

1 **This is the peer reviewed version of the following article:**
2

3 ***Di Marco, M., Ferrier, S., Harwood, T.D., Hoskins, A.J., Watson, J.E.M. Wilderness areas***
4 ***halve the extinction risk of terrestrial biodiversity. Nature 573, 582–585 (2019).***
5

6 **which has been published in final form at <https://doi.org/10.1038/s41586-019-1567-7>.**

7 **This article may be used for the purposes of academic research, subject always to the**
8 **full conditions of use. Any further use is subject to permission from Springer Nature.**
9

10
11
12 **Wilderness areas halve the extinction risk of terrestrial biodiversity**

13 Moreno Di Marco^{1,2}, Simon Ferrier³, Tom D Harwood³, Andrew J Hoskins⁴, James EM
14 Watson^{5,6}

15
16 ¹CSIRO Land and Water, EcoSciences Precinct, 41 Boggo Road, Dutton Park Qld 4102,
17 Australia

18 ²Dept. of Biology and Biotechnology, Sapienza University of Rome, viale dell'Università 32,
19 00185 Rome, Italy

20 ³CSIRO Land and Water, Black Mountain Laboratories, Clunies Ross St., Canberra, ACT
21 2601, Australia

22 ⁴CSIRO Health and Biosecurity, Australian Tropical Science & Innovation Precinct, James
23 Cook University, James Cook Drive, Townsville QLD 4810, Australia

24 ⁵Centre for Biodiversity and Conservation Science, The University of Queensland, 4072
25 Brisbane, Queensland, Australia.

26 ⁶Global Conservation Program, Wildlife Conservation Society, 2300 Southern Boulevard,
27 Bronx, New York 10460, USA

28 **Reducing the unsustainable rate of global biodiversity loss is one of the major challenges**
29 **that humanity faces¹, as the consequences of biological annihilation would be irreversible**
30 **for humankind²⁻⁴. While the ongoing erosion of ecosystems^{5,6} and the species that**
31 **comprise them^{7,8} is now well documented, little is known about the role Earth's remaining**
32 **wilderness areas play in mitigating the global biodiversity crisis. Here we show that**
33 **retaining this remaining wilderness is essential for the international conservation agenda,**
34 **using an innovative approach to modelling biodiversity persistence from habitat**
35 **condition and spatial variation in species composition. Wilderness areas act as a buffer**
36 **against species loss, with extinction risk being less than half, on average, for species within**
37 **wilderness communities compared to those in non-wilderness communities. While all**
38 **wilderness areas have an intrinsic conservation value^{9,10}, we identify areas on every**
39 **continent that make the highest relative contribution to the persistence of biodiversity.**
40 **Alarminglly, these highly important areas - where habitat loss would have more dramatic**
41 **biodiversity impact - are poorly protected. Given the high rates of global wilderness loss¹⁰**
42 **these areas urgently require targeted protection to ensure the long-term persistence of**
43 **biodiversity, alongside efforts aimed at protecting and restoring more degraded**
44 **environments.**

