Effect sizes and standardization in neighbourhood models of forest stands: potential biases and misinterpretations

Peter Stoll[1], David J. Murrell[*] & David M. Newbery

Institute of Plant Sciences, University of Bern, Altenbergrain 21, CH-3013 Bern, Switzerland

[*] Centre for Biodiversity and Environmental Research, Department of Genetics, Evolution and Environment, University College London, Darwin Building, Gower Street, London WC1E 6BT, UK

Running head: effect sizes and standardization

[1] E-mail: peter.stoll@ips.unibe.ch

1 **Summary**

2 **1.** Effects of conspecific neighbours on survival and growth of trees have been found to be

3 related to species abundance. Both positive and negative relationships may explain

4 observed abundance patterns. Surprisingly, it is rarely tested whether such relationships

5 could be biased or even spurious due to transforming neighbourhood variables or

6 influences of spatial aggregation, distance decay of neighbour effects and standardization

7 of effect sizes.

8 **2.** To investigate potential biases, communities of 20 identical species were simulated with

9 log-series abundances but without species-specific interactions. No relationship of

10 conspecific neighbour effects on survival or growth with species abundance was expected.

11 Survival and growth of individuals was simulated in random and aggregated spatial

12 patterns using no, linear, or squared distance decay of neighbour effects.

13 **3.** Regression coefficients of statistical neighbourhood models were unbiased and unrelated

14 to species abundance. However, variation in the number of conspecific neighbours was

15 positively or negatively related to species abundance depending on transformations of

16 neighbourhood variables, spatial pattern and distance decay. Consequently, effect sizes

17 and standardized regression coefficients, often used in model fitting across large numbers

18 of species, were also positively or negatively related to species abundance depending on

19 transformation of neighbourhood variables, spatial pattern and distance decay.

20 **4.** Tests using randomized tree positions and identities provide the best bench marks by

21 which to critically evaluate relationships of effect sizes or standardized regression

22 coefficients with tree species abundance. This will better guard against potential

23 misinterpretations.

24

25 **Key-words:** community dynamics; multiple regression; neighbourhood model; population

26 dynamics; tropical forest.

27   **Introduction**

28   Whether or not conspecific negative density dependence (CNDD) at small

29   neighbourhood scales shapes species abundances in tropical tree communities at larger scales

30   is far from resolved and we probably should not even expect the answer to be simple. In

31   principle, there are several possibilities. First, the strength of CNDD is unrelated to

32   abundance. Second, the strength of CNDD is negatively related to abundance (strong CNDD

33   for abundant but weak CNDD for rare species). This would prevent abundant species

34   becoming even more abundant and thereby competitively excluding other species. Moreover,

35   it would confer a rare-species advantage and possibly lead to a community compensatory

36   trend (CCT, Connell et al. 1984). Third, the strength of CNDD is positively related to

37   abundance (strong CNDD for rare but weak for abundant species). This would explain the

38   rarity and low abundance of the species with strong CNDD and the high abundances of

39   species with weak CNDD (Comita *et al.* 2010). There remain though two further possibilities

40   which are that either a mix of positive and negative processes is operating, or the observed

41   relationships are simply spurious (i.e. the result of a statistical artefact).

42   In an empirical study, Newbery & Stoll (2013) showed negative effects of conspecific

43   neighbours on absolute growth rate of medium-sized trees. The argument was that reduced

44   growth of an individual tree will – other factors being equal – translate into survivorship and

45   fecundity reductions and hence affect species abundances. Nevertheless, direct effects of

46   conspecifics on survival could be more relevant for population dynamics of different species

47   within communities. Therefore, the tests reported here simulate both individual survival and

48   growth rate and use a framework of neighbourhood analysis similar to that of Newbery &

49   Stoll (2013) to show that all possible relationships of the strength of CNDD and abundance

50   may emerge without any species-specific or effects of abundances. Moreover, we show that

51   potential biases do not depend on the nature of the dependent variable.

52   Relationships between the strength of CNDD and abundance were investigated using a

53   simple, spatially explicit and individual-based, model which simulated identical species

54   without any species-specific interactions. Thus, any relationships between the strength of

55   CNDD and abundance in communities simulated under these assumptions would not be

56   expected. Nevertheless, relationships do emerge because of interfering effects of spatial

57   patterns and distance decay (i.e. the functional form relating neighbour effects to distance

58   from focal trees, Fig. 1) and, perhaps more importantly, due to transforming (e.g. log-

59   transformation) and/or scaling (e.g. standardization or z-transformation) of the input

60   variables. For example, if rare species have lower variability in the number of conspecifics in

61   their local neighbourhoods compared to common species, scaling is expected to decrease

62   effect sizes (or standardized partial correlation coefficients) of rare relative to common

63   species, possibly leading to spurious negative relationships between the strength of CNDD

64   and abundances. Scaling is recommended (e.g. Schielzeth 2010) and applied especially in

65   hierarchical Bayesian modelling to speed up or even ensure numerical convergence (e.g.

