1

Plant demographic knowledge is biased towards short-term studies of

temperate-region herbaceous perennials

4

Authors:

Gesa Römer*[1,2], Johan P. Dahlgren[1,2], Roberto Salguero-Gómez[3], Iain M. Stott[4] and Owen R. Jones*[1,2]

8

Affiliations:

[1] *Interdisciplinary Centre on Population Dynamics (CPop), University of Southern, Denmark, Campusvej 55, 5230 Odense M, Denmark.*

[2] *Department of Biology, University of Southern, Denmark, Campusvej 55, 5230 Odense M, Denmark.*

[3] *Department of Zoology, University of Oxford. 11a Mansfield Road, Oxford, OX1 3SZ, United Kingdom.*

[4] *School of Life Sciences, University of Lincoln, Brayford Pool, Lincoln. LN6 7TS, United Kingdom.*

18

*Corresponding authors:

Gesa Römer gesa.roemer@gmail.com; Owen R. Jones jones@biology.sdu.dk

21

Summary

(1) Plant population dynamics research has a long history, and data collection rates have increased through time. The inclusion of this information in databases enables researchers to investigate the drivers of demographic patterns globally and study life history evolution.

(2) Studies aiming to generalise demographic patterns rely on data being derived from a representative sample of populations. However, the data are likely to be biased, both in terms of the species and ecoregions investigated and in how the original studies were conducted.

(3) Matrix population models (MPMs) are a widely-used tool in plant demography, so an assessment of publications that have used MPMs is a convenient way to assess the distribution of plant demographic knowledge. We assessed bias in this knowledge using data from the COMPADRE Plant Matrix Database, which contains MPMs for almost 800 plant species.

(4) We show that tree species and tropical ecoregions are under-represented, while herbaceous perennials and temperate ecoregions are over-represented. In addition, there is a positive association between the number of studies per country and the wealth of the country. Furthermore, we found a strong tendency towards low spatiotemporal replication: More than 50% of the studies were conducted over fewer than 4 years, and only 17% of the studies have replication across >3 sites. This limited spatiotemporal coverage means that the data may not be representative of the environmental conditions experienced by the species.

(5) Synthesis: The biases and knowledge gaps we identify are a challenge for the progress of theory and limit the usefulness of current data for determining patterns that would

46   be useful for conservation decisions, such as determining general responses to climate

47   change. We urge researchers to close these knowledge gaps with novel data collection.

48   Keywords

49   Biogeography, comparative biology, population growth rate, population projection matrix,

50   transient dynamics

Introduction

Population ecologists aim to understand and predict population dynamics using demographic data that includes the vital rates of survival, reproduction, and development. Their efforts include examining population responses to changes in climate, land use, and management (Silva *et al.*, 1991; Buhler *et al.*, 1997; Eriksson *et al.,* 2002; Morris *et al.*, 2008; Colautti & Barrett, 2013). Demographic data are also crucial for robust population viability analyses of threatened and invasive species (Morris & Doak, 2002; Hansen & Wilson, 2006; Rueda-Cediel *et al.*, 2019). Besides single-species studies, researchers have conducted comparative analyses investigating broad demographic and life history patterns among species. These comparative analyses have aided the development of general theories of life history variation, including r-K selection theory (MacArthur & Wilson, 1967; Gunderson, 1980), Grime's C-S-R triangle (Grime, 1974; Silvertown *et al.* 1992), Stearns' fast-slow continuum (Stearns, 1992; Franco & Silvertown, 1996; Salguero-Gómez *et al.*, 2016) and reproductive strategies continuum (Salguero-Gómez 2017). The empirical exploration of these themes requires large quantities of data from diverse species experiencing a wide range of environmental conditions.

Comparative analyses often rely on the collation of published data to obtain sufficient sample sizes. There are numerous recent examples of this (*e.g.,* Iriondo, 2009; Dalgleish *et al.*, 2010; Bullock *et al.*, 2012; Burns *et al.*, 2013), and large-scale collaborative efforts to collate global demographic and life history and related data are increasingly common (*e.g.,* GBIF: The Global Biodiversity Information Facility; Wright *et al.*, 2004; Loh *et al.*, 2005; NERC Centre for Population Biology, 2010; Kattge *et al.*, 2011; Salguero-Gómez *et al.*, 2015). These databases provide a rich resource for workers focussing on life history strategies and demographic performance.

74   One of the most frequently used tools to describe a species' demography and life history are

75   matrix population models (MPMs, Crone *et al.*, 2011). MPMs depict a population's life cycle

76   in terms of survival, reproduction, and transitions among discrete life stages (Leslie, 1945;

77   Lefkovitch, 1965; Caswell, 2001). MPMs are particularly useful because they have well-

78   understood mathematical properties, and measures derived from MPMs are comparable across

79   diverse species (Silvertown *et al.*, 1993; Caswell, 2001; Salguero-Gómez & Kroon, 2010). The

80   COMPADRE Plant Matrix Database (Salguero-Gómez *et al.*, 2015) is the most comprehensive

81   database of plant studies using MPMs and thus reflects our collective knowledge of plant

82   demography. The contents of databases like COMPADRE were not explicitly collected for

83   inclusion in large databases but rather for the disparate purposes of the many original studies.

84   Although these large databases may be an unbiased (or even complete) sampling of the

85   literature, researchers likely focus on species or geographical areas of particular interest. The

86   resulting data collections are likely to be similarly taxonomically, geographically, or

87   methodologically biased.

88   Bias of this nature has far-reaching consequences for our understanding of plant demography

89   and could limit the usefulness of databases like COMPADRE for comparative analyses. To

90   identify potential biases in plant demographic data and discuss their implications, we used

91   COMPADRE to address the following questions: (1) When have the studies been published?

92   (2) Where is the research done? (3) For which species and populations do we have demographic

93   data? (4) How are the MPMs constructed? More precisely, we tested the following nine

94   hypotheses (H1-H9):

95

96   WHEN HAVE THE STUDIES BEEN PUBLISHED?

97   We expected (**H1**) to see that the proportion of published plant ecology articles that use MPMs

98   has increased through time, reflecting the growing importance of demographic research within

99   plant ecology.

100

101   WHERE IS THE RESEARCH DONE?

102   To assess the potential geographic bias of plant demographic studies, we focussed on

103   continental, ecoregion, and country-level biases. We also examined the relationship between

104   the number of studies and the country's wealth where the study was carried out (as indicated

105   by gross domestic product, GDP). We expected (**H2**) Europe and North America to be over-

106   represented and, likewise, that temperate ecoregions (which characterise these continents)

107   would dominate. Further, we expected (**H3**) wealthier countries would be over-represented

108   compared to their poorer counterparts since they have more funds for research.

109

110   FOR WHICH SPECIES AND POPULATIONS DO WE HAVE DEMOGRAPHIC DATA?

111   We expected (**H4**) the representation of growth forms would not be proportional to their natural

112   abundance, with herbaceous perennials being over-represented and some growth forms being

113   almost absent. This potential bias is important because, if true, it would limit opportunities to

114   make general inferences on the demography of poorly-represented growth forms. Furthermore,

115   we expected to find (**H5**) a tendency to preferentially study threatened species because they are

116   of particular interest in population ecology (Morris & Doak 2002). We also expected (**H6**) a

117   trend towards choosing flourishing populations (*i.e.*, those with $\lambda > 1$) for data collection,

118   reflecting the researchers' desire to ensure the long-term viability of their project.

119

120   HOW ARE THE MPMS CONSTRUCTED?

6

121   The usefulness of individual demographic studies for comparative analyses is sometimes

122   limited by the methods used to construct the MPMs. We explored this by examining within-

123   study spatio-temporal replication and MPM dimension, which can influence demographic

124   quantities calculated from MPMs, including asymptotic population growth rate (Salguero-

125   Gómez & Plotkin, 2010). We expected (**H7**) low rates of temporal and spatial replication,

126   meaning that the data may not represent adequately the environmental conditions experienced

127   by the population/species. We expected (**H8**) that matrix dimension would vary widely, with a

128   tendency for the MPMs of long-lived species such as trees to have a greater dimension.

129   Finally, we analysed the prevalence of a widespread simplification approach used in

130   parameterisation: the assumption of transition constancy in two or more consecutive stages.

131   This can occur, for example, when reproduction or survival in consecutive stages (*e.g.*, small,

132   medium, large plants) is assumed to be constant. Although researchers may justify this

133   simplifying assumption based on data limitations, estimates derived from such simplified

134   MPMs may be inaccurate, limiting their usefulness in comparative work. Studies featuring

135   analyses of life expectancy or generation time (Gaillard *et al.*, 2005; Staerk *et al.*, 2019), ageing

136   trajectories (Baudisch, 2011; Baudisch & Stott, 2019), and transient population dynamics (Stott

137   *et al.*, 2011) are all likely to be marred by the widespread use of this assumption. Despite this

138   problem, we expected (**H9**) a large proportion of studies to parameterise matrices using average

139   transition probabilities and fecundity estimates across stages in this way.

140   We discuss the implications of the biases we identify for several applications. Our results

141   highlight plant demographic knowledge gaps for assessing general patterns, and we encourage

142   researchers to close these taxonomic, biogeographic, and methodological gaps going forward.

143

Materials and methods

To quantify potential biases in our knowledge of plant demography, we used the COMPADRE Plant Matrix Database version 5.0.0 (Salguero-Gómez *et al.*, 2015). Although COMPADRE also contains data on red and brown algae, and lichens, we restrict our analysis to plants (*i.e.*, land plants and green algae) (Cavalier-Smith, 1981). We analysed our data using R version 4.0.4 (R Core Team, 2021).

Data in COMPADRE are organised by research publication such that particular species can appear multiple times in different articles, and a single publication can include several species. We derived our sample from 641 articles on 746 species. Most articles (547) focussed on single species while 94 focussed on multiple species (2-30 species). In some cases, the archived MPM represents the element-by-element average across several transitions (*e.g.*, the average of 5 years of data). However, COMPADRE also often includes data for the individual transitions (*e.g.*, annual transitions are most commonly used, and COMPADRE thus often includes data on the transition from year 1 to 2, and another for year 2 to 3 and so on). Similarly, articles often include several MPMs representing different experimental treatments and/or different spatial areas for a given year or set of years. Our data set included 925 species-by-article combinations and a total of 9,022 MPMs. In addition to the MPMs, we use COMPADRE metadata on geolocation, ecoregion, growth form, taxonomy, and study timeframe, as well as the MPM projection interval, to examine potential temporal, biogeographic and taxonomic biases. We analyse this data in several ways. For most analyses, we use the entire data set but, due to specific requirements for some analyses, we subset the data for some parts of our study. For full transparency, we include the analysis code as supplementary information.