45

46 **Main text**

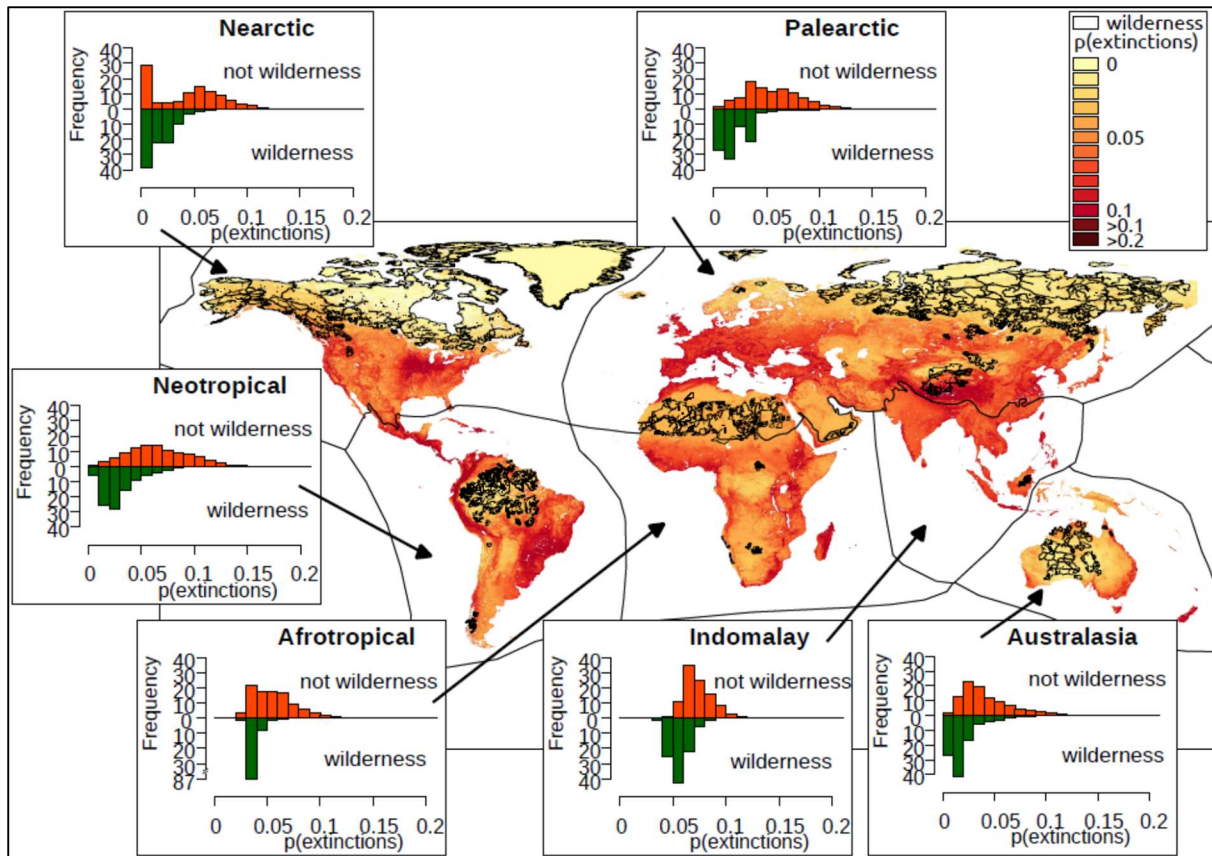
47 Wilderness areas, where industrial levels of human disturbance are absent or minimal^{9,10}, are
48 the last stronghold of intact ecosystems across Earth, but their extent has been increasingly
49 eroded with >10% of wilderness converted to human uses since the early 1990s^{10,11}. Yet,
50 little is known about the role wilderness plays in supporting biodiversity persistence, as
51 reflected in the absence of wilderness targets in the international environmental Agenda¹².
52 Here we address this knowledge gap, and provide the first estimate of the global significance

53 of wilderness areas for the persistence of terrestrial biodiversity. We use communities of
54 vascular plants and invertebrates as biodiversity surrogates, as these highly diverse and
55 customarily understudied^{13,14} groups represent the largest part of terrestrial biodiversity in
56 terms of species numbers and biomass (~60% of the species are invertebrates¹⁵, ~80% of the
57 biomass is from plants¹⁶).

58 We take advantage of an innovative approach¹⁷ to map the β -diversity of biological
59 communities – i.e. spatial variation in their species composition – based on generalised
60 dissimilarity modelling^{18,19}. Instead of delineating discrete community types, this method
61 assigns each location across the terrestrial surface of the Earth (represented here as a 1 km
62 grid cell) to a continuum of spatial turnover in biological composition. This approach predicts
63 the proportion of species which any two locations shared when both had intact habitat, as a
64 function of the environmental differences and the geographic separation of these locations.
65 Building on such prediction, and the current condition of habitats, we invoke the species-area
66 relationship (SAR) to estimate the proportion of species, in any given community, that are
67 expected to persist over the long term across the landscape^{20,21}. The complement to this
68 estimate represents the proportion of species committed to extinction – i.e. to disappear from
69 their entire distribution if habitat condition does not improve. For simplicity, we refer to the
70 set of species represented in a wilderness cell as a “wilderness community”, and the set of
71 species represented in a cell falling outside wilderness as a “non-wilderness community”.
72 Importantly, the continuous nature of our β -diversity estimates reflects the reality that a
73 certain proportion of species in a given wilderness community will also occur in cells found
74 outside wilderness, and vice-versa.

