

1 **Environmental correlates of species rank – abundance distributions in**
2 **global drylands**

3
4 Werner Ulrich^{1,*}, Santiago Soliveres², Andrew D. Thomas³, Andrew J. Dougill⁴ &
5 Fernando T. Maestre⁵

6
7 ¹Chair of Ecology and Biogeography, Nicolaus Copernicus University in Toruń
8 Lwowska 1, 87-100 Toruń, Poland, e-mail: ulrichw@umk.pl.

9 ²Institute of Plant Sciences, University of Bern, Alterbengrain 21, 3013 Bern,
10 Switzerland, e-mail: santiago.soliveres@ips.unibe.ch.

11 ³Department of Geography and Earth Sciences, Aberystwyth University, SY23 3DB,
12 UK, e-mail: ant23@aber.ac.uk

13 ⁴School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, UK, e-mail:
14 a.j.dougill@leeds.ac.uk

15 ⁵Área de Biodiversidad y Conservación, Biología y Geología, Física y Química
16 Inorgánica, Escuela Superior de Ciencias Experimentales y Tecnología, Universidad
17 Rey Juan Carlos, 28933 Móstoles, Spain, e-mail: fernando.maestre@urjc.es.

18
19 **Author for correspondence* (Phone: 0048 56 611 2649, e-mail: *ulrichw@umk.pl*)

20
21 Category: Community Ecology

22
23
24 Author contributions: SS, ADD, AJD, and FTM collected the data and provided the raw
25 data base. WU performed the data analysis. WU wrote the first draft of the manuscript,
26 and all authors contributed substantially to revisions.

28 **Abstract**

29 Theoretical models predict lognormal species abundance distributions (SADs) in stable
30 and productive environments, with log-series SADs in less stable, dispersal driven
31 communities. We studied patterns of relative species abundances of perennial vascular
32 plants in global dryland communities to: i) assess the influence of climatic and soil
33 characteristics on the observed SADs, ii) infer how environmental variability influences
34 relative abundances, and iii) evaluate how colonisation dynamics and environmental
35 filters shape abundance distributions. We fitted lognormal and log-series SADs to 91
36 sites containing at least 15 species of perennial vascular plants. The dependence of
37 species relative abundances on soil and climate variables was assessed using general
38 linear models. Irrespective of habitat type and latitude, the majority of the SADs
39 (70.3%) were best described by a lognormal distribution. The lognormal was associated
40 with low annual precipitation, higher aridity, high soil carbon content, and higher
41 variability of climate variables and soil nitrate. Our results do not corroborate models
42 predicting the prevalence of log-series SADs in dryland communities. As lognormal
43 SADs were particularly associated with sites with drier conditions and a higher
44 environmental variability, we reject models linking lognormality to environmental
45 stability and high productivity conditions. Instead our results point to the prevalence of
46 lognormal SADs in variable and stressful ecosystems, which are generally shaped by
47 strong habitat filters and limited colonisation. This suggests that drylands may be
48 resilient to environmental changes because the many species with intermediate relative
49 abundances could take over ecosystem functioning if the environment becomes
50 suboptimal for dominant species.

51

52 **Running title:** Rank – abundance distributions of dryland plants

53 **Key Words:** aridity, species abundance, competition, lognormal distribution, log-series
54 distribution, habitat filtering, soil fertility, climate

55

56

57 **Introduction**

58

59 Since its introduction to ecology by Raunkiær (1909), species abundance
60 distributions (SADs) have been extensively studied (reviewed in McGill et al., 2007;
61 Matthew and Whittaker, 2014, 2015). They provide an exhaustive description of the
62 distribution of species abundances within an ecological community (Magurran, 2004;
63 McGill et al., 2007; Dornelas et al., 2011; Matthews and Whittaker, 2015) and have
64 been linked to differential resource use and competitive strength (Sugihara, 1980;
65 Tokeshi, 1998; Pueyo, 2006), disturbance regimes (Gray and Mirza, 1979), stochastic
66 processes (May, 1975, Šizling et al., 2009), or species-specific dispersal rates (Hubbell,
67 2001; Zillio and Condit, 2007).

68 SADs can be grouped into two particular classes of distributions: the log-series and
69 the lognormal (Fig. 1; Connolly et al., 2005; Ulrich et al., 2010, 2016). The lognormal is
70 characterized by a comparably high number of species with intermediate abundance and
71 smaller numbers of very abundant and very rare species (Fig. 1). In turn, the log-series
72 lacks a distinct group of subdominant species, typically showing less even distributions
73 (Fig. 1). Although it is difficult to relate these models to a particular underlying
74 mechanism (cf. McGill et al., 2007; Ulrich et al., 2010; Cheng et al., 2012; Locey and
75 White, 2013, but see Alonso et al., 2008), lognormal SADs are more likely to be found
76 in closed communities with low temporal and spatial species turnover and a high
77 proportion of species with intermediate abundances (that is the lognormal represents
78 communities with a proper ‘middle class’ of species) (Magurran and Henderson, 2003)
79 if they are shaped 1) by multiple stochastic processes, independent of niche
80 differentiation, resource use, or competitive ability as predicted by the central limit
81 theorem of statistics (Preston, 1948; May, 1975; Connolly, et al. 2005; Šizling et al.,
82 2009), 2) by sequential niche partitioning, where competitive strength with respect to
83 dominant niche axes governs the distribution of species abundances (MacArthur, 1957;
84 Sugihara, 1980; Tokeshi, 1998; Pueyo, 2006), or 3) by environmental filters, such as
85 climate and soil characteristics that select for certain species and species combinations
86 and limit colonisation (Green and Plotkin, 2007; Zillio and Condit, 2007; Maire et al.,
87 2012). On the other hand, log-series SADs are expected to occur 1) in open colonisation
88 driven communities with high degrees of dispersal and species turnover (Volkov et al.,

89 2005; Zillio and Condit, 2007; Hirao et al., 2012) or 2) in incomplete samples from
90 larger species pools (Fisher et al., 1943).

91 Species abundance distributions have often been theoretically linked to
92 environmental conditions and gradients (reviewed in Magurran, 2004; McGill et al.,
93 2007; Dornelas et al., 2011). Some authors assume that lognormal SADs prevail in
94 stable, undisturbed environments, while log-series SADs will be found in disturbed
95 habitats with higher temporal or spatial variability (e.g. Gray et al., 1979; Gray and
96 Mirza, 1979; Hamer et al., 1997; Hill and Hamer, 1998; but see Nummelin, 1998).
97 Whittaker (1975) and Hubbell (1979) linked lognormal SADs to higher environmental
98 productivity. Consequently, log-series SADs should predominate at unproductive, e.g.
99 arid, sites. However, the direct influence of environmental conditions on abundance
100 distributions has been very rarely studied empirically. The few existing studies mainly
101 focus on community recovery after severe disturbances (Mouillot et al., 2000), gradients
102 of environmental pollution (e.g. Gray et al., 1979; Death, 1996; Qu et al., 2008), and
103 successional stages (e.g. Whittaker, 1965; Bazzaz, 1975; Zaplata et al., 2013). Taken
104 together, current evidence indicates that a directional shift from log-series towards
105 lognormal SADs may occur with increasing intensity of interspecific competitive
106 interactions and habitat stability (Tilman, 1982; Lan and Bai, 2012).

107 Our knowledge about plant species abundance distributions stems mainly from
108 work done in forests (Hubbell, 1979; Morlon et al., 2009; Ulrich et al., 2015) and
109 temperate grasslands (Bazzaz, 1975; Maire et al., 2012). With the exception of
110 Whittaker's (1965) classical report of a lognormal SAD for Arizona desert plants,
111 similar distributions in arid, semi-arid and dry-subhumid regions (drylands hereafter)
112 have so far not been studied. Drylands, including a variety of habitat types like
113 grasslands, scrublands and savannahs, occupy more than 40% of the terrestrial surface
114 area (Safriel and Adeel, 2005) and are vulnerable to human disturbances (Maestre et al.
115 2012a) and changing climate (Körner, 2000; Reynolds et al., 2007; Dai, 2013), which in
116 turn affect nutrient cycles (Maestre et al., 2012b). We do not know whether the
117 abundance patterns observed in forests can be generalised to drylands, and how changes
118 in environmental conditions affect the SADs of dryland communities. As plant
119 abundances are directly related to important ecosystem functions in drylands, like
120 primary production and nutrient cycling (Gaitán et al., 2014; Maestre and Escudero,
121 2009), such knowledge can also greatly contribute to our understanding of the

122 consequences of global change on ecosystem functioning in these areas (Maestre et al.,
123 2012a; Maire et al., 2012).