WHEN HAVE THE STUDIES BEEN PUBLISHED?

168   To assess temporal trends in the publication of demography-focussed articles in plant ecology

169   (**H1**), we examined the estimated proportion of articles published in the Journal of Ecology

170   between 1991-2019 that used MPMs. We chose the *Journal of Ecology* as a proxy for the field

171   of plant ecology because it is among the leading and the oldest journals for this field and is thus

172   likely to reflect the temporal development of the discipline. To do this, we downloaded

173   metadata for all of the journal articles for 1991-2019 from the Web of Science. We queried this

174   dataset with the search terms [*projection model* OR *matrix model* OR *MPM*] to identify those

175   that used MPMs. We then compared graphically the estimated percentage of articles that

176   included MPMs.

177   To gauge the completeness of COMPADRE's data holdings, we compared COMPADRE's

178   currently available data (to February 2019) to the COMPADRE team's curated list of articles

179   targeted for eventual inclusion (data provided by Haydee Hernández-Yañez, pers. comm.,

180   2019).

181

182   **WHERE IS THE RESEARCH DONE?**

183   *Biogeography.*

184   We characterised biases in the distribution of studies among continents and ecoregions (**H2**).

185   We first quantified the density of studied species (*i.e.*, number of studied species per unit area

186   (n/km^2) of each country). We then compared the distribution of species among ecoregions in

187   COMPADRE with the estimated actual species distribution in nature among those ecoregions

188   worldwide using Pearson's Chi-squared tests (hereafter, χ^2-tests) and *post-hoc* proportion z-

189   tests (using `prop.test` in R). We could do this because COMPADRE assigns each studied

190   population to one or more of Olson's 14 ecoregions (Olson *et al.*, 2001). In some cases,

191   populations were assigned to multiple closely-related ecoregions (*e.g.*, different types of

9

192   temperate forests). To simplify the analysis, we collapsed Olson's ecoregions into five broader

193   categories: tropical (Olson's ecoregions TMB, TDB, TSC and TGV, see Table S1 for

194   explanation), temperate (TBM, TCF, TGS), Mediterranean/desert (MED, DES), tundra/boreal

195   (BOR, TUN, MON), and wetland (MAN, FGS). We extracted the estimated number of species

196   naturally occurring in each region from Kier (Kier *et al.*, 2005). Kier *et al.* (2005) did not

197   include bryophytes or algae, so we excluded them from this comparison.

198   *Country-specific wealth (GDP).*

199   We used country-level per-capita GDP for 2017 (World Bank, 2018) to examine whether

200   wealthy countries are overrepresented in COMPADRE (**H3**). To do this, we used a Poisson

201   generalised linear model (GLM) (log-link) with log-transformed GDP as the explanatory

202   variable and the number of demographic studies as the response variable. Log-transformation

203   of GDP was necessary to improve the fit of the model. We included only countries with at least

204   one demographic study on plants to avoid a zero-inflated model.

205

FOR WHICH SPECIES AND POPULATIONS DO WE HAVE DEMOGRAPHIC DATA?

208   *Taxonomy.*

209   To characterise potential biases in taxonomy and growth form (**H4**), we first analysed the

210   distribution of taxa in COMPADRE among the taxonomic categories of angiosperm *vs.*

211   gymnosperm, monocot *vs.* eudicot (for the 849 angiosperms only), and Family. We then used

212   the COMPADRE database metadata variable `OrganismType` (hereafter, growth form),

213   which includes a range of paraphyletic growth form categories such as "tree", "herbaceous

214   perennial", and "shrub". We compared the distribution of angiosperms *vs.* gymnosperms in

215   COMPADRE with estimates of their diversity across all plant species from Campbell *et al.*

216   (2018) and the numbers of eudicots and monocots with numbers derived from Evert *et al.*

217   (2013). We linked numbers of species within families extracted from COMPADRE with the

218   number of species per Family listed in The Plant List (2010) for the same families. FitzJohn *et*

219   *al.* (2014) and Willis (2017) estimate how many plant species in the world are woody, which

220   we compared with the number of trees and shrubs in COMPADRE. We compared the number

221   of epiphytes with estimates from (Zotz, 2013). As above, we used χ^2-tests and post-hoc

222   proportion tests for these comparisons.

223   *Conservation status.*

224   To characterise potential bias in the conservation statuses of species studied (**H5**), we obtained

225   the IUCN Red List categories (IUCN 2019) for all species in our data set and in The Plant List,

226   using the R package `rredlist` v. 0.4.0 (Chamberlain, 2017). We compared the distribution

227   of COMPADRE species in each Red List category with the corresponding distribution of all

228   species in The Plant List using a χ^2-test.

229   *Population growth rates.*

230   To assess whether researchers tend to collect demographic data on growing or declining

231   populations and whether researchers tend to study populations that are in a "boom phase" (**H6**),

232   we examined the asymptotic population growth rates (λ) calculated from each MPM. We first

233   filtered the data to include only studies that spanned at least five years, were experimentally

234   unmanipulated, and for which λ could be calculated (*i.e.*, the MPMs contained no missing

235   values and did not violate ergodicity and irreducibility assumptions (Stott *et al.*, 2010)). We

236   then fitted an ordinary least squares regression with λ as the response variable and year as the

237   explanatory variable. The slope of this model indicates the temporal trend in λ: a negative trend

238   supports our hypothesis that researchers preferentially work on initially flourishing sites where

239   population growth rates then decline over time.

240

241 **HOW ARE THE MPMS CONSTRUCTED?**

242 *Temporal and spatial replication.*

243 To explore potential biases in temporal and spatial replication (**H7**), we examined the

244 frequency distributions of study length (years) and the number of spatially distinct populations

245 (as defined by the original article authors) for all species-by-article combinations in

246 COMPADRE.

247 *Matrix model dimension.*

248 To examine how the MPM dimension chosen by modellers varies systematically among growth

249 form and ecoregion (**H8**), we compared the frequency distribution of the MPM dimension

250 across these variables. As above, we tested for an association statistically using Poisson GLMs.

251 *Averaging matrix model elements.*

252 In COMPADRE, MPMs (**A**) are defined as the sum of three submatrices, $\mathbf{A} = \mathbf{U} + \mathbf{F} + \mathbf{C}$,

253 where **U** represents survival-dependent transitions (*e.g.*, growth, stasis, shrinkage, ageing), **F**

254 describes sexual reproduction, and **C** represents clonal reproduction. We used these three

255 submatrices to address **H9** by characterising the prevalence of averaging across stage-/age-

256 classes. This approach was only possible for matrices without missing values. We assessed the

257 number of consecutive life cycle stages in each MPM that contained averaged rates of survival,

258 reproduction and clonality, using the following approaches:

259 To estimate the degree of survival averaging, we calculated stage-specific survival probability

260 as the column sums of the **U** matrix. When the survival probabilities across stages were all

261 different, we categorised the MPM as not containing averaged rates ("no averaging"). When

262 several survival probabilities were consecutively identical across up to 50% of stages, we

263 assumed they had been averaged over those stages ("≤50% averaging"). Finally, when more

12

264   than half of the stages have consecutively identical survival probabilities, we assumed that they

265   had been averaged ("> 50% averaging").

266   To estimate the degree of averaging for sexual reproduction, we calculated stage-specific

267   sexual reproduction as the column sums of the **F** matrix. We classified the degree of averaging

268   into three categories in the same way as for survival. Whereas every survival column must sum

269   to >0, reproduction columns may include zero values. These may be before the stage of first

270   reproduction (pre-reproductive) or after (post-reproductive). Post-reproductive stages with

271     $\mathbf{F} = 0$ are more likely to exist because researchers did not observe reproduction during

272   fieldwork, rather than due to the method of averaging over multiple life-stages. Our

273   assessments include only 'reproducing' stages (i.e., where reproduction >0). We estimated the

274   degree of clonality averaging in the same way, using the **C** matrix.

13

Results

WHEN HAVE THE STUDIES BEEN PUBLISHED?

The proportion of articles focussing on plant demography (**H1**) in the *Journal of Ecology* has increased slightly in the past decades. However, the year-to-year variation is high, and the slope is not significantly different from zero (linear model: slope = 0.034 ± 0.044, $F_{1,21}$ = 0.601, P = 0.447; Fig. 1).

WHERE IS THE RESEARCH DONE?

Biogeography.

As expected (**H2**), geographical bias was obvious, with study density being greatest in Europe (23.67 studies per million km2) and North America (17.56 studies per million km^2), while Oceania, South American, Asian, and African countries were relatively poorly-represented, with 6.69, 5.10, 1.66, 1.02 studies per million km^2, respectively (see also Fig. 2B; Fig.S2). Interestingly, a post-hoc analysis showed that the dominance of Europe and North America in the COMPADRE database has increased since the year 2000: 75% of articles post-2000 focussed on these regions compared to 63% pre-2000 (Table S2).

As hypothesised (**H2**), there was a significant difference between the species distribution in COMPADRE compared to in nature (χ^2-test: χ^2 = 327.79, d.f. = 4, P <0.001). Indeed, species from temperate ecoregions are significantly over-represented in COMPADRE (48% compared to an estimated 24% of species inhabiting these regions in nature; proportion test: χ^2 = 267.71, d.f. = 1, P <0.001; Fig. 2A). In contrast, tropical ecoregions are significantly under-represented (27% vs. 42%; proportion test: χ^2 = 77.432, d.f. = 1, P <0.001). Similar results were apparent for wetlands (1% vs. 5%; proportion test: χ^2 = 23.426, d.f. = 1, P <0.001) and tundra and boreal

14

298   ecoregions (6% vs. 13%; proportion test: $\chi^2 = 40.335$, d.f. = 1, P <0.001). Species from

299   Mediterranean and desert ecoregions are represented in approximately the same proportion as

300   in nature (18% in COMPADRE *vs*. an estimated 17%: proportion test: $\chi^2 = 1.273$, d.f. = 1,

301     $P = 0.259$).