66   Gelman & Hill 2007).

67   Motivation to investigate the relationships between the strength of CNDD and

68   abundance more carefully using simulations came from the opposite outcomes of two recent

69   publications. A consistent *negative* relationship between the strength of CNDD (i.e. effect

70   sizes derived from statistical neighbourhood models) and abundance (total basal area of

71   species) in randomization tests was shown by Newbery & Stoll (2013). By contrast, a strong

72   *positive* relationship between the strength of CNDD and abundance was found by Comita *et*

73   *al*. (2010). Whilst such different results are interesting, and might be explained by different

74   underlying biological mechanisms operating on different species at different locations, before

75   making such a conclusion possible differences arising from artefacts and biases of the

76   statistical methods should first be ruled out.

Materials and Methods

77

78   A completely neutral forest without any species-specific effects was simulated. Initial

79   size distributions of individuals (basal area, *ba*) were log-normal with mean 2 and standard

80   deviation 1, and simulations were initialized with no spatial dependency in individual size.

81   Individuals of 20 identical species with log-series abundances (i.e. 2827, 1408, 935, 699, 557,

82   462, 395, 344, 305, 273, 248, 226, 208, 192, 179, 167, 157, 147, 139, 132) were placed on

83   plots (200 x 400 m) either randomly or with aggregated spatial patterns. The aggregated

84   pattern was realized by dispersing individuals around 'parent trees' (assigned random

85   locations according to a homogeneous Poisson process), using a Gaussian dispersal kernel

86   with mean 0 and standard deviation 3 m. Thus the species distributions were modeled as a

87   Thomas cluster process, which in turn is a special case of a Neyman-Scott cluster process

88   (Neyman & Scott 1952), and this method means species are spatially independent of one

89   another.

90   Individual survival was simulated in three steps. First, a linear predictor (*y*) for survival

91   was simulated for individuals within a border of 20 m using the following multiple regression

92   equation:

93

94    $$y = \beta_0 + \beta_1 \log(ba) + \beta_2 \sum_r heterospecific\ neighbours + \beta_3 \sum_r conspecific\ neighbours \qquad (1)$$

95

96   with β_0 to β_3 the regression coefficients, *ba* the initial size (basal area) of individuals

97   and the two neighbour terms simply summing the number of heterospecific or conspecific

98   neighbours within a neighbourhood radius (*r*) of 20 m without taking size or relative size

99   differences between focal individuals and neighbours into account. The regression

100   coefficients were chosen to lead to roughly 50% mortality for each species. Specifically, $\beta_0 =$

101   5, $\beta_1 = 2.5$, and $\beta_2 = \beta_3 = -0.05$. Second, the linear predictor was converted to individual

102   survival probabilities using the inverse logit transformation. Third, binomially distributed

103   errors were used to convert the probabilities to the binary variable survival (0's and 1's) by

104   drawing from binomially distributed random numbers. Survival was then used in logistic

105   regressions as the dependent variable. Regressions were run for each species separately.

106   Standardized regression coefficients (*b*) were obtained from regressions with independent

107   variables standardized (or scaled) by subtracting their mean and dividing by their standard

108   deviation. Unlike in the neighbourhood analysis for absolute growth rate (*agr*) as dependent

109   variable (see below), fitted neighbourhood radii were fixed for the logistic regressions at 20

110   m, because best fitting neighbourhood radii for rarer species were sometime smaller than the

111   simulated 20 m radius. To investigate effects of transformations, the same multiple regression

112   approach as described above, but now with log-transformed neighbour terms was used:

113

114    $$y = \beta_0 + \beta_1 \log(ba) + \beta_2 \log(1 + \sum_r heterospecific\,neighbours) + \beta_3 \log(1 + \sum_r conspecific\,neighbours) \quad (2)$$

115

116   with β_0 and β_1 as above, and $\beta_2 = \beta_3 = -1.3$. The β's of the neighbour terms needed to be

117   adjusted in order to maintain 50% mortality. Again, unstandardized (β) and standardized (*b*)

118   regression coefficients were estimated by logistic regressions (general linear models with

119   binomially distributed error terms). Finally, β_3 and b_3, as well as variability in numbers of

120   conspecific neighbours within the neighbourhood radius *r* for each species, were correlated

121   with species abundances (i.e. log(number of individuals of each species at the plot level)).