75 We found that wilderness areas act as a buffer against extinction risk. The global
76 probability of species extinction in non-wilderness communities (mean = 5.6%; sd= 2.8%) is
77 over twice as high as that of species in wilderness communities (mean = 2.1%; sd = 1.6%).

78 The buffering effect that wilderness has on extinction risk was found in every biogeographic
79 realm²², but was higher for those realms with larger extents of wilderness remaining, such as
80 the Palearctic (Fig. 1; Extended Data Table 1). Wilderness areas included the vast majority of
81 communities facing low extinction risk in the Nearctic, Palearctic, Neotropical, and
82 Australasian realms, where wilderness still has substantial coverage. The little remaining
83 wilderness of the Afrotropical realm also covered low-risk areas, even if some areas of low-
84 risk were found outside it. Communities in the IndoMalay realm faced the highest overall risk
85 of extinction and had the lowest wilderness coverage of all realms, confirming worldwide
86 concerns for the biodiversity of that region²³. The buffering effect of wilderness areas on
87 extinction risk was confirmed when looking separately at communities of vascular plants
88 (Extended Data Figure 1) and invertebrates (Extended Data Figure 2), with plants showing
89 higher extinction risk values overall. This result was also confirmed when we accounted for
90 the potential effect of habitat connectivity (Extended Data Figure 3), where the average
91 extinction risk for non-wilderness communities (mean = 6.9%; sd = 2.9%) was once again
92 twice as high as that of wilderness communities (mean = 3.5%; sd = 1.7%).



93

94 **Fig 1 Global probabilities of species extinction for communities of invertebrates and**
 95 **vascular plants associated with 1 km grid cells.**

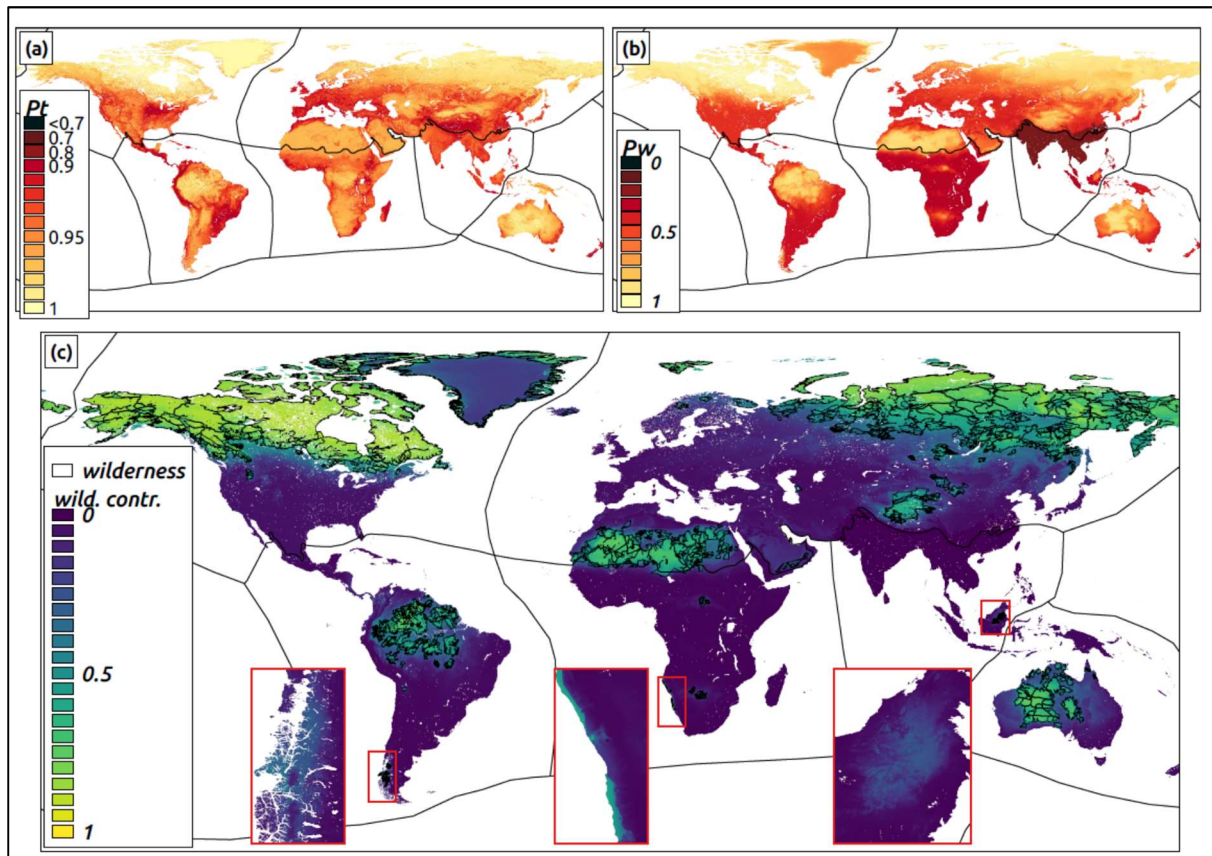
96 *The underlying map reports the estimated proportion of native species, originally associated*
 97 *with a particular grid cell, expected to disappear completely from their distribution due to*
 98 *the current condition of the habitats where they occur. The histogram bars represent the*
 99 *relative frequency distribution of the extinction risk values registered within areas of*
 100 *wilderness (green bars) and non-wilderness (orange bars) for each biogeographic realm.*