124 Here we evaluate how environmental factors affect the SADs of 91 dryland
125 communities from all continents except Antarctica and from three different vegetation
126 types obtained within an international, large-scale dryland survey (Maestre et al., 2012b,
127 Delgado-Baquerizo et al., 2013). We focus on the gradient between the log-series and
128 the lognormal type SAD. Based on the available knowledge, we assumed that highly
129 variable environmental conditions would favour unstable and dispersal-driven
130 communities (reviewed in Fraterrigo and Rusak, 2008), while water-rich, productive
131 environments favour stable, competition driven communities (Whittaker, 1979;
132 Hubbell, 1979). These assumptions lead to three basic starting hypotheses regarding
133 dryland plant communities:

- 134 (1) Dryland communities generally show log-series SADs, as they are dominated by
135 habitat filtering and perform under low productive environments,
- 136 (2) woody communities are dominated by fewer species, in accordance to previously
137 observed forests data and, therefore, their SADs fit better to log-series distributions than
138 those of grasslands, which are more evenly distributed (ref),
- 139 (3) more arid, and therefore less productive, communities are dominated by log-series
140 SADs,
- 141 (4) lognormal SADs dominate in species rich, communities, and
- 142 (5) log-series SADs are linked to both increased environmental variability and
143 decreased importance of habitat filtering.

144

145 **Materials and methods**

146 Study sites and sampling protocol

147

148 Field data were obtained from 230 sites established across precipitation gradients in 17
149 countries from five continents (Argentina, Australia, Botswana, Brazil, Chile, China,
150 Ecuador, Iran, Israel, Kenya, Mexico, Morocco, Peru, Spain, Tunisia, USA and
151 Venezuela). Sites were chosen to cover a wide spectrum of abiotic (climatic, soil type,
152 slope) and biotic (type of vegetation, total cover, species richness) features
153 characterizing drylands worldwide. These sites include the 224 sites used in Maestre et
154 al. (2012b) plus six additional sites in Botswana surveyed in 2012. We restricted our

155 study to arid, semi-arid and dry-subhumid ecosystems, defined as sites with an aridity
156 index (precipitation/potential evapotranspiration) between 0.05 and 0.65. The sites
157 cover all major biogeographic regions and four basic vegetation types (woodlands,
158 savannahs, scrublands, and grasslands). All study sites were sampled quantitatively
159 following the same protocol. At each site, we surveyed 80 1.5 m × 1.5 m quadrats along
160 four 30-m long transects separated eight meters from each other (see Maestre et al.
161 2012b for full methodology). In each quadrat, we measured the cover of perennial plant
162 species and used the total counts to construct the respective vectors of relative
163 abundances. Thus all abundance distributions are based on complete censuses.

164 A low number of species per site increases the noise in the SAD fits (Wilson et al.
165 1998), while selecting a high minimum number of species greatly reduces the number
166 of sites (and vegetation types) considered making statistical inferences challenging. As a
167 compromise, we retained 91 of the study sites, which had ≥ 15 species of perennial
168 vascular plants. Nevertheless, and to assess the robustness of our analysis, we compared
169 the results obtained from these sites with those obtained from an extended data set (166
170 sites) including at least 10 species (as recommended by Ulrich et al, 2010 as the lower
171 limit for reliable fits) and from a reduced data set (55 sites) including at least 20 species
172 per site. As the results from these three data sets were qualitatively similar, we only
173 report the results obtained with the 91 sites having 15 species or more. We show the
174 results obtained with the reduced and extended data sets in the electronic supplement.

175

176 Biotic and abiotic factors

177 Using a stratified sampling design, we sampled the top 7.5 cm of the soil from up to
178 three different microhabitats per site. These microhabitats always included a location
179 with bare soil (i.e. devoid of perennial vascular plants), as well as sites dominated by
180 perennial vegetation (e.g. under trees, shrubs or grasses, depending on the dominant
181 growth forms present within each site). Five samples were collected from each
182 microsite, yielding between 10 and 15 samples per site. Soil samples were air-dried at
183 room temperature, sieved (< 2 mm fraction) and analysed in the laboratory to obtain a
184 range of physio-chemical analyses. In each soil sample we measured pH, organic
185 carbon, available phosphorus, and nitrate content as described in Maestre et al. (2012b).
186 These variables were selected because they are either appropriate surrogates of overall
187 soil fertility and nutrient availability for plants in drylands (carbon and nitrogen

188 variables; Whitford, 2002) or they are surrogates of abiotic variables that control
189 nutrient transformations and availability in soils (e.g. pH; Reth et al., 2005). Thus, we
190 expect them to be important factors influencing the relative abundance distributions of
191 plant species. Soil variables were pooled to a single site-level value by weighting the
192 values found underneath vegetation or in bare ground areas by their respective cover
193 within the site (cf. Maestre et al., 2012b). As a measure of habitat variability, we
194 calculated for the four soil variables their respective coefficients of variations based on
195 the 10-15 samples obtained per site.

196 We also obtained climatic data for each site using Worldclim
197 (<http://www.worldclim.org>; Hijmans et al., 2005). From this database, we extracted the
198 altitude of each site, the mean annual temperature and precipitation, and their annual
199 seasonality. As we expected to see changes in relative abundances along climatic
200 gradients, particularly along the gradient from moist to dry, we calculated the UNEP
201 aridity index as the quotient of annual precipitation and evapotranspiration. To give a
202 more readily interpretable result, we used the aridity level (1- aridity), which is directly
203 related to aridity (higher values indicate higher aridity conditions). Aridity was
204 estimated using the Global Aridity Index (Global-Aridity) dataset ([http://www.cgiar-](http://www.cgiar-csi.org/data/global-aridity-and-pet-database)
205 [csi.org/data/global-aridity-and-pet-database](http://www.cgiar-csi.org/data/global-aridity-and-pet-database); Zomer et al., 2008; Trabucco and Zomer,
206 2009), which is based on the interpolations provided by the Worldclim database.

207

208 Fitting of relative abundances

209 We fitted lognormal (fit_{norm}) and log-series (fit_{lser}) models to the observed SADs as
210 in Ulrich et al. (2010). For this task we used rank- log abundance (Whittaker) plots that
211 show the log-transformed species abundances for each species ranked in declining
212 abundance order (Fig. 1). These plots are superior to classical distribution (Preston)
213 plots for fitting as they do not lose information and are not biased due to the grouping of
214 species (Nekola et al., 2008, Ulrich et al., 2010). For each rank – log abundance plot we
215 used a maximisation algorithm (implemented in the software application RAD 2.0,
216 Ulrich 2013) that iteratively encapsulates parameter values to find the ones that
217 minimise the average least square differences of observed and predicted relative
218 abundance, respectively

$$219 \quad fit = \frac{\sum_{i=1}^S (\ln A_{i,obs} - \ln A_{i,pred})^2}{S} \quad (1)$$

220 where $A_{i,obs}$ and $A_{i,pred}$ are the respective relative abundances of species i in the

221 community of S species. Fits for all communities are contained in the electronic
222 supplement. We used least squares differences for fitting as they put comparably high
223 weight on rare and abundant species (Connolly and Dornelas, 2011) thus increasing the
224 power to discriminate between the lognormal and the log-series models (Ulrich et al.,
225 2010). In this respect, we note that major axis and reduced major axis have less
226 discriminative power in the present context as both methods put higher weight on
227 species with intermediate abundance.

228 As *fit* (eq. 1) χ^2 equals the residuals sums of squares we compared the relative fits of
229 both distributions using the corrected Akaike information criterion in the form

$$230 \quad AICc = 2k + S \ln fit + \frac{2k(k+1)}{S-k-1} \quad (2)$$

231 The lognormal SAD has $k = 3$ free parameters (richness S , shape, and error), the
232 logseries is a four parameter model (S , α , X , and error). We used ΔAIC_c to identify the
233 best fitting model and assigned models with $\Delta AIC_c > |10|$ as fitting significantly better
234 (Burnham and Anderson 2002). As species differ in the probability to obtain particular
235 least squares values (Connolly and Dornelas, 2011), least squares fitting applied to non-
236 linear data might introduce a statistical bias when comparing SADs of different species
237 richness. We minimized this possible bias two-fold: first, we always compared the two
238 model fits for the same community and second, we included species richness as an extra
239 predictor in our analyses. Locey and White (2013) highlighted the problem of
240 comparing SADs from communities with different species richness and total abundance.
241 Here we minimize this problem as we always fit both models to the same community
242 and subsequently compare the respective relative fits among communities.

