma = 0.01`; one cluster per species:  
 79 `mother_points = 1`.

80

81

## 82 References

- 83 Baddeley, A., Rubak, E. & Turner, R. (2015). *Spatial Point Patterns: Methodology and Applications with*  
84 *R*. Chapman and Hall/CRC, s.l.
- 85 Green, J.L. & Ostling, A. (2003). Endemics-area relationships: the influence of species dominance and  
86 spatial aggregation. *Ecology*, **84**, 3090–3097.
- 87 Hankin, R.K.S. (2007). Introducing `untb`, an R package for simulating ecological drift under the unified  
88 neutral theory of biodiversity. *Journal of Statistical Software*, **22**.
- 89 Harte, J. (2011). *Maximum Entropy And Ecology: A Theory of Abundance, Distribution, and Energetics*.  
90 Oxford University Press, U.S.A., Oxford.
- 91 Harte, J. & Kinzig, A.P. (1997). On the implications of species-area relationships for endemism, spatial  
92 turnover, and food web patterns. *Oikos*, **80**, 417–427.
- 93 Harte, J., Smith, A.B. & Storch, D. (2009). Biodiversity scales from plots to biomes with a universal  
94 species–area curve. *Ecology Letters*, **12**, 789–797.
- 95 Harte, J., Zillio, T., Conlisk, E. & Smith, A. (2008). Maximum entropy and the state-variable approach to  
96 macroecology. *Ecology*, **89**, 2700–2711.
- 97 He, F. & Hubbell, S.P. (2011). Species-area relationships always overestimate extinction rates from  
98 habitat loss. *Nature*, **473**, 368–371.
- 99 Hsieh, T.C., Ma, K.H. & Chao, A. (2016). `iNEXT`: an R package for rarefaction and extrapolation of species  
100 diversity (Hill numbers). *Methods in Ecology and Evolution*, **7**, 1451–1456.
- 101 Keil, P., Storch, D. & Jetz, W. (2015). On the decline of biodiversity due to area loss. *Nature*  
102 *communications*, **6**.
- 103 Kinzig, A.P. & Harte, J. (2000). Implications of endemics-area relationships for estimates of species  
104 extinctions. *Ecology*, **81**, 3305–3311.
- 105 Kitzes, J. & Wilber, M. (2016). `macroeco`: reproducible ecological pattern analysis in Python. *Ecography*,  
106 **39**, 361–367.
- 107 Münkemüller, T. & Gallien, L. (2015). `VirtualCom`: a simulation model for eco-evolutionary community  
108 assembly and invasion. *Methods in Ecology and Evolution*, **6**, 735–743.
- 109 Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., Minchin, P.R., O’Hara,  
110 R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E. & Wagner, H. (2017). `vegan`:  
111 *Community Ecology Package*.
- 112 Pereira, H.M., Leadley, P.W., Proença, V., Alkemade, R., Scharlemann, J.P., Fernandez-Manjarrés, J.F.,  
113 Araújo, M.B., Balvanera, P., Biggs, R., Cheung, W.W. & others. (2010). Scenarios for global  
114 biodiversity in the 21st century. *Science*, **330**, 1496–1501.
- 115 Prado, P.I., Miranda, M.D. & Chalom, A. (2016). `sads`: *Maximum Likelihood Models for Species*  
116 *Abundance Distributions*.

117 Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke,  
118 L.F., Jackson, R.B., Kinzig, A., Leemans, R., Lodge, D.M., Mooney, H.A., Oesterheld, M., Poff,  
119 N.L., Sykes, M.T., Walker, B.H., Walker, M. & Wall, D.H. (2000). Global biodiversity scenarios  
120 for the year 2100. *Science*, **287**, 1770–1774.

121