302   *Country-specific wealth (GDP).*

303   As expected (**H3**), the numbers of articles per country and per capita GDP are positively

304   correlated (Poisson GLM: Null Deviance = 1494.7, Residual Deviance = 1144.3, d.f. = 1, 54,

305   P <0.001, Fig. 2C). Here, 73% of countries (161 out of 222) are not represented in

306   COMPADRE. The United States of America, with 202 research articles, dominates

307   COMPADRE, followed by Mexico with 88 articles, Sweden (33 articles), Australia and

308   Canada (both 26 articles), Spain (25 articles), Japan (22 articles), Czech Republic and the

309   United Kingdom (both 21 articles), and Brazil (19 articles). Interestingly, several of the

310   wealthiest countries are not represented in COMPADRE (*e.g.*, China, Iceland, and Ireland).

311   The positive correlation between the number of studies in COMPADRE and country GDP (Fig.

312   2C) remains statistically significant even when we remove the two outliers with most studies

313   (USA and Mexico).

314

315   **FOR WHICH SPECIES AND POPULATIONS DO WE HAVE DEMOGRAPHIC**

316   **DATA?**

317   *Taxonomy.*

318   As expected (**H4**), the representation of species in COMPADRE does not well-reflect natural

319   diversity. COMPADRE categorises species as angiosperm (91%), gymnosperm (6%), or "non-

320   seed plants" (3%) which includes ferns, mosses etc. (Fig. 2D). According to Campbell *et al.*

321   (2018), the corresponding figures in nature are 88% (angiosperm), 1% (gymnosperm) and 11%

15

322   (non-seed plants). Thus, COMPADRE's taxonomic representativity is different than that found

323   in nature (χ^2 test: $\chi^2 = 1282.7$, d.f. = 2, $P < 0.001$). In fact, COMPADRE over-represents

324   gymnosperms (proportion test: $\chi^2 = 1127.1$, d.f. = 1, $P < 0.001$) and under-represents the non-

325   seed plants (proportion test: $\chi^2 = 64.39$, d.f. = 1, $P < 0.001$).

326   Of COMPADRE's angiosperms, 74% are eudicots and 26% monocots (Fig 2E), which

327   approximately reflects Evert *et al.*'s (2013) estimate of their natural diversity distribution (69%

328   eudicot, 31% monocot). Our χ^2 test nevertheless indicated the COMPADRE distribution was

329   significantly different to the natural distribution (χ^2 test: $\chi^2 = 9.79$, d.f. = 1, $P = 0.002$).

330   To better understand the distribution of COMPADRE species across plant families, we

331   examined the five largest eudicot families (according to The Plant List, 2010): Compositae

332   (Asteraceae), Leguminosae (Fabaceae), Orchidaceae, Rosaceae, and Rubiaceae (Fig. 2F). The

333   χ^2 test showed that the distributions differed between COMPADRE and in nature patterns

334   (χ^2test: $\chi^2 = 24.161$, d.f. = 4, $P < 0.001$). However, this difference is driven by the Rubiaceae

335   and Orchidaceae, which are significantly under-represented (Rubiaceae: proportion test:

336    $\chi^2 = 13.616$, d.f. = 1, $P < 0.001$; Orchidaceae: proportion test: $\chi^2 = 7.0719$, d.f. = 1, $P = 0.008$).

337   The other families are fairly proportionately represented in COMPADRE (all $P > 0.05$).

338   Half of the species in COMPADRE are herbaceous perennials (462 of 932 species), and only

339   26% are woody plants. This figure contrasts with current estimates that 45-48% of the world's

340   vascular plant species are woody (FitzJohn *et al.*, 2014; Willis, 2017).

341   *Conservation status.*

342   The IUCN has assessed only 29% (n = 220) of the species in COMPADRE. Of these, contrary

343   to our expectations (**H5**), COMPADRE's content with respect to the Red List status reflects

344   current Red List assessments (χ^2 test: $\chi^2 = 4.054$, d.f. = 4, $P = 0.399$). Most of COMPADRE's

345   species are assessed as Least Concern (62%), with the rest falling into one of the threatened

346   categories (Vulnerable, Endangered, or Critically Endangered) (Fig. 3A).

347   *Population growth rates.*

348   As predicted (**H6**), there is a slight but statistically significant tendency to study growing

349   populations (*i.e.*, $\lambda > 1$), the effect size is small and driven by the skewed nature of the data

350   (Fig. 3B; two-sided t-test on $\lambda = 1$: $t = 10.941$, d.f. $= 2312$, $P = 0.001$). The overall mean value

351   for $\log \lambda$ was 0.013 (standard deviation $= 0.45$). Our analysis of ordinary least-squares

352   regression slopes from the subset of populations with at least a 5-year time-series shows no

353   tendency for researchers to select populations where λ is initially high but then decreases,

354   leading to negative slope values (Fig. 3C; t-test: $t = -0.020$, d.f. $= 192$, $P = 0.984$).

355

356   **HOW ARE THE MPMS CONSTRUCTED?**

357   *Temporal and spatial replication.*

358   As expected (**H7**), most studies in COMPADRE are short-term (Fig. 4A). The modal study

359   duration is three years, while the median is four years. The mean is slightly longer (5.48 years),

360   reflecting the skewed nature of the distribution. Some long-term exceptions include studies of

361   the shrub *Cassia nemophila* (Silander, 1983) and the tree *Acer saccharum* (Lin & Augspurger,

362   2008), which both span 51-years. The study duration varies by ecoregion, with tropical and

363   marine studies tending to be shorter than those from other ecoregions (Fig. S3A). There was

364   minimal variation in study duration among growth forms and, contrary to expectation, trees (a

365   typically long-lived growth form) are not studied for longer periods than typically shorter-lived

366   growth forms (Fig. S4A).

367   As expected, most studies have a low degree of spatial replication (Fig 4B). Most studies are

368   carried out at a single site, though the distribution is skewed (mean $= 2.82$, median $= 2$,

369   range = 1-60). Interestingly, species studied at four or more sites are mainly herbaceous

370   perennials (57%).

371   Each MPM is obtained from a different study site or year. Therefore, the number of MPMs

372   present in a study may be an indicator of the range of environmental conditions captured by

373   the study. However, the number of MPMs per study does not vary much among ecoregions

374   (Fig S3A and S3B) or growth form (Fig S4A and B).

375   *Matrix model dimension.*

376   We expected (**H8**) that the MPM dimension would vary widely and would be greater for long-

377   lived groups like trees. MPM dimension ranges between two and 60 (for *Rhododendron*

378   *ponticum*, (Travis *et al.*, 2011) but is left-skewed (mean = 6.51, median = 5; Fig. 4C).

379   Furthermore, typical matrix dimension varied significantly among ecoregions (ANOVA on log

380   matrix dimension: $F = 17.967$, d.f. = 5 and 884, $P <0.001$) (Fig. S3C), with tropical species

381   tending to have slightly larger matrices than temperate or Mediterranean ones (t-test on log

382   matrix dimension: $t = 6.855$, d.f. = 375.36, $P <0.001$; though this effect is likely driven by large

383   tropical tree matrices). Finally, MPM dimension varies systematically across growth forms

384   (ANOVA on log matrix dimension: $F = 22.718$, d.f. = 6 and 730, $P <0.001$) with tree MPMs

385   tending to be larger (mean = 9.48, median = 8, range = 2-60) than other growth forms (Fig.

386   S4C), thus supporting our initial hypothesis (**H8**).

387   *Averaging matrix model elements.*

388   Contrary to our expectation (**H9**), researchers do not often appear to average survival rates

389   across life cycle stages when parameterising MPMs. In COMPADRE, 77% of MPMs have no

390   averaging of survival across consecutive stages at all, and only 4% have more than half of their

391   stage-specific survival rates averaged across consecutive stages (Fig. 4D). A similar pattern

392   follows for fecundity: 79% have no averaging and a smaller proportion (5.1%) have over half

18

393   of their stage/age-specific fecundity rates averaged (Fig. 4E). Fewer than 1% of the MPMs

394   showed averaging in clonality rates (Fig. 4F).

395

Discussion

397   Before analysing large heterogeneous databases, it is essential to understand their potential

398   biases and inconsistencies. Under-representation of particular ecoregions or taxonomic groups

399   may lead to incorrect generalisations if those under-represented regions or groups have distinct

400   demographic behaviour. Improved knowledge of systematic biases in databases like

401   COMPADRE, and a recognition of their impact on inferences, will improve our understanding

402   of the natural world. Researchers should carefully consider potential systematic biases in these

403   large-scale datasets, especially when conducting comparative studies that seek to generalise

404   across disparate taxa and geographic regions.

405   COMPADRE v.5.0.0. contains data for 746 species, representing 0.002% of the ~370,000

406   extant plant species (not including green algae) (The Plant List, 2010). Although COMPADRE

407   covers only a fraction of plant diversity, we show that it contains the majority of published

408   MPM-based plant demographic work (Fig. 1B). Thus, this database is a valid indicator of our

409   knowledge of MPM-based plant demography and reveals demographic knowledge gaps for

410   most species. We note, however, that the vast majority of these studies are published in the

411   English language literature. Amano *et al.* (2016) found that a third of the literature in

412   biodiversity conservation was non-English, and that half provide neither the title nor the

413   abstract in English. About 16% of this corpus is unsearchable using English keywords, thus

414   remaining hidden. Assuming that there is a similar pattern for plant demography literature,

415   some knowledge gaps in the English-speaking research community could undoubtedly be

416   closed by engaging researchers familiar with non-English language literature to assist with

417   contributions to COMPADRE.

418

WHEN HAVE THE STUDIES BEEN PUBLISHED?

420   The number of articles that focus on plant demography has steadily accumulated since the

421   1970s (Fig. 1B). Our hypothesis that an increasing proportion of plant ecology research would

422   be demography-focussed (**H1**) was supported, based on our survey of articles published in the

423   *Journal of Ecology* (Fig 1A), although the relationship was rather noisy. The downturn in the

424   rate of increase in MPM-related publication in recent years (since about 2015) may be due to

425   the increasingly important role that integral projection models (IPMs, Easterling *et al.*, 2000)

426   play in plant demographic research.