243 Ulrich et al. (2010) studied a third basic shape, the power function, and found it to
244 be rarely realised in natural communities except for some forest tree data. Nevertheless,
245 we checked the frequency of power function SADs in the global dryland data set. Our
246 data confirmed the results of Ulrich et al. (2010) and revealed a low power to
247 discriminate between log-series and power function shapes. Thus, we did not consider
248 this model here, but present respective numbers of best and worst fits of all three
249 models (lognormal, log-series and the power function) in the electronic supplement
250 (Table A9).

251 An auxiliary measure of model fit is the skewness of the abundance distribution (γ).
252 The symmetrical lognormal is not skewed. Unsymmetrical lognormal SADs have nearly
253 always an excess of rare species, and consequently a negative skewness (McGill, 2003).

254 The log-series has an excess of relatively abundant species (associated with a positive
255 skewness) mostly in the case of incomplete sampling. An excess of relatively rare
256 species (negative skewness) has been theoretically linked to communities characterised
257 by high colonisation dynamics (Zillio and Condit, 2007).

258 As an approximate measure of SAD variance, the concept of evenness is closely
259 related to the distribution of relative abundances (McGill et al., 2007). We assessed the
260 evenness (E) in species abundances using the Shannon diversity metric H : $E = H/\ln(S)$.
261 p_{norm} , skewness, and evenness values for each site are available from figshare (Maestre
262 et al., 2015).

263

264 Statistical analyses

265 -Assessing the relationships between SADs and site productivity and species richness
266 We used ordinary least squares general linear model analysis (GLM) in AICc model
267 selection to link the ΔAIC_c scores (eq. 2) to environmental data. Environmental data
268 included those variables directly or indirectly related to site productivity, such as
269 elevation, temperature, rainfall, soil pH, organic C, available P and nitrate. We added
270 species richness as an additional covariate, to evaluate the relationship between the
271 richness of each community and its observed SAD. Our SAD fits and predictors were
272 moderately spatially autocorrelated (Moran's $I < 0.5$). However, the global distribution
273 of sites studied would cause any spatially explicit modelling, like simultaneous
274 autoregression modelling or similar techniques, to artificially concentrate a large part of
275 the variance in environmental data in the spatial distance matrix, masking thereby the
276 underlying influences of the environment (Hawkins, 2012). However, and to account
277 for the spatial structure present in our data, we included the dominant eigenvector of the
278 associated geographical distance matrix as an additional predictor in the GLM analyses
279 (Hawkins, 2012). This dominant spatial eigenvector covered the large scale spatial
280 structure of the sites and explained 85% of total variance in the geographical distance
281 matrix.

282 We selected as the most parsimonious models those with the lowest AICc, using
283 the model selection routine of SAM 4.0 (Rangel et al., 2010). To verify our first to
284 starting hypotheses on the dependence of abundance distributions on environmental
285 states we related ΔAIC_c , skewness, and evenness to latitude (and squared latitude),

286 climatic and soil variables. Our second hypothesis was then tested by analyzing the
287 relationship between the ΔAIC_c of each community and its species richness.

288

289 -Evaluating the relationship between SADs and environmental (soil and climate)
290 variability

291 As our third starting hypothesis is about the influence of environmental variability we
292 run separate models using the coefficients of variation of these environmental variables
293 as predictors. Pearson correlation coefficients between predictor variables were always
294 lower than 0.7, and therefore multicollinearity problems in our analyses are unlikely.
295 Because vegetation type is strongly linked to temperature and precipitation, we did not
296 include vegetation type as a categorical variable into the regression models to avoid
297 multicollinearity problems. To account for possible non-linearity and non-normal error
298 structures, we compared these results with those obtained from generalised linear
299 modelling using log-link functions and Poisson error structure. As this latter approach
300 did not improve our results and was largely consistent with the main analyses shown
301 here, we only present them in the electronic supplement (Tables A7 and A8).

302 We used additive variance partitioning to assess the effects of single environmental
303 predictors on ΔAIC_c , skewness, and evenness. The data used for the present study are
304 available from figshare (Maestre et al., 2015).

305

306 **Results**

307 -General patterns of species abundance distributions in drylands

308 At the global scale the lognormal model fitted better ($\Delta AIC_c < 0$) for 64 (70.3%) and
309 definitely better ($\Delta AIC_c < -10$) for 58 of the 91 communities with at least 15 species
310 (40.7%; Table 1). Only 10 communities (10.0%) were definitely better fitted by a log-
311 series ($p_{norm} > 10$) while 23 communities (25.3%) scored intermediate ($-10 \leq p_{norm} \leq 10$).
312 Although we found a prevalence of lognormal distributions in each vegetation type
313 (Table 1), there was slight indication that these differ with respect to SAD fit (one-way
314 ANOVA: $F_{3,87} = 3.7$, $P = 0.02$). Tukey post-hoc comparisons point to grasslands as
315 having a lower proportion of lognormal type communities (Table.1). Including sites
316 with as few as 10 species made the results increasingly noisy (electronic supplement
317 Table A1) while at ≥ 20 species per site (Table A2) results were qualitatively identical to
318 those presented above.

319

320 -Assessing the relationships between SADs and site productivity and species richness
321 There was a significant latitudinal gradient in ΔAICc indicating better fits of the
322 lognormal in the Mediterranean communities (GLM $r^2 = 0.17$, $P < 0.01$). South
323 American communities tended to be better fitted by the log-series than Old World
324 communities (GLM $r^2 = 0.11$, $P < 0.05$). Evenness peaked around the equator and
325 decreased with increasing latitude (GLM quadratic regression $r^2 = 0.08$, P [quadratic
326 regression term] = 0.01), while skewness did not significantly vary with latitude ($r^2 =$
327 0.03, n.s.). After accounting for the effects of species richness and spatial
328 autocorrelation, average annual precipitation was negatively linked to the fit of the
329 lognormal model (Table 2, Table A4), and explained 8% of the variance in ΔAICc .
330 Communities best described by a log-series occurred along the whole gradient of
331 precipitation while better fits of the lognormal were largely restricted to values of
332 annual precipitation below 650 mm (Fig. 2a, ANOVA $F_{1,89} = 5.1$, $P < 0.05$, Fig. A2).
333 However, within those communities with aridity levels > 0.5 there was a trend towards
334 log-series-distributed SADs at increased arid environments (Fig. 2b, GLM $r^2 = 0.05$, $P <$
335 0.05). This trend was supported by the reduced data set (at least 20 species per site
336 included: Fig. A2, GLM $r^2 = 0.16$, $P < 0.01$). Among the soil variables, only carbon was
337 consistently included in the regression models for ΔAICc (Table 2, Tabs. A3, A4), and
338 explained 6% of the variance. ΔAICc decreased with increasing soil carbon content
339 (Table 2) indicating a better fit of the lognormal in richer soils. This carbon influence
340 was also corroborated by GLM Poisson regression (Table A7). Finally, we found
341 ΔAICc to be positively linked to available phosphorus (Table 2, 6% variance
342 explanation, and Table A7).

343 Positive and negative skewness measure the proportions of abundant and rare
344 species, respectively. AICc model selection pointed to carbon content (Table 2) as
345 affecting skewness, although this variable explained less than 5% of variance and
346 consequently was insignificant in the reduced data set (Table A4) and the GLM Poisson
347 model (Table A7). Evenness was negatively linked to soil carbon content (11% of
348 variance explained) and these results were consistent regardless of the data subset used
349 (Table 2, Tables A3, A4, A6).