101

102 Given the continuous nature of our β -diversity predictions, wilderness habitat made a
 103 relative contribution (p_c) to the persistence of species in both wilderness and non-wilderness
 104 communities (Fig. 2). As expected, species persistence in wilderness communities was highly
 105 dependent on wilderness habitat (global mean p_c across wilderness communities was 68%),
 106 but many non-wilderness communities also had some degrees of dependency on habitat
 107 found within wilderness (global mean p_c across non-wilderness communities was 13%). This
 108 was especially the case for communities in the Amazon basin, and those found close to the

109 southern border of the Palearctic and Nearctic wilderness, where the survival of species was
110 largely dependent (up to 90%) on habitat in good condition inside wilderness areas.
111 Biogeographic realms characterised by larger extents of wilderness, such as the Nearctic and
112 the Palearctic, hosted communities with a higher dependency on wilderness habitat (mean p_c
113 was 48% in the Nearctic and 31% in the Palearctic). In particular, high dependency on
114 wilderness habitat was found for communities in northern America, northern Asia, the
115 Amazon basin, and the arid and semi-arid areas in northern Africa and central Australia.
116 Realms with limited wilderness remaining, such as the Afrotropics, showed contrasting
117 patterns. In some cases, such as the Kalahari savannas, the remaining wilderness made a
118 generally limited contribution to biodiversity persistence. In other cases, such as the Namib
119 Desert, the remaining wilderness made high contributions to persistence and acted as habitat
120 refugia for the biota found in that area.