122   In the simulations of individual growth, different distance decays and relative size

123   differences were also taken into account because competition is often size- and distance

124   dependent. Size- and distance dependencies could also have been analysed for survival as

125   dependent variable. However, since the parallel analyses yielded essentially similar

126   conclusions, we present analyses with different distance decays for growth only. For each

127   individual, one single growth increment (absolute growth rate, *agr*) was simulated for

128   individuals within a border of 20 m using the following multiple regression equation:

129

130    $$\log(agr) = \beta_0 + \beta_1 \log(ba) + \beta_2 \log(1 + \sum_r ba_{HET}/w) + \beta_3 \log(1 + \sum_r ba_{CON}/w) + error \qquad (3)$$

131

132   with $w = 1$ (no distance decay), $w =$ distance (linear distance decay) or $w =$ distance^2

133   (squared distance decay, Fig. 1). The neighbourhood terms (ba_{HET} and ba_{CON}) summed the

134   basal areas of bigger heterospecific (HET) or bigger conspecific (CON) neighbours within a

135   neighbourhood radius (r) of 20 m. The random error term was N $(0, 0.3)$. Regression

136   coefficients were $\beta_0 = -0.1$, $\beta_1 = 0.3$ and $\beta_2 = \beta_3 = -0.2$. To verify the simulations, test runs

137   with random errors set to N $(0, 0)$ were performed. The simulations were realized using C++

138   (computer code is given in Appendix A of the supplementary material).

139   Neighbourhood models (as in Stoll & Newbery 2005) were then fitted to the simulated

140   data over all possible combinations of radii for HET and CON neighbours using R (R

141   Development Core Team 2012) and parameter estimates taken from those models yielding

142   the highest adjusted R^2-values. Five runs with different seeds were performed and estimates

143   of regression coefficients from best fitting neighbourhood models, effect sizes (Cohen 1988;

144   Nakagawa & Cuthill 2007) or standardized regression coefficients (e.g. Warner 2012)

145   averaged across the five runs. Effect sizes (i.e. squared partial correlation coefficients, $t^2/[t^2 +$

146   residual degrees of freedom], t = t-value) and standardized regression coefficients ($b = \beta$'s

147   obtained from regressions with standardized variables by subtracting their mean and dividing

148   by their standard deviation) were then correlated with species abundances (i.e. plot level

149   basal area, BA, log-transformed). Standardized regression coefficients can also be calculated

150   from unstandardized β's as $b = \beta * SD_X$, if the dependent variable itself is not standardized

151   (e.g. survival). In the case of continuous dependent variables (e.g. *agr*), however, the

152   dependent variable itself is often also standardized as well. In these cases, variability in the

153   dependent variable is also involved in standardizing regression coefficients and $b = \beta * SD_X /$

154   SD_Y. In both cases, a positive correlation of b with abundance implies that less abundant, rare

155   species have stronger CON effects – β is more negative – (as in Comita $et\ al.$ 2010), whereas

156   a negative relationship implies more abundant species have stronger CON effects (as in

157   Newbery & Stoll 2013). Note, however, that possible correlations of b with abundance may

158   be biased due to correlations of SD_X (or additionally SD_Y) with abundance. But if β's are

159   negative, large SD_X lead to more negative b-values and the relationship with abundance may

160   switch direction not because of a difference in the strength of conspecific interactions

161   between rare and common species, but because of differences in the variability of number

162   and abundance of conspecific neighbours. Moreover, because the simulations and analyses

163   for both survival and agr as dependent variables are based on a multiple regression approach,

164   the basic consequences described above (i.e. possible biases in standardized regression

165   coefficients because of differences in SD_X between rare and common species) are essentially

166   the same independent of the nature of the dependent variable.

167

168   **Results**

169   There were no significant regressions for conspecific density-dependent effects

170   (regression coefficient β_3 in Eq. 1) on survival and species abundance (Fig. 2) regardless of

171   whether untransformed or log-transformed number of conspecific neighbours were used to

172   quantify the neighbourhood. However, variability in number of conspecific neighbours was

173   positively correlated with abundance if untransformed (Eq. 1) but negatively related if log-

174   transformed (Eq. 2). Consequently, standardized regression coefficients were negatively

175   correlated with abundance if number of conspecific neighbours was quantified on the

176   untransformed but positively correlated with abundance if number of conspecific neighbours

177   were log-transformed. Frequency distributions for the rarest (n = 132) and most common (n =

178   2827) simulated species on untransformed and log-transformed scales (Fig. 3) demonstrate

179   that log-transforming the number of conspecific neighbours for rare species (small values)

180   expands variability but compresses the variability in number of conspecific neighbours for

181   common species (large values). This variability in number of conspecific neighbours (SD_X)

182   increases from rare to common species on untransformed scales but decreases from rare to

183   common species on transformed scales.

184   There were no significant regressions for conspecific density-dependent effects on

185   growth (regression coefficient β_3 in Eq. 3) and species abundance (plot level basal area)

186   regardless of distance decay or spatial pattern (Fig. 4). Variation in parameter estimates was

187   largest for squared distance decay and random spatial pattern. Best fitting radii for bigger

188   conspecific neighbours were unbiased in neighbourhood models without distance decay and

189   random spatial pattern (Table 1). However, in the aggregated pattern and with linear distance

190   decay they were slightly underestimated. With estimates (mean ± SD) of 15.9 ± 2.6 in the

191   random spatial pattern and 14.5 ± 3.2, the underestimation was more pronounced with

192   squared distance decay.