350

351 -Evaluating the relationship between SADs and environmental (soil and climate)
352 variability
353 The relative fit of the lognormal model increased with increasing seasonality in
354 temperature (Table 3, A5, A6, A8) while seasonality in precipitation had no significant
355 effect (Table 3, Tables A5, A6, A8). Despite of the lack of clear regressive trends
356 linking AICc and soil variability (Table 3, Tables A5, A6, A8), our data indicate a
357 distinction of model fit with respect to nitrate variability (Fig. 3a, Fig. A4A).
358 Communities fitted better by a log-series were largely restricted to low nitrate
359 variability. Further, lognormal communities significantly decreased in skewness (Fig.
360 3b, $r^2 = 0.17$, Fig. A4B) and increased in evenness (Fig. 3c, $r^2 = 0.16$, Fig. A4C) at
361 higher nitrate variability, while there were no such trends for log-series communities
362 (Figs. 3b, c).

363

364 **Discussion**

365 -General patterns of species abundance distributions in drylands
366 Contrary to our first starting hypothesis (arid communities should be dominated by log-
367 series SADs), our study adds dryland plants to the group of communities with a
368 prevalence of lognormal abundance distributions (e.g. Tokeshi, 1998; Magurran and
369 Henderson, 2003; Connolly et al., 2005; Ulrich et al., 2010). Irrespective of dryland
370 habitat type (Table 1), we found that nearly 2/3 of the communities studied were fitted
371 better by the lognormal model, which predicts a relative excess of species with
372 intermediate abundance. This finding is in line with the only comparable study by
373 Whittaker (1965) on desert plant communities, but contrasts to results obtained with
374 forest tree communities (Ulrich et al. 2010). Also Leigh (1999), Morlon et al. (2009),
375 and Ulrich et al. (2016) have reported log-series abundance distributions to prevail
376 particularly in tropical forest communities. Our results do not exclude the possibility
377 that abundance distributions of dryland vegetation types, in general, differ from more
378 humid forest communities. Therefore our results demand caution about the
379 generalisation of abundance patterns obtained from single ecosystems types and their
380 transfer to dryland ecosystems.

381 The contrasting results from forest and the present dryland studies call for a
382 mechanistic explanation. The forest data studied by Morlon et al. (2009) and Ulrich et
383 al. (2015) represent to a large extent secondary succession forests and plantations. These

384 are generally characterised by small numbers of highly abundant and larger numbers of
385 rare species, and thus lack the group of intermediately abundant species that
386 characterizes a lognormal distribution (Preston, 1948). Such communities show a
387 comparably low degree of evenness and this community organisation is more in line
388 with a log-series. Studies on boreal forests, containing a relatively low number of very
389 abundant species (often even mono-stands) also reported log-series distributions
390 (Whittaker, 1960). Similarly, in species-rich coral reefs (Connolly et al., 2005) and in
391 tropical and relatively pristine forest communities (Hubbell, 1979; Volkov et al., 2003;
392 Cheng et al., 2012) lognormally organised communities seem to prevail. While our
393 study sites comprise areas with different degrees of human activities, none of the
394 studied sites are subject to intensive management areas such as cropping, fertilization or
395 planting of species (Maestre et al., 2012b). Thus, our results and those from the
396 literature indicate that less impacted ecosystems have a higher probability to follow
397 lognormal species abundance distributions. Consequently, these dryland systems tend to
398 accumulate a ‘middle class’ of species with intermediate relative abundances. Having
399 such a class may make these systems more resistant to functional disturbance because
400 these species might take over ecosystem functioning if the environment becomes
401 suboptimal for the dominant ones, potentially enhancing the resilience to environmental
402 changes (Walker et al., 1999).

403 About a quarter of the communities (25.3%, Table 1) were roughly equally fitted by
404 both models. This pattern is in line with previous reports (e.g. Hughes, 1986; Magurran
405 and Henderson, 2003; Ulrich and Ollik, 2004; Dornelas and Connolly, 2008; Vergnon et
406 al., 2012), who observed that SADs may be compound functions that capture
407 contrasting parts of local communities and patterns of community assembly. These
408 SADs might comprise on one side the stable elements of resident species following a
409 lognormal distribution and on the other site so-called satellite species having a high
410 temporal dynamic and thus being best described by the log-series (Magurran and
411 Henderson, 2003). Surprisingly, up to now there is no systematic empirical study on
412 how well the compound model fits to SADs in communities across a variety of habitat-
413 types and differing environmental conditions. Apart from the dynamics model of
414 Hughes (1986) and recent work on speciation driven neutral communities (Vergnon et
415 al., 2012) and hidden niche models (Barabás et al., 2013) focusing on multimodality,
416 there is also no explicit theoretical model to predict the precise SAD shape.

417 The large proportion of intermediate SADs also indicates that lognormal and log-
418 series SADs rather mark both endpoints of a continuum within which very different
419 dominance structures might be realised (Magurran and Henderson, 2003). We speculate
420 that the position within this continuum provides information about the trade-off between
421 species interactions and colonisation – extinction dynamics by which a focal community
422 is shaped. This trade-off should be triggered by the regional species pool size (the
423 colonisation pressure), but also by environmental drivers that act as filters for potential
424 colonisers. Both processes position a focal community into this continuum of SAD
425 shapes. The fact that nearly half of our communities ranked intermediate on this
426 continuum makes it probable that dryland communities are assembled by the interplay
427 of colonisation dynamics and competitive interactions.

428

429 Environmental triggers

430 Based on the global positive co-variation of species richness and productivity
431 (Whittaker, 1975; Currie, 1991, but see Adler et al., 2011), Whittaker (1975) and
432 Hubbell (1979) initiated the idea that SADs are linked to productivity gradients, with
433 increasing lognormality at higher levels of productivity. Therefore, we expected to see a
434 negative correlation of our AICc measure with average precipitation and a respective
435 positive correlation with aridity (hypothesis 2), as plant cover and productivity decrease
436 with increasing aridity (Safriel and Adeel, 2005; Delgado-Baquerizo et al., 2013). This
437 was not the case, as rather we found the opposite pattern between AICc and annual
438 precipitation (Table 2, Fig. 2a), and also a slightly negative effect of species richness on
439 AICc (Table 2). Interestingly, Ulrich et al. (2015) reported a similar negative correlation
440 of the fit of the lognormal distribution with precipitation and also with
441 evapotranspiration in global forest communities. Therefore, both results do not
442 corroborate the productivity hypothesis.

443 This finding links the occurrence of lognormally distributed communities to sites
444 with higher environmental (in this case water) stress. Ecological theory mainly predicts
445 a connection of stress with the log-series, although we note that existing evidence for
446 this assumption is scarce (Gray et al., 1979; Gray and Mirza, 1979; Death, 1996; McGill
447 et al., 2007; Qu et al., 2008). Our results point to strong effects of habitat filtering, and
448 consequently limited dispersal in stressful environments as the major process shaping
449 SADs. Average conditions filter specific sets of species (Wiens and Graham, 2005), and

450 the abundance rank orders are established in a subsequent step by the interplay of
451 species interactions, reproductive success, and local extinction (McGill et al., 2007).
452 Therefore, variability in environmental conditions appears to be more important for the
453 variation in species composition and abundances between sites than average conditions
454 (Violle et al., 2012). Indeed, we found significant, albeit contrasting, relationships
455 between AICc and the variability in temperature (Table 3). These results are partly in
456 accordance with our third hypothesis (i.e. log-series SADs should be linked to both
457 increased environmental variability and decreased importance of habitat filtering), and
458 indicate the existence of trade-offs in habitat variability with regard to certain
459 abundance distributions, thus complicating the simple environmental variability –
460 lognormal view (Gray et al., 1979; Hamer et al., 1997; Hill and Hamer, 1998).