121



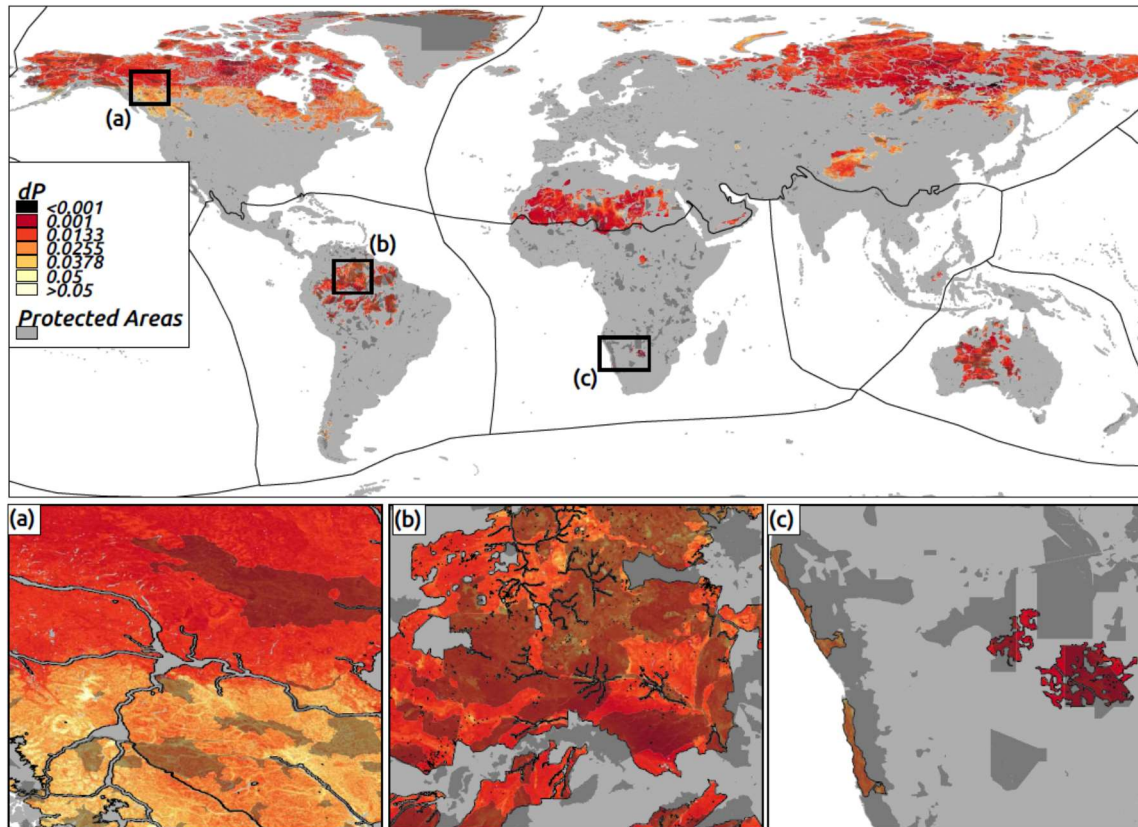
122

123 **Fig. 2 Relative contribution of wilderness areas to the persistence of plant and**
124 **invertebrate communities.**

125 Map (a) reports the probability of persistence (p_i) of species associated with any given grid
126 cell, accounting for the entire habitat surface of that community. Map (b) reports the
127 probability of persistence (p_w) considering only the habitat retained within wilderness. Map
128 (c) reports the proportional contribution ($p_c = p_w/p_i$) that wilderness areas make to the
129 persistence of species within each community. Note that maps (a) and (b) have a different
130 legend scale, to ensure readability.

131

132 We assessed the impact that direct loss of a given wilderness location would have on
133 biodiversity persistence ('delta persistence', δp ; Fig. 3). We found the potential losses in
134 probability of persistence were typically in the range 0.19%-3.65% worldwide (95% range of
135 δp values for wilderness locations). The extent to which a wilderness block represents the
136 biological diversity of a particular region (as described in the previous paragraph) was
137 reflected in the estimated reduction in species persistence that would result from habitat
138 degradation. Loss of wilderness areas characterised by a more unique biota (i.e. exhibiting
139 high endemism), and/or representing the last remaining good-quality habitat for a particular
140 biota, had a far higher impact on species persistence. For example, relatively low impact was
141 predicted to result from the loss of individual wilderness grid cells in the Kalahari savannas,
142 whereas higher impact was predicted from the loss of wilderness cells in the Namib Desert.
143 This relates to the different levels of endemism characterising these two areas, and the fact
144 that biodiversity in the Kalahari wilderness is surrounded by communities facing relatively
145 low risk of extinction, while that in the Namib wilderness is surrounded by communities
146 facing higher risk (Fig. 1).