193   Variability in local conspecific neighbour density (within 20 m) varied depending on

194   distance decay and spatial pattern (Fig. 5). A strong negative regression with abundance

195   emerged without distance decay in both spatial patterns. With linear distance decay, the

196   regression was not significant with random spatial pattern but still negative in the aggregated

197   pattern. With squared distance decay, the regression switched to positive in the random

198   pattern, but it was not significant in the aggregated pattern.

199   As a consequence of variation in local conspecific neighbour density, effect sizes (Fig.

200   6) and standardized regression coefficients (b_3, Fig. 7) showed various relations with

201   abundance depending on distance decay and spatial pattern. Without distance decay both

202   effect sizes and standardized regression coefficients were positively related with abundance,

203   regardless of spatial pattern. This was also the case for effect sizes and linear distance decay,

204   whereas standardized regression coefficients were not significantly related with abundance in

205   random spatial pattern but still positively related with abundance in the aggregated pattern.

206   For squared distance decay, both effect sizes and standardized regression coefficients were

207   negatively related with abundance in random spatial patterns but unrelated in aggregated

208   patterns. Apparently, the squared distance decay cancelled the effect of aggregation.

209

210   **Discussion**

211   The simulations and neighbourhood analyses with individual survival or growth as

212   dependent variable showed that estimates of regression coefficients (β) were unrelated to

213   species abundances independent of transformations, spatial pattern and distance decay — as

214   expected based on the simulations of identical species without species-specific interactions.

215   However, variability in local density of conspecifics (SD_X) showed various relationships with

216   species abundances depending on transformations of neighbourhood variables, degree of

217   spatial pattern and form of distance decay. As a consequence (i.e. $b = \beta * SD_X$), relationships

218   between effect sizes, or standardized regression coefficients (b), and species abundances were

219   either non-significant, positive or negative.

220   If untransformed scales are used to quantify conspecific neighbourhoods, relationships

221   with variability and abundance are expected to be generally positive (Fig. B2 in Appendix B)

222   at least in the cases and range of abundances investigated here. In these cases, relationships

223   between standardized effect sizes with abundance will be negative. If, however, log-

224   transformed scales are used to quantify conspecific neighbourhoods, relationships with

225   variability and abundance (Fig. B2) can be modified in all possible ways, i.e. be absent,

226   positive or negative, depending on spatial pattern, exact form of distance decay, but also on

227   whether or not relative size differences are taken into account. There are many and

228   sometimes rather non-transparent possibilities making it very difficult to systematically

229   evaluate the published literature on neighbourhood models and possible relationships

230   between the strength of CNDD and species abundance, especially where details of how

231   variables were handled are incompletely reported, and data have not been archived to allow

232   independent checks.

233   By using neighbourhood models without distance decay and unstandardized input

234   variables, in single-species analyses, a negative relationship between CNDD and forest-level

235   abundance was found, at least for the first of the two 10-year periods analyzed (Newbery &

236   Stoll 2013). Using no distance decay, yet standardizing before fitting their models, Lin *et al*.

237   (2012) found positive relationships over their dry-season interval. Using an exponential

238   distance decay, Comita *et al*. (2010) centered (subtracted the mean) but did not standardize

239   (divide by standard deviation) their input variables (L. Comita, pers. comm.) and found a

240   strong positive relationship too. Whereas Lin *et al*. (2012) fitted mixed models using

241   maximum likelihood estimation, i.e. without any prior information being involved, Comita *et*

242   *al*. (2010) used a hierarchical Bayesian analysis with non-informative priors distributed

243   according to the scaled inverse-Wishart function. This conjugate distribution models the

244   covariance matrix of the species-level regression. Nevertheless, both studies did find positive

245   relationships, thereby apparently supporting one another's conclusions.

246   The specific scale and distribution of the priors used by Comita *et al*. (2010) might

247   have introduced additional critical information that determined in part the estimation of their

248   coefficients, in a similar way as standardization did in our simulations, and may also have

249   done for Lin *et al*. (2012). Gelman & Hill (2007) discuss the use of the inverse Wishart

250   distribution in some detail, and highlight in particular the need to confirm that Bayesian

251   priors are indeed non-informative across the same ranges of independent variables that result

252 in the posterior probabilities. Dennis (1996) has discussed fundamental issues concerning the

253 use of non-informative priors and Bayesian analysis for ecology in general.

254   The results of Newbery & Stoll (2013) dealt with effects of conspecific neighbours (as

255 large tree abundance) on growth of small trees, whereas those of Comita *et al.* (2010)

256 concerned conspecific neighbour effects (as either local tree seedling density or tree

257 abundance) on survival of those seedlings. Their result could be more generally important if

258 confirmed to be fully robust to statistical treatment. It might then support the notion that

259 fundamentally different density dependent processes are likely operating at the seedling as

260 opposed to the small-tree stage in tropical forest dynamics (Uriarte *et al.* 2004 a, b; Newbery

261 & Stoll 2013).