461 Only variability in soil carbon content entered the best fit regression model, and
462 thus soil variability appeared to be much less influential than climate variability as a
463 driver of the variation found in the SADs. However, nitrate variability (Fig. 3) might act
464 differently, determining thresholds for community structure. We were surprised to find
465 log-series SADs to be limited to soils with low nitrate variability (Fig. 3a). As nitrate
466 variability also caused a negative skewness (Fig. 3b) and an increased community
467 evenness (Fig. 3c), it apparently forces communities towards lognormal abundance
468 structures with a small number of very rare species. These SADs are not predicted from
469 colonisation driven models that possess a heavy tail of relatively rare species, for
470 instance neutral models without dispersal limitation (Hubbell, 2001; Zillio and Condit,
471 2007). Our results thus clearly point to variability as a mechanism promoting the
472 emergence of lognormal distributions (Fig. 3a) and limiting local colonisation dynamics
473 (Figs. 3b, c). Consequently, our findings do not corroborate the opposed variability –
474 log-series model that predicts disturbed or unstable sites to have log-series distributed
475 communities (Gray et al., 1979; Zillio and Condit, 2007). A mechanistic explanation for
476 this result invokes that high small-scale soil variability induces the development of a
477 patchy community organisation with many intermediate and low abundant species that,
478 when pooled to samples, nevertheless exhibit a higher evenness than expected from a
479 homogeneous environment (equivalent to statistical averaging, Lehman and Tilman,
480 2000). Such a patchy distribution of soil nutrients is often exacerbated by even light
481 levels of grazing and shifts seen towards increased shrub canopy cover (Berkeley et al.,
482 2005). Further this patchy distribution prevents species from becoming locally very

483 abundant, thus reducing the number of dominant species in line with the spatial storage
484 effect (Sears and Chesson, 2007). Alternative explanations for the prevalence of
485 lognormal SADs in more heterogeneous environments is the generalized lack of
486 competition hierarchy (intransitive competition) in drylands, which increases co-
487 dominance of a relative large number of species and is enhanced by environmental
488 heterogeneity (Soliveres et al. 2015). Alternatively, temporal storage effects (Chesson
489 2000) could prevent the dominance of a single species and should become more
490 frequent with rainfall or temperature variability. Temporal storage effects, however, do
491 not seem a plausible explanations for the prevalence of lognormal SADs in drylands, as
492 the variability of both temperature and rainfall caused SAD distributions to better fit
493 log-series rather than lognormal distributions (Table 3). Regardless of the underlying
494 mechanism, the trigger for the negative skewness is not only caused by an increased
495 number of very rare species but also by the low number of very abundant species.
496 Indeed, small-scale soil variability is known to induce vicariant plant species
497 composition and phylogenetic structure (Schreeg et al., 2010; Ulrich et al., 2014),
498 reducing the dominance of the most competitive species. In turn, dispersion-driven
499 variability in species composition favours log-series abundance distributions. Thus
500 variability in community composition induced by environmental factors and dispersal
501 might act in opposite directions. We hypothesise that if environmental variability also
502 affects composition, the outcome might be unpredictable and often intermediate
503 between both types of dominance order.

504 The above picture is complicated by the fact that our environmental variables
505 accounted for at most 35% of the variances in dominance structure (Table 3). This is the
506 point where biotic interactions might step in. As the species found within each plot had
507 already passed the abiotic habitat filters captured by our environmental variables,
508 observed species composition and dominance structure already contain part of the
509 environmental variance, leaving species interactions to explain the residual variance in
510 SAD shapes. In this respect, dryland plant communities worldwide are predominantly
511 shaped by mutualistic, particularly facilitative, interactions (Soliveres and Maestre
512 2014). Interestingly, mutualistic interactions have been largely neglected in the SAD
513 literature, which has focused on competition as the major process shaping dominance
514 structures (McGill et al., 2007). Many competition based models (reviewed in Tokeshi,
515 1998; but see Mouillot et al., 2000) predict lognormal type SADs. As there are no

516 models that include the interplay of competition and mutualistic species interactions, it
517 remains unclear whether and to what degree the observed residual variance in SAD
518 shapes (> 65%) can be explained by both types of interactions.

519 However, a low impact in terms of variance explanation does not mean that an
520 environmental predictor is of low or even no influence. This predictor might severely
521 and selectively constrain species abundance and also filter for possible species
522 combinations. Consequently, such predictors might invoke strong selective pressures on
523 species causing the long-term reshaping of community structure. Unfortunately
524 respective long-term effects of low impact environmental drivers are not well known. In
525 this respect we need data on the temporal change in abundance distributions in habitats
526 of stable environmental conditions. Such data might allow for an assessment of the real
527 impact of environmental drivers on community structure.

528

529 Conclusions

530 Composition and dominance orders of dryland plant communities are influenced by
531 a manifold of possible drivers. Our results do not point to productivity as a driver
532 towards lognormal abundance distributions in drylands. Rather, we identified the small
533 scale variability in soil characteristics to be of major importance for the maintenance of
534 community evenness and the type of SAD. This variability, in combination with arid
535 habitat conditions, is supported by the presence of a proper ‘middle class’ of
536 abundances. Factors increasing this small-scale soil variability might therefore also
537 contribute to the stability of dryland plant communities.

538

539 Acknowledgements

540 We specially thank Manuel Delgado-Baquerizo, Miguel Berdugo, Matthew A. Bowker,
541 Donaldo Bran, Omar Cabrera, José A. Carreira, Alex Cea, Mohamed Chaieb, Abel A.
542 Conceição, Mchich Derak, Carlos I. Espinosa, Adriana Florentino, Juan Gaitán, Wahida
543 Ghiloufi, Susana Gómez-González, Beatriz Gozalo, Julio R. Gutiérrez, Elizabeth
544 Guzmán, Rosa M. Hernández, Elisabeth Huber-Sannwald, Miguel García-Gómez,
545 Mohammad Jankju, Rebecca L. Mau, Maria Miriti, Jorge Moneris, Victoria Ochoa,
546 Ana Prado-Comesaña Vicente Polo, Aníbal Prina, Eduardo Pucheta, José Luis Quero,
547 David A. Ramírez, Roberto Romão, Duilio Torres, Cristian Torres-Díaz, James Val,
548 Enrique Valencia, Deli Wang and Eli Zaady for their contribution to the database used.

549 This research was funded by the European Research Council under the European
550 Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement
551 242658 (BIOCOM). The Ciencia y Tecnología para el Desarrollo (CYTED) program
552 funded networking activities (EPES, Acción 407AC0323). WU was supported by the
553 Polish National Science Centre (grant 2014/13/B/NZ8/04681). FTM acknowledges
554 support from the Salvador de Madariaga program of the Spanish Ministry of Education,
555 Culture and Sports (PRX14/00225), and from a Humboldt Research Award from the
556 Alexander von Humboldt Foundation.

557

558 **References**

- 559 Adler, P.B. et al. .2011. Productivity is a poor predictor of plant species richness.
560 *Science* 333, 1750–1753.
- 561 Alonso, D., Ostlingm A., Etienne R,S, 2008, The implicit assumption of symmetry
562 and the species abundance distribution. *Ecol. Lett.* 11, 93–105.
- 563 Barabás, G., D’Andrea, R., Rael, R., Meszema, G., Ostling, A. 2013. Emergent
564 neutrality of hidden niches? *Oikos* 122, 1565-1572.
- 565 Bazzaz, F.A. 1975. Plant species diversity in old-field successional ecosystems in
566 southern Illinois. *Ecology* 56, 485-488.
- 567 Berkeley, A., Thomas, A.D., Dougill, A.J. 2005. Spatial dynamics of biological soil
568 crusts: bush canopies litter and burial in Kalahari rangelands. *African J. Ecol.* 43,
569 137-145.
- 570 Burnham, K.P., Anderson, D.R. 2002. Model selection and multimodel inference: a
571 practical information-theoretic approach, 2nd edition. Springer, New York.
- 572 Cheng, J., Mi, X., Nadrowski, K., Ren, H., Zhang, J., Ma, K. 2012. Separating the effect of
573 mechanisms shaping species-abundance distributions at multiple scales in a
574 subtropical forest. *Oikos* 121, 236-244.
- 575 Connolly, S.R., Hughes, T.P., Bellwood, D.R., Karlson, R.H. 2005. Community
576 structure of corals and reef fish at multiple scales. *Science* 309, 1363–1365.
- 577 Connolly, S.R., Dornelas, M. 2011. Fitting and empirical evaluation of models for
578 species abundance distributions, in: Magurran, A.E., McGill, B.J. (Eds.)
579 *Biological diversity: frontiers in measurement and assessment*. Oxford University
580 Press Oxford, pp. 123-140.

581 Currie, D.J. 1991. Energy and large-scale patterns of animal- and plant-species richness.
582 Am. Nat. 137, 27–49.

583 Dai, A. 2013. Increasing drought under global warming in observations and models.
584 Nature Climate Change 3.,52-58.

585 Death, R.G. 1996. The effect of habitat stability on benthic invertebrate communities:
586 the utility of species abundance distributions. *Hydrobiol.* 317, 97-107.