147

148 **Fig. 3 Relative contribution of each wilderness grid cell to the estimated probability of**
 149 **persistence of species within invertebrate and vascular plant communities.**

150 The map reports the estimated impact of loss of a given 1km wilderness pixel, in terms of the
 151 consequent reduction in global species persistence (δp). Shaded polygons represent terrestrial
 152 protected areas. The inset maps report details of example wilderness areas in the Nearctic,
 153 Neotropics (b), and Afrotropics (c) realms.

154

155 In every biogeographic realm, with the exception of the IndoMalay, there were
 156 communities whose entire persistence depended mostly ($p_c > 70\%$) on habitat found within
 157 wilderness (Supplementary Table 1). For these communities, even the loss of a single grid
 158 cell of wilderness can have high impact on the persistence of species (δp up to 14% in the
 159 Neotropics). While the highest average δp values were found in the Nearctic and Palearctic,
 160 there were at least some highly valued blocks in every realm (Table 1; Extended Data Figure
 161 4). Overall these high-value blocks of wilderness were spread across different biome types,
 162 from arid environments to tropical moist forests, but were most stark in tropical and
 163 subtropical forests and shrublands. Yet, we found the level of protection²⁴ for wilderness

164 areas that make the highest contribution to species persistence differed very little from that of
165 other wilderness areas (Extended Data Figure 5). While the average δp value across protected
166 wilderness grid cells was slightly higher than random in all realms, with the exception of the
167 Neotropics, the difference in terms of effect size was small in the Afrotropics (*Cohen's d* =
168 0.33) and negligible elsewhere (*Cohen's d* < 0.2; Extended Data Table 2). This means that
169 wilderness areas where habitat loss would have the highest impact on biodiversity are not
170 better protected than other wilderness (i.e. 18.45% protection as a global average).

171

172 **Table 1** Contribution that wilderness areas make to overall species persistence within
 173 biological communities in each biogeographic realm.

174

Realm	Mean p_c	Max p_c	Block ID	Biome	Area km ²	Mean δp	Max δp
AA	0.239	0.792	26267	Trop & Subtrop Grass, Savan & Shrub*	12835	0.034	0.120
			25285	Trop & Subtrop Grass, Savan & Shrub*	41426	0.020	0.060
			26185	Trop & Subtrop Grass, Savan & Shrub*	38742	0.019	0.050
			25429	Trop & Subtrop Grass, Savan & Shrub*	25120	0.018	0.055
			25865	Trop & Subtrop Grass, Savan & Shrub*	40790	0.018	0.052
AT	0.031	0.717	27623	Deserts & Xeric Shrublands	20510	0.024	0.053
			29333	Deserts & Xeric Shrublands	20548	0.023	0.055
			20550	Trop & Subtrop Grass, Savan & Shrub*	83161	0.012	0.043
			19928	Trop & Subtrop Grass, Savan & Shrub*	44205	0.005	0.035
			27743	Trop & Subtrop Grass, Savan & Shrub*	20673	0.004	0.019
IM	0.007	0.455	21258	Trop & Subtrop Moist Broad Forests**	27837	0.013	0.073
			21094	Trop & Subtrop Moist Broad Forests**	43918	0.012	0.078
NA	0.481	0.900	9218	Temperate Conifer Forests	36061	0.036	0.102
			12514	Temperate Conifer Forests	18704	0.036	0.083
			8926	Tundra	34585	0.036	0.101
			7597	Boreal Forests/Taiga	24821	0.036	0.091
			12141	Temperate Conifer Forests	273538	0.033	0.091
NT	0.162	0.796	33835	Temperate Broadleaf & Mixed Forests	78296	0.026	0.139
			33404	Temperate Broadleaf & Mixed Forests	22240	0.023	0.078
			20311	Trop & Subtrop Moist Broad Forests**	13449	0.021	0.052
			24334	Trop & Subtrop Moist Broad Forests**	18693	0.018	0.053
			24997	Trop & Subtrop Moist Broad Forests**	20306	0.018	0.045
PA	0.309	0.865	16393	Deserts & Xeric Shrublands	19778	0.035	0.057
			15588	Deserts & Xeric Shrublands	23076	0.034	0.052
			7356	Tundra	43756	0.032	0.091
			8102	Boreal Forests/Taiga	26615	0.032	0.075
			16476	Deserts & Xeric Shrublands	51814	0.032	0.061