262   Because of conceptual similarities of neighbourhood analyses of Comita *et al.* (2010)

263 and those of others (e.g. Uriarte *et al.* 2004a, b; Lin *et al.* 2012), the analysis presented here

264 could be more widely relevant. Since standardization can lead to spurious relationships

265 between CNDD and species abundances, its potential influence needs to be carefully

266 considered when interpreting relationships of small-scale effects of conspecific neighbours

267 on larger scale abundance patterns within diverse tree communities. Similarly, care should be

268 taken when specifying and justifying prior information in hierarchical Bayesian analyses. Our

269 recommendation, following from Newbery & Stoll (2013), is that tests that randomize tree

270 positions and identities indeed provide the best benchmark by which to critically evaluate and

271 judge relationships between effect sizes, or standardized regression coefficients, and tree

272 species abundances.

273

276 **Supplementary Materials**

277 **Appendix A**

278 Documented computer code used for the simulations. A detailed description of input

279 parameters and simulation output is provided in the file named growth_files.rtf.

280 **Appendix B**

281 Relationships of variability in local neighbourhoods with species abundances in different

282 competitive scenarios, spatial patterns and various distance decays.

283

284 **References**

285 Cohen, J. (1988) *Statistical power analysis for the behavioral sciences,* Second edn. L.
286   Earlbaum Associates, USA.

287 Comita, L.S., Muller-Landau, H.C., Aguilar, S. & Hubbell, S.P. (2010) Asymmetric density
288   dependence shapes species abundances in a tropical tree community. *Science,* **329,** 330-
289   332.

290 Connell, J.H., Tracey, J.G. & Webb, L.J. (1984) Compensatory recruitment, growth, and
291   mortality as factors maintaining rain forest tree diversity. *Ecological Monographs,* **54,**
292   141-164.

293 Dennis, B. (1996) Discussion: Should ecologists become Bayesians? *Ecological*
294   *Applications,* **6,** 1095-1103.

295 Gelman, A. & Hill, J. (2007) *Data analysis using regression and multilevel/hierarchical*
296   *models.* Cambridge University Press, Cambridge.

297 Lin, L.X., Comita, L.S., Zheng, Z. & Cao, M. (2012) Seasonal differentiation in density-
298   dependent seedling survival in a tropical rain forest. *Journal of Ecology,* **100,** 905-914.

299 Nakagawa, S. & Cuthill, I.C. (2007) Effect size, confidence interval and statistical
300   significance: a practical guide for biologists. *Biological Reviews,* **82,** 591-605.

301 Newbery, D.M. & Stoll, P. (2013) Relaxation of species-specific neighborhood effects in
302   Bornean rain forest under climatic perturbation. *Ecology,* **94,** 2838-2851.

303 Neyman, J. & Scott, E.L. (1952) A theory of the spatial distribution of galaxies.
304   *Astrophysical Journal,* **116,** 144-163.

305 R Development Core Team (2012) R: A language and environment for statistical computing.
306 R Foundation for Statistical Computing, Vienna, Austria.

307 Schielzeth, H. (2010) Simple means to improve the interpretability of regression coefficients.
308 *Methods in Ecology and Evolution,* **1,** 103-113.

309 Stoll, P. & Newbery, D.M. (2005) Evidence of species-specific neighborhood effects in the
310 dipterocarpaceae of a Bornean rain forest. *Ecology,* **86,** 3048-3062.

311 Uriarte, M., Condit, R., Canham, C.D. & Hubbell, S.P. (2004a) A spatially explicit model of
312 sapling growth in a tropical forest: does the identity of neighbours matter? *Journal of*
313 *Ecology,* **92,** 348-360.

314 Uriarte, U., Canham, C.D., Thompson, J. & Zimmerman, J.K. (2004b) A neighborhood
315 analysis of tree growth and survival in a hurricane-driven tropical forest. *Ecological*
316 *Monographs,* **74,** 591-614.

317 Warner, R.M. (2012) *Applied statistics: from bivariate through multivariate techniques.*
318 Sage, Los Angeles, USA.

319

319   TABLE 1. Average best fitting radii ± standard deviation (SD) for bigger conspecifics in

320   neighbourhood models (Eq. 3) across 20 species with identical initial size distributions

321   and log-species abundances and random or aggregated spatial patterns.

322

Distance decay	Spatial pattern					
	random			aggregated		
no	20.0	±	0.0	19.8	±	0.4
linear	19.6	±	0.5	19.1	±	1.0
squared	15.9	±	2.6	14.5	±	3.2

323
324

Figure legends

324

325 Figure. 1. Distance decay of neighbourhood effects. In the cut-off model (dashed), the sizes

326 of bigger neighbours with a distance < cut-off are summed. In the linear distance decay

327 (black), the sizes of bigger neighbours are weighed by 1/distance. This is similar to an

328 exponential distance decay (red), which, however, gives somewhat more weight at

329 intermediate distances. A decay of $1/distane^2$ (blue) yields a very rapidly decreasing function.