587 Delgado-Baquerizo, M. et al. 2013. Decoupling of soil nutrient cycles as a function of
588 aridity in global drylands. *Nature* 502, 672-676.

589 Dornelas, M., Connolly, S.R. 2008. Multiple modes in a coral species abundance
590 distribution. *Ecol. Lett.* 11, 1008-1016.

591 Dornelas, M., Soykan, C.U., Umland, K.I. 2011 Biodiversity and disturbance in:
592 Magurran, A.E., McGill, B.J. (Eds.) *Biological diversity: frontiers in*
593 *measurement and assessment.* Oxford University Press Oxford, pp. 237-251.

594 Fisher, R.A., Corbet, A.S., Williams, C.B. 1943. The relation between the number of
595 species and the number of individuals in a random sample of an animal
596 population. *J. Anim. Ecol.* 12, 42–58.

597 Fraterrigo, J.M., Rusak, J.A. 2008. Disturbance-driven changes in the variability of
598 ecological patterns and processes. *Ecol. Lett.* 11, 756-770.

599 Gaitán, J.J., Oliva, G.E., Bran, D.E., Maestre, F.T., Aguiar, M.R., Jobbagy, E.B.,
600 Buono, G.G., Ferrante, D., Nakamatsu, V.B., Ciari, G., Salomone, J.M., Massara,
601 V. 2014. Vegetation structure is as important as climate to explain ecosystem
602 function across Patagonian rangelands. *J. Ecol.* 102, 1419–1428.

603 Gray, J.S., Waldichuk, M., Newton, A.J., Berry, R.J., Holden, A.V., Pearson, T.H.
604 1979. Pollution-induced changes in populations [and discussion]. *Phil. Trans. R.*
605 *Soc. Lond. B* 286, 545–561.

606 Gray, J.S., Mirza, F.B. 1979. A possible method for the detection of pollution-induced
607 disturbance on marine benthic communities. *Marine. Poll. Bull.* 10, 142-146.

608 Green, J., Plotkin, J.B. 2007. A statistical theory for sampling species abundances. *Ecol.*
609 *Lett.* 10, 1037–1045.

610 Hamer, K.C., Hill, J.K., Lace, L.A., Langman, A.M. 1997. Ecological and
611 biogeographical effects of forest disturbance on tropical butterflies of Sumba
612 Indonesia. *J. Biogeogr.* 24, 67–75.

613 Hawkins, B.A. 2012. Eight (and a half) deadly sins of spatial analysis. *J. Biogeogr.* 39,
614 1-9.

615 Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A. 2005. Very high
616 resolution interpolated climate surfaces for global land areas. *Intern. J. Climat.* 25,
617 1965-1978.

618 Hill, J.K., Hamer, K.C. 1998. Using species abundance models as indicators of habitat
619 disturbance in tropical forests. *J. Appl. Ecol.* 35, 458-460.

620 Hirao, T., Murakami, M., Kubota, Y. 2012. Species abundance distributions of moth
621 and beetle assemblages in a cool-temperate deciduous forest. *Ins. Cons. Div.* 6,
622 494–501.

623 Hubbell, S.P. 1979. Tree dispersion abundance and diversity in a tropical tree forest.
624 *Science* 203, 1299-1309.

625 Hubbell, S.P. 2001. *The unified theory of biogeography and biodiversity.* Princeton
626 University Press, Princeton,

627 Hughes, R.G. 1986. Theories and models of species abundance. *Am. Nat.* 128, 879–
628 899.

629 Körner, C.H. 2000. Biosphere responses to CO₂-enrichment. *Ecol. Appl.* 10, 1590–
630 1619.

631 Lehman, C.L., Tilman, D. 2000. Biodiversity stability and productivity in competitive
632 communities. *Am. Nat.* 156, 534–552.

633 Lan, Z., Bai, Y. 2012. Testing mechanisms of N-enrichment-induced species loss in a
634 semi-arid Inner Mongolian grassland: critical thresholds and implications for
635 long-term ecosystem responses. *Phil. Trans. R. Soc. Lond. B* 367, 3125-3134.

636 Leigh Jr. E.G. 1999. *Tropical Forest Ecology.* Oxford University Press, Oxford.

637 Locey, K.J., White, E.P. 2013. How species richness and abundance constrain the
638 distribution of abundance. *Ecol. Lett.* 16, 1177-1185.

639 Maestre, F.T., Escudero, A. 2009. Is the patch-size distribution of vegetation a suitable
640 indicator of desertification processes? *Ecology* 90, 1729-1735.

641 Maestre, F.T., Salguero-Gómez, R., Quero, J.L. 2012. It's getting hotter in here:
642 determining and projecting the impacts of global change on drylands. *Phil. Trans.*
643 *R. Soc. B* 367, 3062–3075.

644 Maestre, F.T. et al 2012b. Plant species richness and ecosystem multifunctionality in
645 global drylands. *Science* 335, 214-218.

646 Maestre, F.T., Ulrich, W., Soliveres, S., Thomas, A.D., Doughill, A.J. 2015. Data from
647 “Environmental correlates of species rank – abundance distributions in global
648 drylands”. Figshare doi, 106084/m9.figshare1450709

649 MacArthur, R. 1957. On the relative abundance of bird species. Proc. Natl. Acad. Sci.
650 USA 43, 293–295.

651 Magurran, A.E., Henderson, P.A. 2003. Explaining the excess of rare species in natural
652 species abundance distributions. Nature 422, 714–716.

653 Magurran, A.E. 2004. Measuring biological diversity. Blackwell Oxford.

654 Maire, V., Gross, N., Börger, L., Proulx, R., Wirth, C., Pontes, L., Soussana, J.-F.,
655 Louault, F. 2012. Habitat filtering and niche differentiation jointly explain species
656 relative abundance within grassland communities along fertility and disturbance
657 gradients. New Phytol. 196, 497–509.

658 Matthews, T.J., Whittaker, R.J. 2014. Fitting and comparing competing models of the
659 species abundance distribution: assessment and prospect. Frontiers Biogeogr. 6,
660 67–82.

661 Matthews, T.J., Whittaker, R.J. 2015. On the species abundance distribution in applied
662 ecology and biodiversity management. J. Appl. Ecol. 52, 443–454.

663 May, R.M. 1975. Patterns of species abundance and diversity, in: Cody, M.L.,
664 Diamond, J.M. (Eds.) Ecology and evolution of communities. Cambridge
665 University Press, pp. 81–120.

666 McGill, B.J. 2003 Does Mother Nature really prefer rare species or are log-left-skewed
667 SADs a sampling artefact? Ecol. Lett. 6, 766–773.

668 McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benecha, H.K.,
669 Dornelas, M., Enquist, B.J., Green, J.L., He, F., Hurlbert, A.H., Magurran, A.E.,
670 Marquet, P.A., Maurer, B.A., Ostling, A., Soykan, C.U., Ugland, K.I., White, E.P.
671 2007. Species abundance distributions: moving beyond single prediction theories
672 to integration within an ecological framework. Ecol. Lett. 10, 995–1015.

673 Morlon, H., White, E.P., Etienne, R.S., Green, J.L., Ostling, A., Alonso, D., Enquist,
674 B.J., He, F., Hurlbert, A., Magurran, A.E., Maurer, B.A., McGill, B.J., Olf, H.,
675 Storch, D., Zillio, T. 2009. Taking species abundance distributions beyond
676 individuals. Ecol. Lett. 12, 488–501.