427

WHERE IS THE RESEARCH DONE?

429   *Biogeography.*

430   Plant biodiversity is unevenly distributed. Equatorial regions are usually relatively species-

431   rich, with declining biodiversity towards the poles for most plant groups (Gaston, 2000).

432   Biodiversity and endemism hotspots are concentrated in the tropics, on equatorial islands, and

433   in the southern hemisphere (Myers *et al.*, 2000; Kier *et al.*, 2009). Our finding that ~73% of

434   demographic studies are carried out in the mainly temperate regions of the western northern

435   hemisphere (Fig. 2A & B) contrasts with those hotspots, thus supporting **H2**. This result is not

436   surprising given similar findings for population dynamics (Amano & Sutherland, 2013; McRae

437   *et al.*, 2017), biodiversity time-series (Dornelas *et al.*, 2018), and the distribution of ecological

438   study sites (Martin *et al.*, 2012). Collectively, these patterns highlight important knowledge

439   gaps for some of the planet's most threatened ecosystems. For example, sub-Saharan Africa

440 and Southeast Asia have among the least ecological data yet show the most rapid decline of

441 terrestrial ecosystems (MEA, 2005). Given that demographic data is important for assessing

442 extinction risk, *e.g.*, by assessing population trends and population viability analyses (IUCN

443 2019; Rodrigues, *et al.*, 2006), the lack of demographic data is a concerning impediment to

444 species-level conservation. Beyond conservation, this biogeographic bias limits our

445 understanding of global ecological trends and drivers of population dynamics and life history

446 evolution.

447 *Country-specific wealth (GDP).*

448 As anticipated (**H3**), the number of demographic articles was positively associated with the

449 per-capita GDP of the country where the work was carried out (Fig. 2C). We expected this

450 because high-GDP countries can invest more in research (van Noorden & Butler, 2019; World

451 Bank, 2019), and researchers tend to conduct research near their home institution for logistical

452 reasons (Coutts *et al.*, 2016). The fact that the relationship is relatively loose reflects the

453 international networks and mobility of some researchers who carry out research away from

454 their home institution. Nevertheless, one way to correct this bias is for funding bodies to

455 encourage plant demographic research, and collaboration with researchers, in understudied and

456 threatened regions.

457

458 **FOR WHICH SPECIES AND POPULATIONS DO WE HAVE DEMOGRAPHIC**

459 **DATA?**

460 *Taxonomy.*

461 As expected (**H4**), most of COMPADRE's species are angiosperms, reflecting the high

462 diversity of this group. However, COMPADRE over-represents gymnosperms and under-

463 represents the non-seed plants. Within the angiosperms, most of COMPADRE's species are

21

464   eudicots. Although this reflects the natural distribution of eudicot *vs.* monocot species, we did

465   detect a statistically significant bias towards the study of eudicots. COMPADRE's species

466   distribution among the major eudicot plant Families approximated the natural distribution,

467   except for Rubiaceae and Orchidaceae, which were under-represented. The pattern for

468   Rubiaceae may be explained by geographical bias since it mainly occurs in the (sub)tropics,

469   which are not well-represented in COMPADRE. As expected (**H4**), we know more about the

470   demography of herbaceous perennials than any other growth form. Trees, which are important

471   both economically (Poore, 2013) and ecologically (Chambers *et al.*, 2001), are under-

472   represented, probably due to the logistical difficulties of studying large, long-lived organisms,

473   but perhaps also due to the relatively low tree species diversity in temperate regions.

474   Although some of these biases may be overcome statistically (*e.g.*, by resampling or

475   rarefaction), the scarcity of demographic data on several growth forms, including ferns, lianas,

476   and bryophytes, drastically reduces our ability to draw general patterns for these growth forms

477   and set them in context with more commonly-studied forms. This issue is particularly

478   troublesome for comparative studies of the evolution of plant life history.

479   *Conservation status.*

480   Demographic models are an indispensable tool to guide management decisions for threatened

481   species (Norris, 2004). Contrary to our expectation (**H5**) that researchers may collectively

482   focus on threatened species, the distribution of demographic studies in COMPADRE well-

483   approximates the distribution of Red List statuses of plants in general: There is no tendency to

484   favour studies of threatened species. However, we should regard this result with caution

485   because it is based on the subset of ~25% of COMPADRE species that have been assessed for

486   the IUCN Red List. The true distribution of species among IUCN Red List categories may be

487   quite different, especially since species endemic to biodiversity hotspots are less likely to have

488   been assessed.

489   *Population growth rates.*

490   There was a slight tendency for researchers to preferentially study growing populations

491   (supporting **H6**) but there was no evidence for a "regression to the mean" effect whereby

492   population growth rates decline along the time-series (contrary to Buckley *et al.*, 2010). We

493   initially expected this tendency because we expected researchers to select obviously flourishing

494   populations to avoid the risk and logistical cost of local extinction. The differences between

495   our results and those of Buckley *et al.* (2010) could be due to differences in data or methods.

496   This effect warrants further investigation because biased sampling (*e.g.*, towards growing

497   populations) could lead to incorrect conclusions about population dynamics in comparative

498   research.

499

500   **HOW ARE THE MPMS CONSTRUCTED?**

501   *Temporal and spatial replication.*

502   The fact that demographic studies tend to have low spatial and temporal replication supports

503   our original hypothesis (**H7**) and confirms previous findings (Crone *et al.*, 2011; Ehrlén *et al.*,

504   2016). Limited spatial replication may affect confidence in inferences made from those models.

505   The geographic distribution of plants varies widely, with some only occurring in specific small

506   areas (*e.g.*, *Iliamna remota* is endemic to the ~8 hectare Langham Island, Illinois, USA;

507   Swinehart & Jacobs, 1998) and others even spanning continents (*e.g.*, *Plantago major*, Sagar

508   & Harper, 1964). Widely distributed species are likely to experience a greater range of

509   environmental conditions than those with small ranges, and demographic data should ideally

510   be collected in representative parts of this range to understand the species' demography more

511   fully. Work by Doak & Morris (2010), Wardle *et al.* (2014), and Römer *et al.* (2021) are good

512   examples of such efforts.

513   The limited temporal extent in most studies is also a concern. Researchers have argued that

514   accurate forecasting of population dynamics typically requires time-series extending well

515   beyond three years, especially because of the demographic impacts of rare extreme weather

516   events (Doak & Morris, 2010; Ehrlén *et al.*, 2016; Teller *et al.*, 2016; Pérez-Llorca, *et al.*,

517   2018). Given the cost and effort required for long-term research, it is not surprising that the

518   temporal extent of studies in COMPADRE is short. In some settings, researchers could use an

519   alternative space-for-time substitution approach to resolve this problem. The approach enables

520   a rapid accumulation of data representing a large range of environmental conditions allowing

521   the modelling of responses to future climate scenarios without the need for long time-series

522   (Blois *et al.*, 2013; Teller *et al.*, 2016; Damgaard, 2019; Römer *et al.*, 2021). However, the

523   approach assumes that drivers of demographic variation across space are equivalent to those

524   that drive temporal variation, which may not be the case (Pickett, 1989). In any case, the low

525   spatial replication in COMPADRE may currently limit the application of this approach.

526   *Matrix model dimension.*

527   Researchers constructing MPMs decide an appropriate dimension for their model, based on

528   factors including species life-history (including longevity or life cycle complexity), the study's

529   purpose, and the amount of data available to parameterise each stage. As expected (**H8**), matrix

530   dimension varies hugely, with a substantial proportion (~20%) having a low dimension of 3 or

531   less. This low dimension could limit utility in some cases. For example, this is likely to be too

532   low for the derivation of measurements relying on the calculation of age trajectories from stage-

533   based MPMs (Cochran & Ellner, 1992; Caswell, 2001) such as Keyfitz's entropy (Keyfitz,

534   1968). Furthermore, other derived metrics, including elasticities (Salguero-Gómez & Plotkin,

535   2010) and some transient measures (Stott *et al.*, 2010), are sensitive to the MPM dimension.

536   Besides influencing individual metrics, the systematic bias in model dimension among growth

537   forms and ecoregion could lead to spurious inferences in multi-species comparative studies if

538   not taken into account.

539   *Averaging matrix model elements.*

540   MPM-derived metrics of population dynamics such as transient dynamics metrics (Stott *et al.*,

541   2010) and measures of life-history (*e.g.*, survival inequality or entropy) are sensitive to

542   homogeneity among the vital rates of stages because peaks and troughs of survival or mortality

543   in certain life stages might be undetected. Despite this, researchers sometimes have no option

544   but to parameterise MPMs with average vital rates across consecutive stages. We expected this

545   would be common (**H9**), but our results show that plant ecology researchers seldom take this

546   approach. This finding is good news because averaging of reproduction or survival could lead

547   to an underestimation of the effects of temporal variation in the environment underlying the

548   vital rates (Stott *et al.*, 2010). This underestimation would be a challenge for applications that

549   are not based on asymptotic properties of the MPM, such as calculations of extinction risk, the

550   stochastic growth rate in population viability analyses, and short-term predictions of population

551   fate (transient dynamics; Stott *et al.*, 2011). Thus, inference from MPMs that have been

552   parameterised in this way could lead to misguided management strategies that miss

553   opportunities to influence the population dynamics in the desired way by manipulating vital

554   rates with strong influences on short-term or stochastic population growth. This averaging of

555   vital rates would also problematic for comparative life-history research, and in particular, work

556   that relies on age-from-stage methods (Cochran & Ellner, 1992; Caswell, 2001) to calculate

557   demographic trajectories and derived measures. We, therefore, encourage researchers to,

558   wherever possible, avoid parameterisation using averaging over consecutive stages.

559

560   **Conclusions**

561   Our current knowledge of global plant demography is based on geographically biased data

562   heavily focused on herbaceous perennials, leaving important knowledge gaps. Demographic

563   studies are constrained by funding, with the temporal length reflecting the typical grant and

564   PhD tenure, with most work concentrated in wealthy countries. We did not find significant bias

565   in conservation status or population growth rate, which indicates that researchers do not focus

566   on species of conservation concern nor growing or shrinking populations. To close the

567   aforementioned knowledge gaps and better understand generalities in life-history strategy and

568   population dynamics, research targeting neglected growth forms and ecoregions is desirable,

569   as is increased spatial and temporal replication within species. Furthermore, an improved

570   understanding of the impact of these biases on model predictions and methodological

571   developments to account for known biases would be helpful.