175

176 *Mean and maximum contribution (p_c) observed for communities in each biogeographic*
 177 *realm, and the mean and maximum reduction in persistence (δp) that would be associated to*
 178 *the loss of habitat in individual locations for each wilderness block. Only the 5 blocks with*
 179 *highest mean δp values are reported for each biogeographic realm (full dataset available in*
 180 *Supplementary Table 1). Only two wilderness blocks remain in the IndoMalay realm, both*
 181 *are listed. Realms acronyms are as follow: AA Australasia, AT Afrotropical, IM Indomalay,*
 182 *NA Nearctic, NT Neotropics, PA Palearctic. Two realms were excluded from analyses:*

183 *Oceanian, and Antarctic. * Tropical & Subtropical Grasslands, Savannas & Shrublands. ***
184 *Tropical & Subtropical Moist Broadleaf Forests.*
185

186

187 Earth's remaining intact ecosystems, increasingly seen as essential for the provision
188 of ecosystem services on which humanity relies²⁵ and for maintaining the bio-cultural
189 connections of indigenous communities²⁶, have been neglected by the biodiversity
190 conservation community. This is largely due to a belief they are less vulnerable to threatening
191 processes, of low species richness, and of low overall diversity¹². Recent analyses on
192 vertebrate taxa²⁷ found that areas of low human impact host fewer restricted-range species
193 than it would be expected by chance. These species might have lost part of their original
194 distribution as a consequence of rapid wilderness loss¹⁰. Our research shows many wilderness
195 areas are today critical in reducing extinction risk for terrestrial biodiversity. These areas are
196 important because they host highly unique biological communities, and/or represent the
197 majority of remaining natural habitats for biological communities that have suffered high
198 levels of habitat loss elsewhere. Alarming, these invaluable areas are not well protected.
199 Our findings point to the need for a targeted retention of the remaining wilderness areas to be
200 coupled with efforts aimed at protecting and restoring important habitats in degraded
201 environments²⁸. We believe it is vital that these two aims are viewed as highly
202 complementary, and non-substitutable, components of a truly integrated approach to
203 promoting the overall persistence of our planet's biodiversity. A strategic expansion of the
204 global protected area estate is needed to preserve those irreplaceable wilderness areas that are
205 at most risk, alongside national land-use legislation and the enforcement of business
206 standards for reducing industrial footprint in intact ecosystems^{8,12}. In addition, regions that
207 have already lost the largest part of their wilderness, such as the IndoMalay and the
208 Afrotropics, require conservation strategies that focus on the restoration of ecosystem

209 integrity²⁹. The value of wilderness in the international biodiversity agenda can be no longer
210 understated if nations are truly committed to achieve the Sustainable Development Goals³⁰.

211 **References**

- 212 1. Johnson, C. N. et al. Biodiversity losses and conservation responses in the
213 Anthropocene. *Science* 356, 270–275 (2017).
- 214 2. Steffen, W. et al. Planetary boundaries: Guiding human development on a changing
215 planet. *Science* 347, 1259855 (2015).
- 216 3. Cardinale, B. J. et al. Biodiversity loss and its impact on humanity. *Nature* 486, 59–67
217 (2012).
- 218 4. Ceballos, G., Ehrlich, P. R. & Dirzo, R. Biological annihilation via the ongoing sixth
219 mass extinction signaled by vertebrate population losses and declines. *Proc. Natl. Acad.*
220 *Sci.* 114, E6089–E6096 (2017).
- 221 5. Watson, J. E. M. et al. Persistent Disparities between Recent Rates of Habitat
222 Conversion and Protection and Implications for Future Global Conservation Targets.
223 *Conserv. Lett.* 9, 413–421 (2016).
- 224 6. Venter, O. et al. Sixteen years of change in the global terrestrial human footprint and
225 implications for biodiversity conservation. *Nat. Commun.* 7, 1–11 (2016).
- 226 7. Newbold, T. et al. Has land use pushed terrestrial biodiversity beyond the planetary
227 boundary? A global assessment. *Science* 353, 288–291 (2016).
- 228 8. Di Marco, M., Venter, O., Possingham, H. P. & Watson, J. E. M. Changes in human
229 footprint drive changes in species extinction risk. *Nat. Commun.* 9, 4621 (2018).
- 230 9. Mittermeier, R. a et al. Wilderness and biodiversity conservation. *Proc. Natl. Acad. Sci.*
231 *USA* 100, 10309–10313 (2003).
- 232 10. Watson, J. E. M. et al. Catastrophic Declines in Wilderness Areas Undermine Global
233 Environment Targets. *Curr. Biol.* 26, 2929–2934 (2016).