- 677 Mouillot, D., Lepretre, A., Andrei-Ruiz, M.-C., Viale, D. 2000. The fractal model: a
678 new model to describe the species accumulation process and relative abundance
679 distribution (RAD). *Oikos* 90, 333-342.
- 680 Nekola, J.C., Šizling, A.L., Boyer, A.G., Storch, D. 2008. Artifacts in the log-
681 transformation of species abundance distributions. *Folia Geobot.* 43, 259–268.
- 682 Nummelin, M. 1998. Log-normal distribution of species abundance is not a universal
683 indicator of rainforest disturbance. *J. Appl. Ecol.* 35, 454-457.
- 684 Preston, F.W. 1948 The commonness and rarity of species. *Ecology* 29. 254–283.
- 685 Pueyo, S. 2006. Self-similarity in species–area relationship and in species–abundance
686 distribution. *Oikos* 112, 156–162.
- 687 Qu, X.-D., Song, M.-Y., Park, Y.-S., Oh, Y.N., Con, T.-S. 2008. Species abundance
688 patterns of benthic macroinvertebrate communities in polluted streams. *Intern. J.*
689 *Limnol.* 44, 11-25.
- 690 Rangel, T.F., Diniz-Filho, J.A.F., Bini, L.M. 2010. SAM: a comprehensive application
691 for spatial analysis in macroecology. *Ecography* 33, 46-50.
- 692 Raunkiær, C. 1909. Livsformen hos Planter paa ny Jord Kongelige Danske
693 Videnskabernes Selskabs Skrifter. Naturvidenskabelig og Matematisk Afdeling
694 7, 1-70.
- 695 Reth, S., Hentschel, K., Drösler, M., Falge, E. 2005. DenNit – Experimental analysis
696 and modelling of soil N₂O efflux in response on changes of soil water content soil
697 temperature soil pH nutrient availability and the time after rain event. *Plant and*
698 *Soil* 272, 349–363.
- 699 Reynolds, J.F., Stafford-Smith, D.M., Lambin, E.F., Turner, B.L., Mortimore, M.,
700 Batterbury, S.P.J., Downing, T.E., Dowlatabadi, H., Fernandez, R.J., Herrick,
701 J.E., Huber-Sannwald, E., Jiang, H., Leemans, R., Lynam, T., Maestre, F.T.,
702 Ayarza, M., Walker, B. 2007. Global desertification: Building a science for
703 dryland development. *Science* 316, 847-851.
- 704 Safriel, U., Adeel, Z. 2005. Dryland systems, in: Hassan, R. et al (Eds.) *Ecosystems and*
705 *human well-being: current state and trends: findings of the condition and trends*
706 *working group.* Island Press Washington DC, pp. 623-662
- 707 Šizling, A.L., Storch, D., Šizlingová, E., Reif, J., Gaston, K.J. 2009. Species abundance
708 distribution results from a spatial analogy of central limit theorem. *Proc. Natl.*
709 *Acad. Sci. USA* 106, 6691–6695.

710 Schreeg, L.A., Kress, W.J., Erickson, D.L., Swenson, N.G. 2010. Phylogenetic analysis
711 of local scale tree soil associations in a lowland moist tropical forest. *Plos One* 5,
712 e13685.

713 Sears, A.L.W., Chesson, P. 2007. New methods for quantifying the spatial storage
714 effect: an illustration with desert annuals. *Ecology* 88, 2240-2247.

715 Soliveres, S., Maestre, F.T. 2014. Plant–plant interactions environmental gradients and
716 plant diversity: A global synthesis of community-level studies. *Perspect. Plant.*
717 *Ecol. Evol. Syst.* 16, 154–163.

718 Sugihara, G. 1980. Minimal community structure: an explanation of species–abundance
719 patterns. *Am. Nat.* 116, 770–787.

720 Tilman, D. 1982. Resource competition and community structure. Princeton Univ Press,
721 Princeton.

722 Tokeshi, M. 1998. Species coexistence. Blackwell Oxford.

723 Trabucco, A., Zomer, R.J. 2009. Global Aridity Index (Global-Aridity) and Global
724 Potential Evapo-Transpiration (Global-PET) Geospatial Database CGIAR
725 Consortium for Spatial Information. Published online available from the CGIAR-
726 CSI GeoPortal at: <http://www.cgiar-csi.org/data/global-aridity-and-pet-database>.

727 Ulrich, W. 2013. RAD 2.0 A Fortran Program for fitting of species – abundance
728 distributions. Published online at www.keib.umk.pl.

729 Ulrich, W., Ollik, M. 2004. Frequent and occasional species and the shape of relative
730 abundance distributions. *Div. Distrib.* 10, 263–269.

731 Ulrich, W., Ollik, M., Ugland, K.I. 2010. A meta-analysis of species – abundance
732 distributions. *Oikos* 119, 1149-1155.

733 Ulrich, W., Kusumoto, B., Shiono, T., Kubota, Y. 2016. Climatic and geographical
734 correlates of global forest tree species abundance distributions and community
735 evenness. *J. Veg. Sci.*, 27, 295-305.

736 Ulrich, W., Piwczynski, M., Zaplata, M.K., Winter, S., Schaaf, W., Fischer, A. 2014.
737 Small-scale spatial variability in phylogenetic community structure during early
738 plant succession depends on soil properties. *Oecologia* 175, 985-995.

739 Vergnon, R., van Nees, R.H., Scheffer, M. 2012. Emergent neutrality leads to
740 multimodal species abundance distributions. *Nature Commun*, 3, #663.

741 Violle, C., Enquist, B.J., McGill, B.J., Jiang, L., Albert, C.H., Hulsfof, C., Jung, V.,
742 Messier, J. 2012. The return of the variance: intraspecific variability in
743 community ecology. *Trends Ecol. Evol.* 27, 1-9.

744 Volkov, I., Banavar, J.R., Hubbell, S.P., Maritan, A. 2003. Neutral theory and relative
745 species abundance in ecology. *Nature* 424, 1035-1037.

746 Volkov, I., Banavar, J.R., He, F., Hubbell, S.P., Maritan, A. 2005. Density dependence
747 explains tree species abundance and diversity in tropical forests. *Nature* 438, 658-
748 661.

749 Walker, B., Kinzig, A., Langridge, J. 1999. Plant attribute diversity resilience and
750 ecosystem function: The nature and significance of dominant and minor species.
751 *Ecosystems* 2, 95–113.

752 Whitford, W.G. 2002. *Ecology of desert systems*. Academic Press California USA.

753 Whittaker, R.H. 1960. *Vegetation of the Siskiyou Mountains Oregon and California*.
754 *Ecol. Monogr.* 30, 279–338.

755 Whittaker, R.H. 1965 Dominance and diversity in land plant communities. *Science* 147,
756 250–260.

757 Whittaker, R.H. 1975. *Communities and Ecosystems* 2nd Ed. MacMillan Publishers
758 New York.

759 Wiens, J.J., Graham, C.H. 2005. Niche conservatism: integrating evolution ecology and
760 conservation biology. *Ann. Rev. Ecol. Syst.* 36, 519–539.

761 Wilson, J.B., Gitay, H., Steel, J.B., King, W. McG. 1998. Relative abundance
762 distributions in plant communities: effects of species richness and of spatial scale.
763 *J. Veg. Sci.* 9, 213-220.

764 Zaplata, M.K., Winter, S., Fischer, A., Kollmann, J., Ulrich, W. 2013. Species-driven
765 phases and increasing structure in early-successional plant communities. *Am. Nat.*
766 181, E17-E27.

767 Zillio, T., Condit, R. 2007. The impact of neutrality niche differentiation and species
768 input on diversity and abundance distributions. *Oikos* 116, 931-940.

769 Zomer, R.J., Trabucco, A., Bossion, D.A., Verchot, L.V. 2008. Climate change
770 mitigation: A spatial analysis of global land suitability for clean development
771 mechanism afforestation and reforestation. *Agricult. Ecosyst. Environm.* 126, 67-
772 80.

773

774

775 **Online resources**

776

777 **Electronic supplementary material**

778 Results from the additional analysis using the extended data set (sites with at least 10
779 species) and the reduced data set (sites with at least 20 species)

780 **Table 1.** Numbers of better fits of the log-series ($\Delta AICc > 10$) and lognormal ($\Delta AICc <$
 781 -10) SAD models for the vegetation types included in the present study. Intermediate
 782 fits refer to $-10 \leq \Delta AICc \leq +10$.