572

Acknowledgements

581

Author's contributions

582

583   GR, RS-G and ORJ conceived the ideas and designed methodology; ORJ and RS-G provided

584   additional data; GR, ORJ and IS analysed the data; GR led the writing of the manuscript. All

585   authors contributed to manuscript writing and gave final approval for publication.

Data availability

586

587   Most of the used data is open access and can be downloaded on the following webpages: The

588   COMPADRE Plant Matrix Database: http://www.compadre-db.org (Salguero-Gómez *et al.*,

589   2015); IUCN Red List data: http://www.iucnredlist.org (IUCN, 2019); The Plant List:

590   http://www.theplantlist.org (The Plant List, 2010); GDP data was obtained from the World

591   Bank: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD (World Bank, 2018). Code for

592   the analyses is included in the supplementary information.

593

594 ## Figure captions

595 Figure 1: Publication trajectories in plant demography. **(A)** The percentage of articles in a

596 sample of plant ecology literature that use matrix population models (MPMs). Note that the

597 time-series starts in 1993 before which abstracts have been not digitised. **(B)** The cumulative

598 number of MPM-based studies by year, archived in the COMPADRE Plant Matrix Database

599 (at February 2019) (blue line) compared to the estimated cumulative number of all published

600 studies containing plant MPMs (black line).

601

602 Figure 2. Geographic and taxonomic biases in the COMPADRE Plant Matrix Database. **(A)**

603 The species distribution among ecoregions in COMPADRE compared to the natural

604 distribution (Trop = tropical; temp = temperate; Med/Des = Mediterranean and deserts;

605 Tund/Bor = tundra and boreal regions; Wetl = Wetlands). **(B)** The distribution of plant

606 demography study density across continents. **(C)** The relationship between country per-capita

607 GDP and the number of plant demography studies. The regression line represents a gamma-

608 error GLM, conditioned on countries having at least one plant demography study. **(D)**

609 Comparison of the species distribution among broad categories of angiosperms, gymnosperms

610 and non-seed plants, in COMPADRE and in nature. **(E)** Comparison of the distribution of

611 angiosperm species among monocot and eudicot categories, in COMPADRE and in nature. **(F)**

612 Comparison of the distribution of species among the five largest dicot families, in

613 COMPADRE and in nature (Comp = Compositae; Legum = Leguminosae;

614 Orchid = Orchidaceae; Rosa = Rosaceae; Rubia = Rubiaceae). **(G)** The distribution of species

615 among growth form categories (Herb = herbaceous perennials; Tree = trees; Shrub = shrubs;

616 Suc = succulents; Palm = palms; An = annuals; Epi = epiphytes; Other = includes mosses and

617 ferns).

618

619   Figure 3. Conservation status and population trend biases in the COMPADRE Plant Matrix

620   Database. **(A)** Comparison of the distribution of species among IUCN Red List conservation

621   status in COMPADRE (blue) and in nature (black) (LC = Least Concern, NT = Near

622   Threatened, VU = Vulnerable, EN = Endangered, CR = Critically Endangered). **(B)** The

623   distribution of population growth rates (λ) for MPMs in COMPADRE. The graph is limited to

624   λ-values between 0 and 4 to show the interesting area around $\lambda = 1$ (and $\lambda > 4$ seems

625   biologically unreasonable and may represent errors). **(C)** The density distribution of the slope

626   of the linear $\lambda \sim$ year relationship for studies with >5 years of data.

627

628   Figure 4: Spatiotemporal replication and MPM construction. **(A)** The distribution of study

629   duration in COMPADRE. **(B)** The distribution of the number of study sites (spatial replication)

630   in COMPADRE. **(C)** The distribution of matrix dimension across MPMs in COMPADRE. The

631   magenta and yellow lines show the mean and median, respectively. **(D, E, F)** Summaries of

632   element averaging in MPM submatrices of **(D)** survival, **(E)** fecundity and **(F)** clonality.

633   *none* = all stages have different survival / fertility / clonality estimates; *≤50%* = apparent

634   averaging with two or more consecutive values are the same, but the number of stages apparent

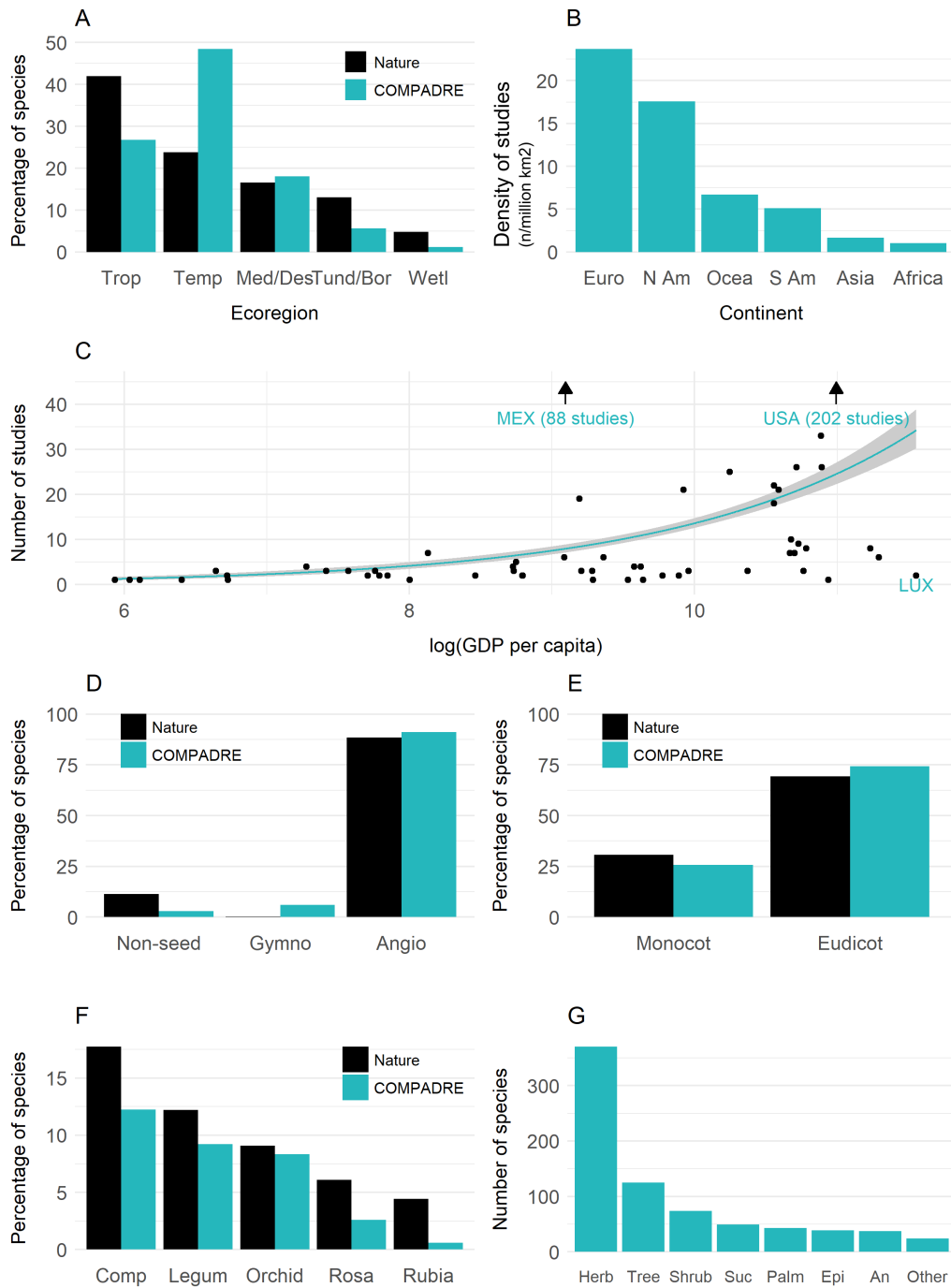635   averaging does not exceed 50%; *>50%* = more than half of the stages have the same value.
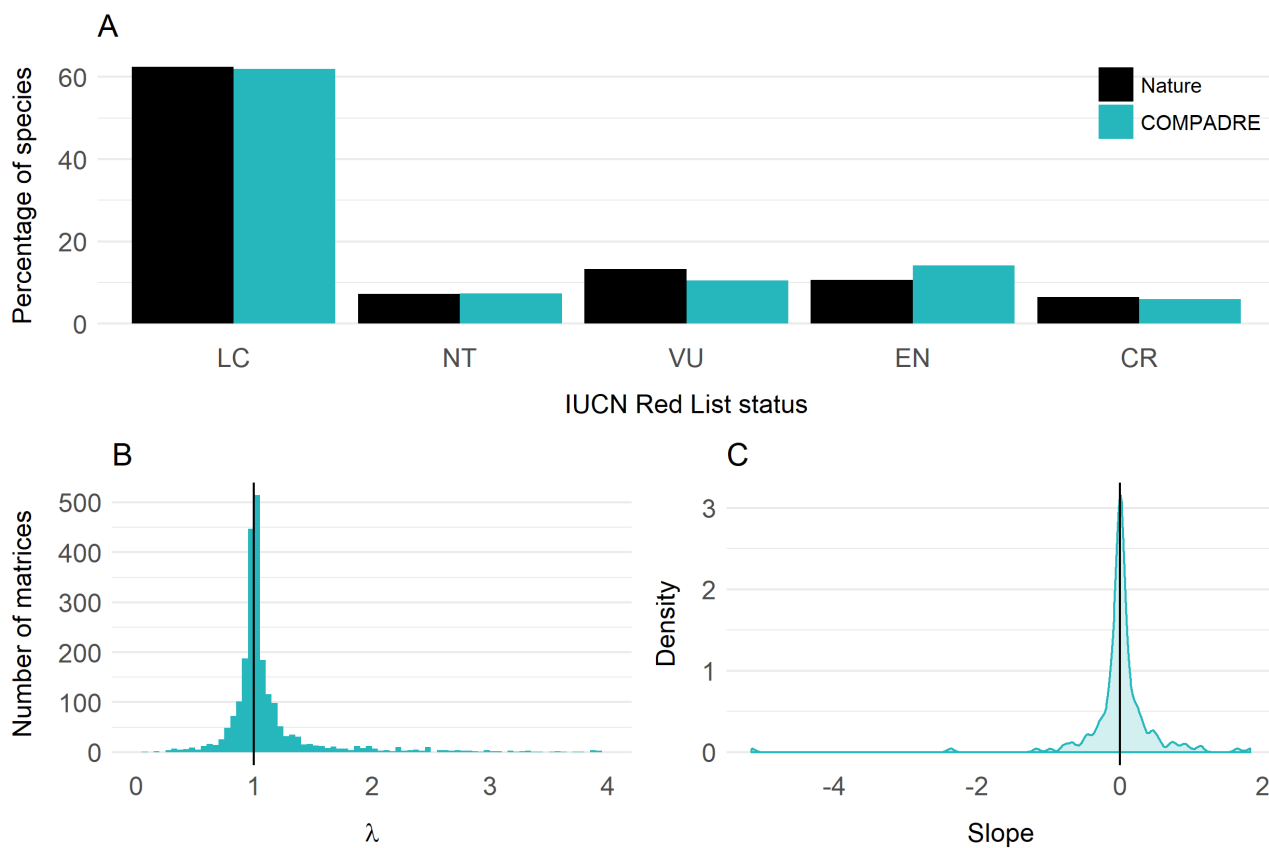
636