330 Beyond 20, all three functions give essentially zero weights.

331

332 Figure. 2. Regressions of conspecific effects (regression coefficient, β_3 in Eq. 1) on

333 individual survival, variability in number of conspecific neighbours within neighbourhood

334 radius ($r = 20$), and standardized regression coefficients (b_3) against species abundances

335 (number of individuals at plot level). Twenty species with identical initial size distributions

336 and log-series abundances were simulated in a random spatial pattern. Note that in the panels

337 of the bottom row, the number of neighbours was log-transformed. Data points are means ($\pm$

338 1 SD) from five replicate simulations. The simulated input value of β_3 (dotted lines) was -

339 0.05 (top left) and -1.3 (bottom left). Continuous lines indicate significant ($P < 0.05$) positive

340 (blue) or negative (red) regressions.

341

342 Figure. 3. Frequency distributions of number of conspecfic neighbours within neighbourhood

343 radius r (20 m) for individuals of rare (n = 132) and common (n = 2827) species in simulated

344 communities with random spatial patterns. Note that the x-axis in the panels of the top row

345 have identical scales. This is also true for the panels of the bottom row. Moreover, the

346 number of conspecific neighbours was log-transformed in the panels of the bottom row. SD

347 indicates the standard deviation of each distribution.

348

349   Figure. 4. Regressions of conspecific negative density dependence (regression coefficient, β_3

350   in Eq. 3) and species abundances (plot level basal area). Twenty species with identical initial

351   size distributions and log-series abundances were simulated without, linear (1/distance) or

352   squared (1/distance^2) distance decay of conspecific neighbour effects within 20 m radius in

353   random or aggregated spatial patterns. Data points are means (± 1 SD) from five replicate

354   simulations. The simulated input value of β_3 was -0.2 (green line).

355

356   Figure. 5. Regressions of variation in conspecific neighbour density (expressed as SD in basal

357   area of bigger conspecifics, ba_{CON} within 20 m) and species abundances (plot level basal

358   area). Twenty species with identical initial size distributions and log-series abundances were

359   simulated with random or aggregated spatial patterns without, linear (1/distance) or squared

360   (1/distance^2) distance decay of conspecific neighbour effects. Data points are means (± 1 SD)

361   from five replicate simulations. Continuous lines indicate significant (P < 0.05) negative (red)

362   or positive (blue) regressions.

363

364   Figure. 6. Regressions of effect sizes (squared partial correlation coefficients of β_3 in Eq. 3)

365   and species abundances (plot level basal area). Twenty species with identical initial size

366   distributions and log-series abundances were simulated with random or aggregated spatial

367   patterns without, linear (1/distance) or squared (1/distance^2) distance decay of conspecific

368   neighbour effects within 20 m radius. Data points are means (± 1 SD) from five replicate

369   simulations. Continuous lines indicate significant (P < 0.05) positive (blue) or negative (red)
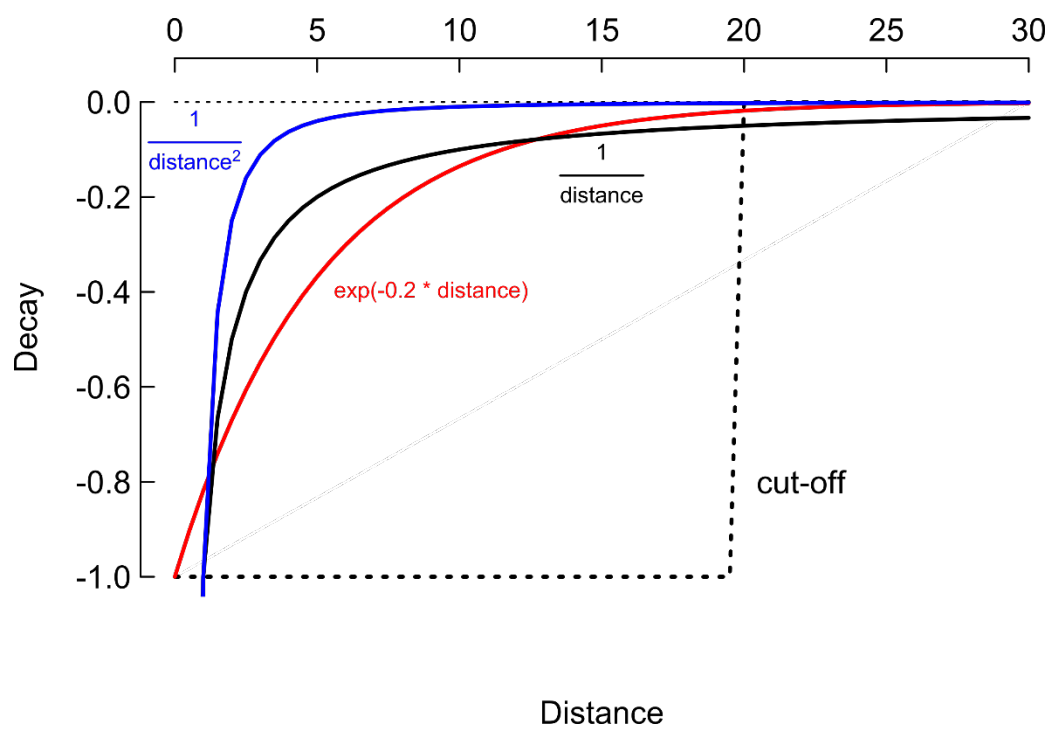
370   regressions.