783

Vegetation type	Better fit of		
	log-series	lognormal	intermediate
Grasslands	4	22	8
Scrublands	6	21	14
Woodlands	0	15	0
Savannah	0	0	1
Total	10	58	23

784

785 **Table 2.** Ordinary least squares (OLS) models to identify relationships between
 786 environmental variables and the relative fits of the lognormal model (ΔAICc), SAD
 787 skewness, and evenness The variables included in the best fit models (lowest AICc) are
 788 in bold type Model beta values and r^2 refer to the beta values and the explained variance
 789 of the respective model N = 91
 790

Variable	ΔAICc	Skewness	Evenness
Spatial eigenvector	-0.03	-0.26	0.25
Elevation	0.01	0.03	0.15
Species richness	-0.24	0.09	0.16
Temperature	0.06	-0.20	0.07
Precipitation	0.17	0.23	-0.12
pH	0.07	-0.11	0.07
Available phosphorus	0.20	-0.01	0.01
Organic carbon	-0.21	0.17	-0.39
Nitrate	-0.09	-0.02	0.03
r^2 (OLS total model)	0.18	0.15	0.28
r^2 (OLS selected model)	0.16	0.14	0.25

791

792 **Table 3.** Ordinary least squares (OLS) models to identify relationships between soil and
 793 climatic variability and the relative fits of the lognormal model (ΔAICc), SAD skewness,
 794 and evenness The variables included in the best fit models (lowest AICc) are in bold
 795 type Model parameters and r^2 refer to the beta values and the explained variance of the
 796 respective model N = 91
 797
 798

Variable	ΔAICc	Skewness	Evenness
Spatial eigenvector	-0.06	-0.28	0.28
Elevation	-0.04	0.07	-0.03
Species richness	-0.21	0.03	0.24
Temperature seasonality	-0.22	0.01	0.01
Precipitation seasonality	-0.07	0.14	0.15
CV pH	-0.04	0.06	-0.12
CV available phosphorus	-0.05	-0.10	0.07
CV organic carbon	0.13	0.15	0.17
CV nitrate	-0.08	-0.39	0.45
r^2 (OLS total model)	0.14	0.26	0.38
r^2 (OLS selected model)	0.12	0.23	0.35

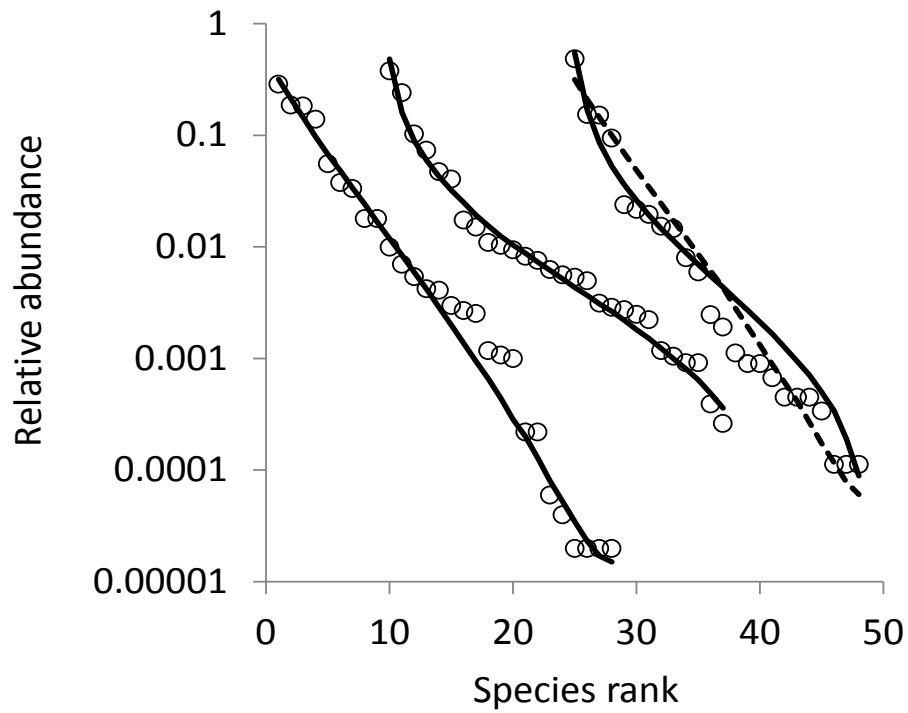
799

800

801

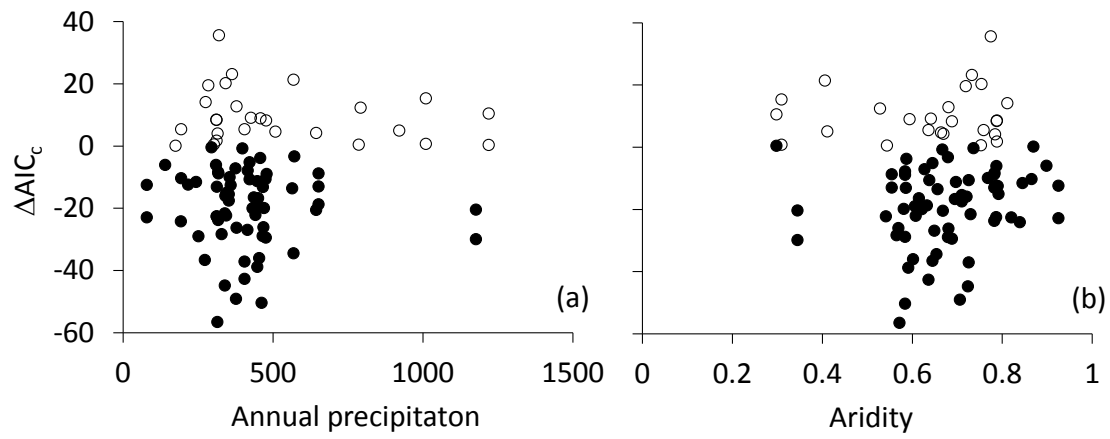
802 **Figure 1.** Three examples of dryland SADs with best fits. From the left: a site from
803 Argentina Pampas and the respective log-series fit, a site from China with the respective
804 lognormal fit, and a site from Spain where both models fit nearly equally well

805



806

807 **Figure 2.** Better fits of the log-series SAD model (open dots) were independent of the
808 degree of precipitation (a) while the lognormal model (black dots) generally fitted better
809 (two exceptions) below 600 mm annual precipitation. Lognormal SADs were found
810 predominately at higher levels of aridity (b).

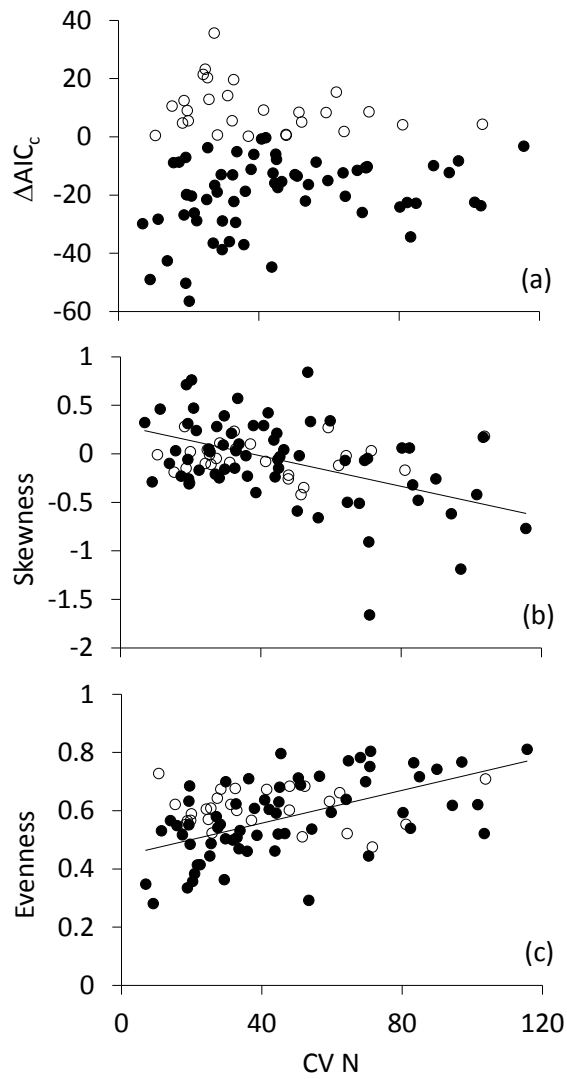


811

812

813 **Figure 3.** Scatter plots of the effect of soil nitrate variability (CV N) on p_{norm} (A), SAD
814 skewness (B), and evenness (C) of the 91 sites having at least 15 species Black and
815 open circles denote sites better fitted by the lognormal SAD and the log-series SAD,
816 respectively. Regression lines for black circles: B: $r^2 = 0.21$, $P < 0.001$, C: $r^2 = 0.25$, $P <$
817 0.001

818



819

820

821