Figures and Tables
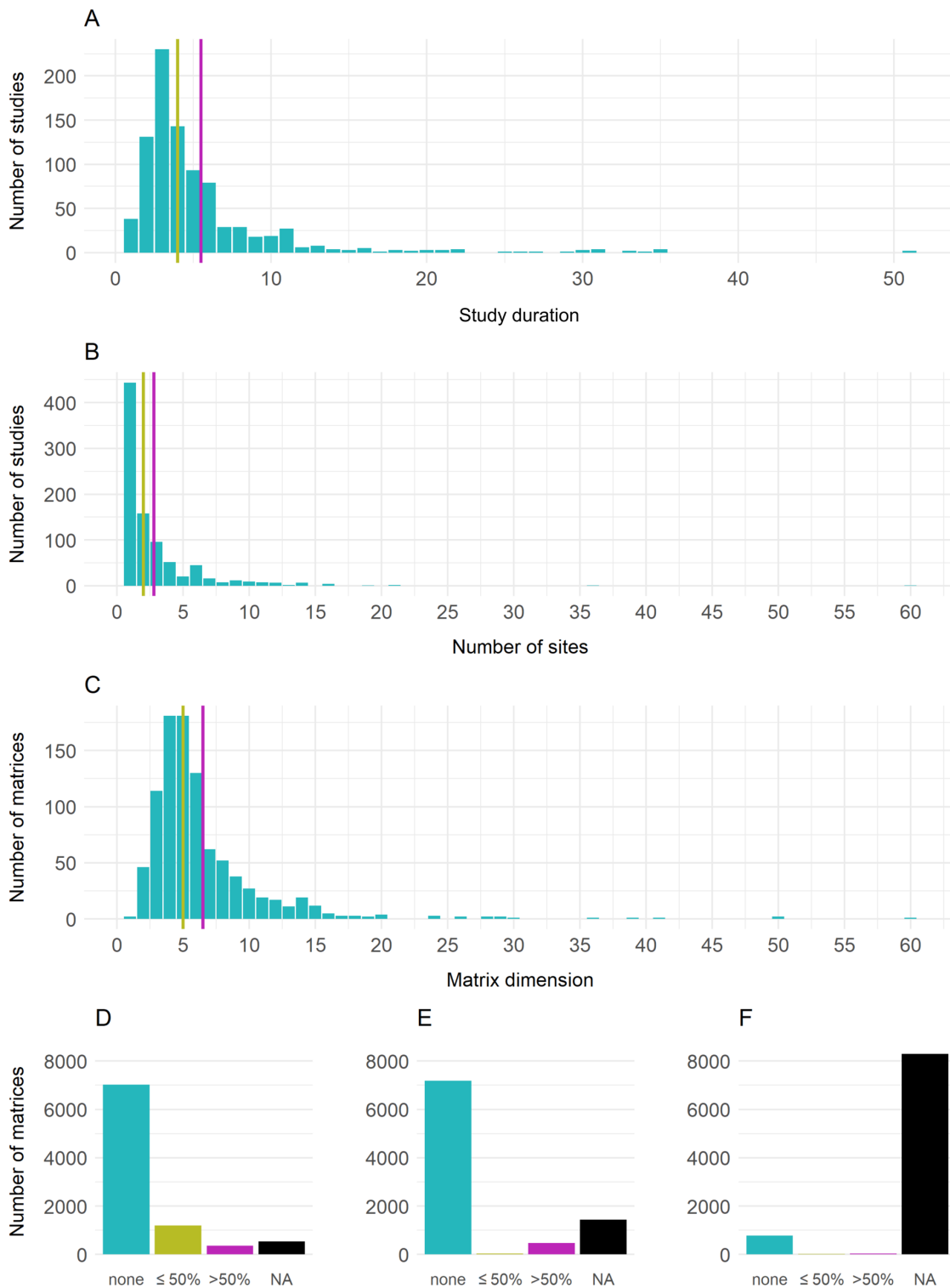
A

B
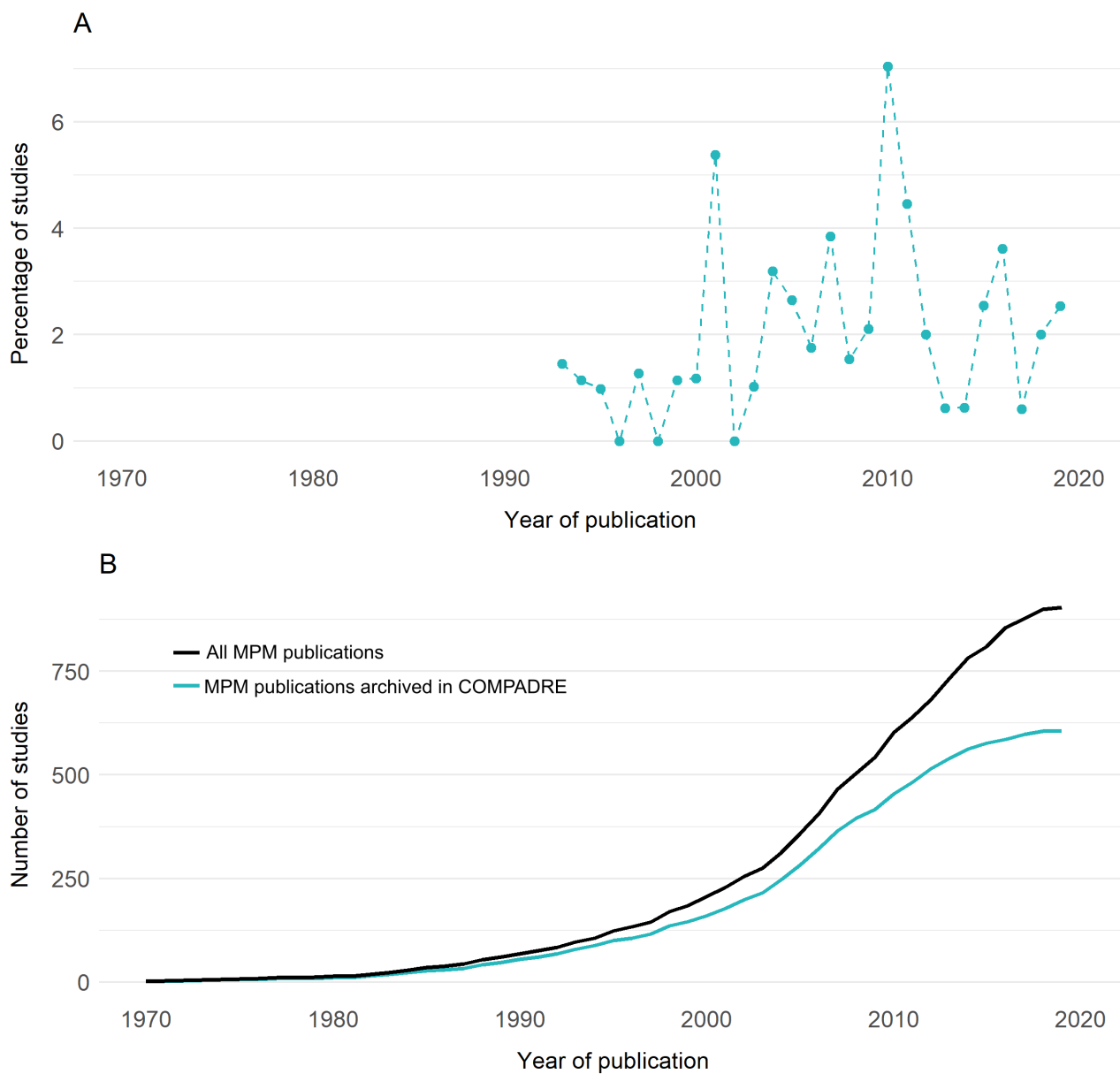
Figure 1.

Figure 2.

645

646

647   Figure 3.

648

649

650

651   Figure 4.

652   **References**

653   Amano, T., González-Varo, J. P., & Sutherland, W. J. (2016). Languages are still a major

654   barrier to global science. *PLoS Biology*, 14(12), e2000933.

655   Amano, T., & Sutherland, W. J. (2013). Four barriers to the global understanding of

656   biodiversity conservation: Wealth, language, geographical location and security.

657   *Proceedings of the Royal Society Series B. Biological Sciences*, 280(1756), 20122649.

658   Baudisch, A. (2011). The pace and shape of ageing. *Methods in Ecology and Evolution*, 2(4),

659   375–382.

660   Baudisch, A., & Stott, I. (2019). A pace and shape perspective on fertility. Methods in

661   Ecology and *Evolution*, 10(11), 1941–1951.

662   Blois, J. L., Williams, J. W., Fitzpatrick, M. C., Jackson, S. T., & Ferrier, S. (2013). Space

663   can substitute for time in predicting climate-change effects on biodiversity. *Proceedings of*

664   *the National Academy of Sciences*, 110(23), 9374–9379.