371

372   Figure. 7. Regressions of standardized regression coefficients (b_3) and species abundances

373   (plot level basal area). Twenty species with identical initial size distributions and log-series

374   abundances were simulated with random or aggregated spatial patterns without, linear

375   (1/distance) or squared (1/distance^2) distance decay of conspecific neighbour effects within

376   20 m radius. Data points are means (± 1 SD) from five replicate simulations. Continuous
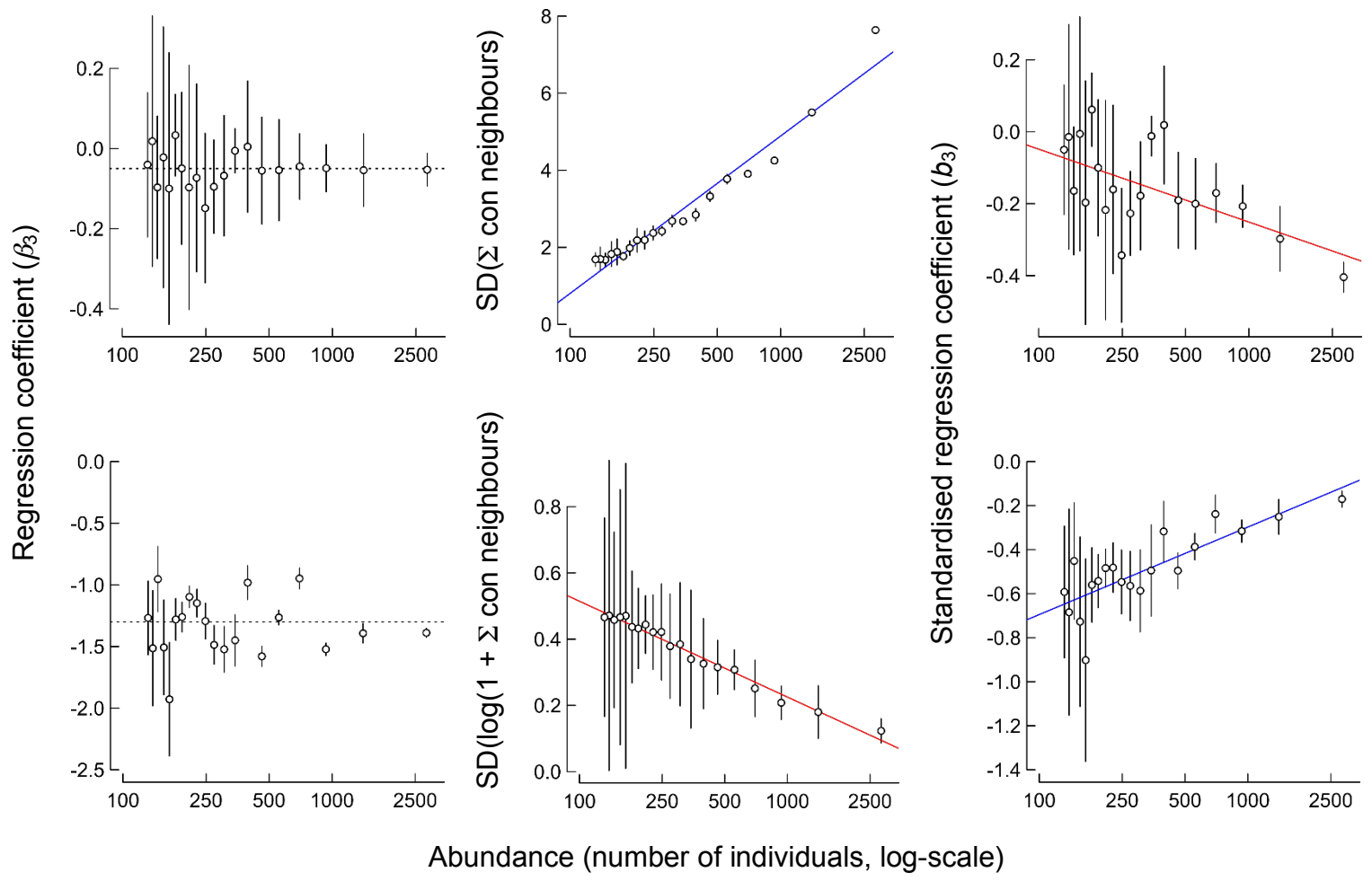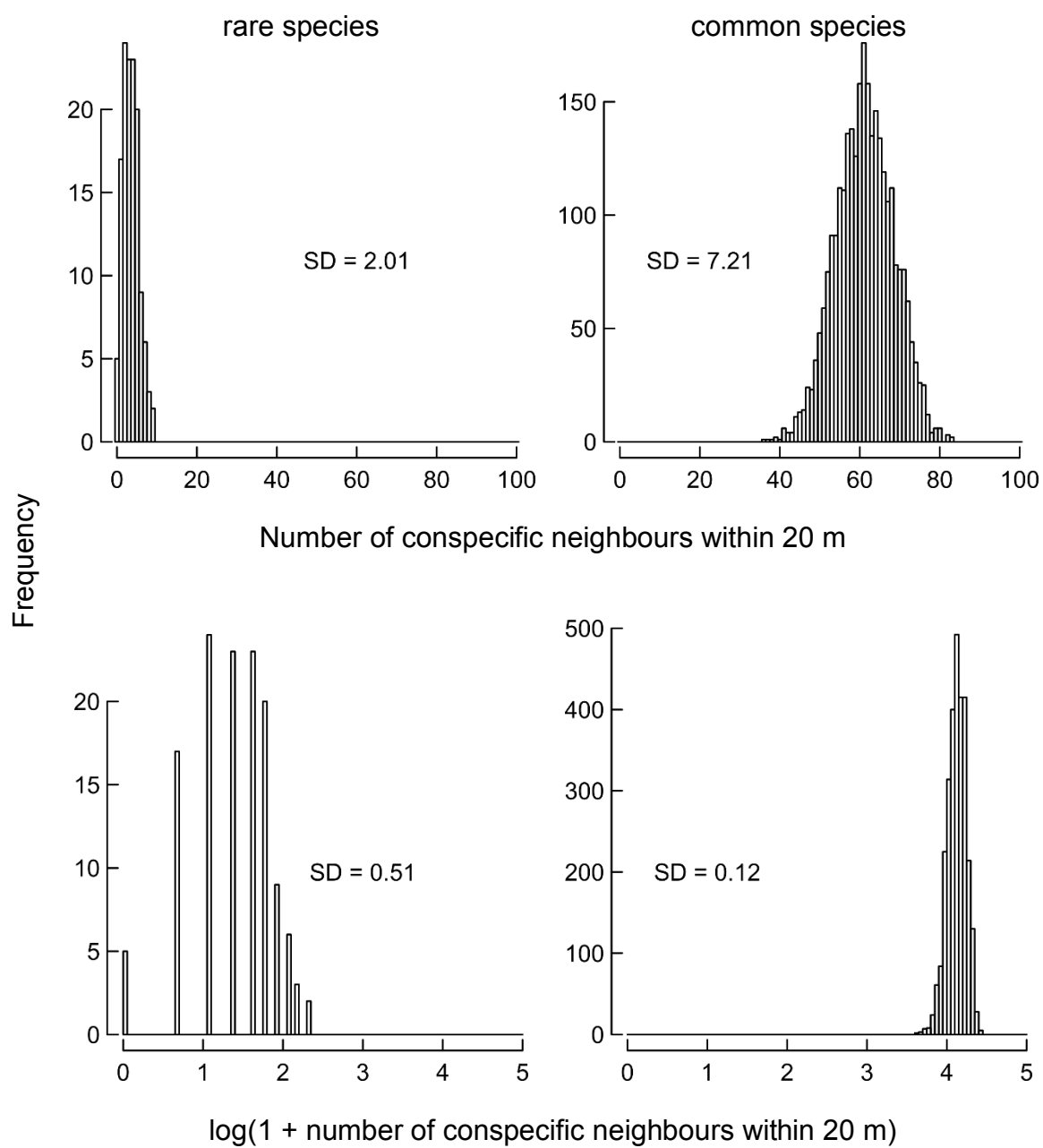
377   lines indicate significant ($P < 0.05$) positive (blue) or negative (red) regressions.

FIG. 1.

FIG. 2

rare species

common species

SD = 2.01

SD = 7.21

Number of conspecific neighbours within 20 m

SD = 0.51

SD = 0.12

log(1 + number of conspecific neighbours within 20 m)

Frequency

FIG. 3

FIG. 4.

random spatial pattern

no decay   linear decay   squared decay

Basal area (log-scale)

aggregated spatial pattern

Standard deviation (log(1 + $\Sigma\,ba_{CON}$))

Basal area (log-scale)

FIG. 5.

FIG. 6.

random spatial pattern

no decay   linear decay   squared decay

Standardized regression coefficient (b_3)

Basal area (log-scale)

aggregated spatial pattern

Basal area (log-scale)

FIG. 7.