665   Buckley, Y. M., Ramula, S., Blomberg, S. P., Burns, J. H., Crone, E. E., Ehrlén, J., . . .

666   Wardle, G. M. (2010). Causes and consequences of variation in plant population growth

667   rate: A synthesis of matrix population models in a phylogenetic context. *Ecology Letters*,

668   13(9), 1182–1197.

669   Buhler, D. D., Hartzler, R. G., & Forcella, F. (1997). Implications of weed seedbank

670   dynamics to weed management. *Weed Science*, 45(3), 329–336.

671   Bullock, J. M., White, S. M., Prudhomme, C., Tansey, C., Perea, R., & Hooftman, D. A. P.

672   (2012). Modelling spread of British wind-dispersed plants under future wind speeds in a

673   changing climate. *Journal of Ecology*, 100(1), 104–115.

674 Burns, J. H. (2008). Demographic performance predicts invasiveness of species in the

675   Commelinaceae under high-nutrient conditions. *Ecological Applications*, 18(2), 335–346.

676 Burns, J. H., Pardini, E. A., Schutzenhofer, M. R., Chung, Y. A., Seidler, K. J., & Knight, T.

677   M. (2013). Greater sexual reproduction contributes to differences in demography of

678   invasive plants and their noninvasive relatives. *Ecology*, 94(5), 995–1004.

679 Campbell, N. A., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Reece, J. B.

680   (2018). *Biology: A global approach* (Eleventh edition, global edition). Boston: Pearson.

681 Caswell, H. (2001). Matrix *population models: Construction, analysis, and interpretation.*

682   (2nd ed.). Sunderland, Mass.: Sinauer Associates.

683 Cavalier-Smith T (1981). Eukaryote kingdoms: seven or nine? *Bio Systems*. 14 (3–4): 461–

684   81.

685 Chamberlain, S. (2017). rredlist: 'IUCN' Red List Client: R package version 0.4.0. Retrieved

686   from https://CRAN.R-project.org/package=rredlist

687 Chambers, J. Q., Higuchi, N., Tribuzy, E. S., & Trumbore, S. E. (2001). Carbon sink for a

688   century. *Nature*, 410(6827), 429.

689 Cochran, M. E., & Ellner, S. (1992). Simple methods for calculating age-based life history

690   parameters for stage-structured populations. *Ecological Monographs*, 62, 345–364.

691 Colautti, R. I., & Barrett, S. C. H. (2013). *Rapid adaptation to climate facilitates range*

692   *expansion of an invasive plant*. Science (New York, N.Y.), 342(6156), 364–366.

693 Conlisk, E., Castanha, C., Germino, M. J., Veblen, T. T., Smith, J. M., & Kueppers, L. M.

694   (2017). Declines in low-elevation subalpine tree populations outpace growth in high-

695   elevation populations with warming. *Journal of Ecology*, 105(5), 1347–1357.

696 Coutts, S. R., Salguero-Gómez, R., Csergő, A. M., & Buckley, Y. M. (2016). Extrapolating

697 demography with climate, proximity and phylogeny: Approach with caution. *Ecology*

698 *Letters*, 19(12), 1429–1438.

699 Crone, E. E., Menges, E. S., Ellis, M. M., Bell, T., Bierzychudek, P., Ehrlén, J., . . . Williams,

700 J. L. (2011). How do plant ecologists use matrix population models? *Ecology Letters*,

701 14(1), 1–8.

702 Dalgleish, H. J., Koons, D. N., & Adler, P. B. (2010). Can life-history traits predict the

703 response of forb populations to changes in climate variability? *Journal of Ecology*, 98(1),

704 209–217.

705 Damgaard, C. (2019). A critique of the space-for-time substitution practice in community

706 ecology. *Trends in Ecology & Evolution*, 34(5), 416–421.

707 Doak, D. F., & Morris, W. F. (2010). Demographic compensation and tipping points in

708 climate-induced range shifts. *Nature*, 467(7318), 959–962.

709 Dornelas, M., Antão, L. H., Moyes, F., Bates, A. E., Magurran, A. E., Adam, D., . . . Hickler,

710 T. (2018). Biotime: A database of biodiversity time series for the Anthropocene. *Global*

711 *Ecology and Biogeography*, 27(7), 760–786.

712 Easterling, M. R., Ellner, S. P., & Dixon, P. M. (2000). Size-specific sensitivity: Applying a

713 new structured population model. *Ecology*, 81(3), 694–708.

714 Ehrlén, J., Morris, W. F., Euler, T. von, & Dahlgren, J. P. (2016). Advancing environmentally

715 explicit structured population models of plants. *Journal of Ecology*, 104(2), 292–305.

716 Eriksson, O., Cousins, S. A.O., & Bruun, H. H. (2002). Land-use history and fragmentation

717 of traditionally managed grasslands in Scandinavia. *Journal of Vegetation Science*, 13(5),

718 743–748.

719   Evert, R. F., Raven, P. H. B. o. p., & Eichhorn, S. E. (2013). *Raven biology of plants* (Eighth

720   edition). New York: W.H. Freeman and Company Publishers.

721   FitzJohn, R. G., Pennell, M. W., Zanne, A. E., Stevens, P. F., Tank, D. C., & Cornwell, W. K.

722   (2014). How much of the world is woody? *Journal of Ecology*, 102(5), 1266–1272.

723   Fowler, N. L., Overath, R. D., & Pease, C. M. (2006). Detection of density dependence

724   requires density manipulations and calculation of lambda. *Ecology*, 87(3), 655–664.

725   Franco, M., & Silvertown, J. (1996). Life history variation in plants: an exploration of the

726   fast-slow continuum hypothesis. *Philosophical Transactions of the Royal Society of*

727   *London. Series B, Biological Sciences*, 351(1345), 1341–1348.

728   Gaillard, J.-M., Yoccoz, N. G., Lebreton, J.-D., Bonenfant, C., Devillard, S., Loison, A., . . .

729   Allaine, D. (2005). Generation time: A reliable metric to measure life-history variation

730   among mammalian populations. *American Naturalist*, 166(1), 119-23; discussion 124-8.

731   Gaston, K. J. (2000). Global patterns in biodiversity. *Nature*, 405(6783), 220–227.

732   GBIF: The Global Biodiversity Information Facility. What is GBIF?. Retrieved from

733   https://www.gbif.org/what-is-gbif

734   Grime, J. P. (1974). Vegetation classification by reference to strategies. *Nature*, 250(5461),

735   26–31.

736   Gunderson, D. R. (1980). Using r-K selection theory to predict natural mortality. *Canadian*

737   *Journal of Fisheries and Aquatic Sciences*, 37(12), 2266–2271.

738   Hansen, M. J., & Wilson, S. D. (2006). Is management of an invasive grass *Agropyron*

739   *cristatum* contingent on environmental variation? *Journal of Applied Ecology*, 43(2), 269–

740   280.

741 Albert, M.J., Giménez Benavides, L., Domínguez Lozano, F. & Escudero, A. (Ed.) (2009).

742 *Poblaciones en peligro: Viabilidad demográfica de la flora vascular amenazada de*

743 *España: Populations in peril: Demographic viability of threatened Spanish vascular flora.*

744 Madrid: Dirección General de Medio Natural y Política Forestal.

745 IUCN 2019. The IUCN Red List of Threatened Species. Version 2019-2. Retrieved from

746 http://www.iucnredlist.org

747 Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., . . . Wirth, C.

748 (2011). TRY - a global database of plant traits. *Global Change Biology*, 17(9), 2905–2935.

749 Keyfitz, N. (1968). *Introduction to the mathematics of population*. Addison-Wesley series in

750 behavioral science: quantitative methods. Reading, Mass.: Addison-Wesley Pub. Co.

751 Kier, G., Kreft, H., Lee, T. M., Jetz, W., Ibisch, P. L., Nowicki, C., . . . Barthlott, W. (2009).

752 A global assessment of endemism and species richness across island and mainland

753 regions. *Proceedings of the National Academy of Sciences*, 106(23), 9322–9327.

754 Kier, G., Mutke, J., Dinerstein, E., Ricketts, T. H., Küper, W., Kreft, H., & Barthlott, W.

755 (2005). Global patterns of plant diversity and floristic knowledge. *Journal of*

756 *Biogeography*, 32(7), 1107–1116.

757 Lefkovitch, L. P. (1965). The study of population growth in organisms grouped by stages.

758 *Biometrics*, 21(1), 1.

759 Leslie, P. H. (1945). On the Use of Matrices in Certain Population Mathematics. Biometrika,

760 33(3), 183.

761 Lin, Y., & Augspurger, C. K. (2008). Long-term spatial dynamics of *Acer saccharum* during

762 a population explosion in an old-growth remnant forest in Illinois. *Forest Ecology and*

763 *Management*, 256(5), 922–928.

764   Loh, J., Green, R. E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V., & Randers, J.

765   (2005). The Living Planet Index: Using species population time series to track trends in

766   biodiversity. *Philosophical Transactions of the Royal Society of London. Series B,*

767   *Biological Sciences*, 360(1454), 289–295.

768   MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton

769   landmarks in biology. Princeton, N.J., Oxford: Princeton University Press.

770   Martin, L. J., Blossey, B., & Ellis, E. (2012). Mapping where ecologists work: biases in the

771   global distribution of terrestrial ecological observations. *Frontiers in Ecology and the*

772   *Environment*, 10(4), 195–201.

773   McRae, L., Deinet, S., & Freeman, R. (2017). The diversity-weighted living planet index:

774   controlling for taxonomic bias in a global biodiversity indicator. *PloS One*, 12(1),

775   e0169156.

776   MEA (2005). *Ecosystems and human well-being: Biodiversity synthesis*. Washington DC:

777   World Resources Institute.

778   Morris, W. F., & Doak, D. F. (2002). *Quantitative conservation biology: Theory and practice*

779   *of population viability analysis*. New York: W. H. Freeman; Basingstoke : Palgrave

780   Macmillan.

781   Morris, W. F., Pfister, C. A., Tuljapurkar, S., Haridas, C. V., Boggs, C. L., Boyce, M. S., . . .

782   Menges, E. S. (2008). Longevity can buffer plant and animal populations against changing

783   climatic variability. *Ecology*, 89(1), 19–25.

784   Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A., & Kent, J. (2000).

785   Biodiversity hotspots for conservation priorities. *Nature*, 403(6772), 853–858.

786   NERC Centre for Population Biology (2010). The Global Population Dynamics Database

787   Version 2. Retrieved from http://www.sw.ic.ac.uk/cpb/cpb/gpdd.html.

788   Norris, K. E. N. (2004). Managing threatened species: the ecological toolbox, evolutionary

789   theory and declining-population paradigm. *Journal of Applied Ecology*, 41(3), 413–426.

790   Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N.,

791   Underwood, E. C., . . . Kassem, K. R. (2001). Terrestrial ecoregions of the world: a new

792   map of life on earth. *BioScience*, 51(11), 933.

793   Pérez-Llorca, M., Fenollosa, E., Salguero-Gómez, R., & Munné-Bosch, S. (2018). What is

794   the minimal optimal sample size for plant ecophysiological studies? *Plant Physiology*,

795   178(3), 953–955.

796   Pickett, S. T. A. (1989). Space-for-Time Substitution as an Alternative to Long-Term

797   Studies: In: Gene E. Likens (Ed.), *Long-Term Studies in Ecology: Approaches and*

798   *Alternatives*. New York NY: Springer New York.

799   Poore, D. (2013). No Timber Without Trees: Routledge.

800   Poorter, L., Zuidema, P. A., Peña-Claros, M., & Boot, R. G. A. (2005). A monocarpic tree

801   species in a polycarpic world: how can *Tachigali vasquezii* maintain itself so successfully

802   in a tropical rain forest community? *Journal of Ecology*, 93(2), 268–278.

803   R Core Team (2021). R: A language and environment for statistical computing. R

804   Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

805   Ramage, M. H., Burridge, H., Busse-Wicher, M., Fereday, G., Reynolds, T., Shah, D. U., . . .

806   Scherman, O. (2017). The wood from the trees: The use of timber in construction.

807   *Renewable and Sustainable Energy Reviews*, 68, 333–359.

808   Raunkiær, C. (1907). *Planterigets Livsformer og deres Betydning for Geografien.: Med 77*

809   *figurer i teksten*. København and Kristiania: Gyldendalske Boghandel - Nordisk Forlag.

810   Raunkiær, C. (1934). *The Life Forms of Plants and Statistical Plant Geography, being the*

811   *collected papers of C. Raunkiær*.: Translated by H. Gilbert-Carter, A. Fausbøll, and A. G.

812   Tansley. Oxford University Press, Oxford. Reprinted 1978 (ed. by Frank N. Egerton).

813   History of ecology. New York: Arno Press.

814   Rodrigues, A. S. L., Pilgrim, J. D., Lamoreux, J. F., Hoffmann, M., & Brooks, T. M. (2006).

815   The value of the IUCN Red List for conservation. *Trends in Ecology & Evolution*, 21(2),

816   71–76.

817   Römer, G., Christiansen, D. M., de Buhr, H., Hylander, K., Jones, O. R., Merinero, S.,

818   Reitzel, K., Ehrlén, J., and Dahlgren, J. P. (2021). Drivers of large-scale spatial

819   demographic variation in a perennial plant. *Ecosphere* 12(1). e03356.

820   Rueda-Cediel, P., Brain, R., Galic, N., & Forbes, V. (2019). Comparative analysis of plant

821   demographic traits across species of different conservation concern: Implications for

822   pesticide risk assessment. *Environmental Toxicology and Chemistry*, 38(9), 2043–2052.

823   Sagar, G. R., & Harper, J. L. (1964). *Plantago major* L., *P. media* L. and *P. lanceolata* L.

824   *Journal of Ecology*, 52(1), 189.

825   Salguero-Gómez, R. (2017). Applications of the fast-slow continuum & reproductive strategy

826   framework of plant life histories. *New Phytologist,* 4, 1618-1624

827   Salguero-Gómez, R., & Plotkin, J. B. (2010). Matrix dimensions bias demographic

828   inferences: Implications for comparative plant demography. *American Naturalist*, 176(6),

829   710–722.

830   Salguero-Gómez, R., Jones, O. R., Archer, C. R., Buckley, Y. M., Che-Castaldo, J., Caswell,

831   H., . . . Vaupel, J. W. (2015). The COMPADRE Plant Matrix Database: an open online

832   repository for plant demography. *Journal of Ecology*, 103(1), 202–218.

41

833   Salguero-Gómez, R., Jones, O. R., Jongejans, E., Blomberg, S. P., Hodgson, D. J., Mbeau-

834   Ache, C., . . . Buckley, Y. M. (2016). Fast-slow continuum and reproductive strategies

835   structure plant life-history variation worldwide. *Proceedings of the National Academy of*

836   *Sciences*, 113(1), 230–235.

837   Salguero-Gómez, R., & Kroon, H. de (2010). Matrix projection models meet variation in the

838   real world. Journal of Ecology, 98(2), 250–254.

839   Silander, J. A. (1983). Demographic variation in the Australian desert cassia under grazing

840   pressure. *Oecologia*, 60(2), 227–233.

841   Silva, J. F., Raventos, J., Caswell, H., & Trevisan, M. C. (1991). Population responses to fire

842   in a tropical savanna grass, *Andropogon semiberbis*: A matrix model approach. *Journal of*

843   *Ecology*, 79, 345–356.

844   Silvertown, J., Franco, M., & McConway, K. (1992). A demographic interpretation of

845   Grime's triangle. *Functional Ecology*, 6(2), 130.

846   Silvertown, J., Franco, M., Pisanty, I., & Mendoza, A. (1993). Comparative plant

847   demography--relative importance of life-cycle components to the finite rate of increase in

848   woody and herbaceous perennials. *Journal of Ecology*, 81(3), 465.

849   Staerk, J., Conde, D. A., Ronget, V., Lemaitre, J.-F., Gaillard, J.-M., & Colchero, F. (2019).

850   Performance of generation time approximations for extinction risk assessments. *Journal of*

851   *Applied Ecology*, 56(6), 1436–1446.

852   Stearns, S. C. (1992). *The evolution of life histories*. Oxford University Press, Oxford. UK.

853   Stott, I., Franco, M., Carslake, D., Townley, S., & Hodgson, D. J. (2010). Boom or bust? A

854   comparative analysis of transient population dynamics in plants. *Journal of Ecology*,

855   98(2), 302–311.

856   Stott, I., Townley, S., Carslake, D., & Hodgson, D. J. (2010). On reducibility and ergodicity

857   of population projection matrix models. *Methods in Ecology and Evolution*, 1: 242-252.

858   Stott, I., Townley, S., & Hodgson, D. J. (2011). A framework for studying transient dynamics

859   of population projection matrix models. *Ecology Letters*, 14, 959–970.

860   Swinehart, A. L., & Jacobs, M. E. (1998). Rediscovery, status and preservation of the

861   endangered Kankakee Globe Mallow (*Iliamna remota*) in Indiana. *Rhodora*, 100, 82–87.

862   Teller, B. J., Adler, P. B., Edwards, C. B., Hooker, G., & Ellner, S. P. (2016). Linking

863   demography with drivers: climate and competition. *Methods in Ecology and Evolution*,

864   7(2), 171–183.

865   The Plant List (2010). The Plant List: A working list of all plant species. Version 1.1.

866   Retrieved from http://www.theplantlist.org

867   Travis, J. M. J., Harris, C. M., Park, K. J., & Bullock, J. M. (2011). Improving prediction and

868   management of range expansions by combining analytical and individual-based modelling

869   approaches. *Methods in Ecology and Evolution*, 2(5), 477–488.

870   van Noorden, R., & Butler, D. (2019). Science in Europe: By the numbers. *Nature*,

871   569(7757), 470–471.

872   Verwer, C., Peña-Claros, M., van der Staak, D., Ohlson-Kiehn, K., & Sterck, F. J. (2008).

873   Silviculture enhances the recovery of overexploited mahogany *Swietenia macrophylla.*

874   *Journal of Applied Ecology*, 45(6), 1770–1779.

875   Wardle, G. M., Buckley, Y. M., & PlantPopNet steering group (2014). *PlantPopNet: A*

876   *Spatially Distributed Model System for Population Ecology* (Ecology Society of Australia

877   2014 Annual Conference). Alice Springs NT, Australia.

878   Willis, K. J. (Ed.) (2017). *State of the world's plants, 2017*. [Richmond, Surrey, UK]: Royal

879   Botanic Gardens Kew.

880   World Bank (2018, October 18). World Development Indicators: GDP per capita (current

881   US$). Retrieved from

882   http://databank.worldbank.org/data/reports.aspx?source=2&type=metadata&series=NY.G

883   DP.PCAP.CD#

884   World Bank (2019). Research and development expenditure (% of GDP). Retrieved from

885   https://data.worldbank.org/indicator/gb.xpd.rsdv.gd.zs?view=map

886   Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., . . . Villar,

887   R. (2004). The worldwide leaf economics spectrum. *Nature*, 428(6985), 821–827.

888   Zotz, G. (2013). The systematic distribution of vascular epiphytes - a critical update.

889   *Botanical Journal of the Linnean Society*, 171(3), 453–481.

890   Zuidema, P. A., & Boot, R. G. A. (2002). Demography of the Brazil nut tree (*Bertholletia*

891   *excelsa*) in the Bolivian Amazon: impact of seed extraction on recruitment and population

892   dynamics. *Journal of Tropical Ecology*, 18(1), 1–31.

893   Zuidema, P. A., Brienen, R. J. W., During, H. J., & Güneralp, B. (2009). Do persistently fast-

894   growing juveniles contribute disproportionately to population growth? A new analysis tool

895   for matrix models and its application to rainforest trees. *American Naturalist*, 174(5), 709–

896   719.

897   Zuidema, P. A., Kroon, H. de, & Werger, M. J. A. (2007). Testing sustainability by

898   prospective and retrospective demographic analyses: evaluation for palm leaf harvest.

899   *Ecological Applications*, 17(1), 118–128

900   Zuidema, P. A. (2000). Demography of exploited tree species in the Bolivian Amazon.

901   *PROMAB scientific series*: Vol. 2. Utrecht: PROMAB.

902

903　## Supporting information

904　Appendix S1: Supplementary figures and tables

905　Appendix S2: Analysis code

906