- 234 11. Allan, J. R., Venter, O. & Watson, J. E. M. Temporally inter-comparable maps of
235 terrestrial wilderness and the Last of the Wild. *Sci. Data* 4, 170187 (2017).
- 236 12. Watson, J. E. M. et al. Protect the last of the wild. *Nature* 563, 27–30 (2018).
- 237 13. Clark, J. A. & May, R. M. Taxonomic bias in conservation research. *Science* 297, 191–
238 192 (2002).
- 239 14. Di Marco, M., Butt, N., Possingham, H. P., Kearney, S. & Watson, J. E. M. Changing
240 trends and persisting biases in three decades of conservation science. *Glob. Ecol.*
241 *Conserv.* 10, (2017).
- 242 15. Chapman, A. D. Numbers of Living Species in Australia and the World. (2009). Report
243 for the Australian Biological Resources Study Canberra, Australia. Available from:
244 [http://www.environment.gov.au/biodiversity/abrs/publications/other/species-](http://www.environment.gov.au/biodiversity/abrs/publications/other/species-numbers/2009/06-references.html)
245 [numbers/2009/06-references.html](http://www.environment.gov.au/biodiversity/abrs/publications/other/species-numbers/2009/06-references.html)
- 246 16. Bar-On, Y. M., Phillips, R. & Milo, R. The biomass distribution on Earth. *Proc. Natl.*
247 *Acad. Sci. USA* 115, 6506–6511 (2018).
- 248 17. Hoskins, A. J. et al. Supporting global biodiversity assessment through high-resolution
249 macroecological modelling: Methodological underpinnings of the BILBI framework.
250 *bioRxiv* 309377 (2019). <https://doi.org/10.1101/309377>
- 251 18. Ferrier, S., Manion, G., Elith, J. & Richardson, K. Using generalized dissimilarity
252 modelling to analyse and predict patterns of beta diversity in regional biodiversity
253 assessment. *Divers. Distrib.* 13, 252–264 (2007).
- 254 19. Ferrier, S. et al. Mapping More of Terrestrial Biodiversity for Global Conservation
255 Assessment. *Bioscience* 54, 1101–1109 (2004).

- 256 20. Allnutt, T. F. et al. A method for quantifying biodiversity loss and its application to a
257 50-year record of deforestation across Madagascar. *Conserv. Lett.* 1, 173–181 (2008).
- 258 21. Di Marco, M. et al. Projecting impacts of global climate and land-use scenarios on plant
259 biodiversity using compositional-turnover modelling. *Glob. Chang. Biol.* 25, 2763–
260 2778 (2019).
- 261 22. Olson, D. M. et al. Terrestrial ecoregions of the world: a new map of life on earth.
262 *Bioscience* 51, 933–938 (2001).
- 263 23. Sodhi, N. S., Koh, L. P., Brook, B. W. & Ng, P. K. L. Southeast Asian biodiversity: an
264 impending disaster. *Trends Ecol. Evol.* 19, 654–60 (2004).
- 265 24. IUCN & UNEP-WCMC. The World Database on Protected Areas (WDPA) version
266 July/2018. Cambridge, UK: UNEP-WCMC (2018). Available at:
267 www.protectedplanet.net.
- 268 25. Potapov, P. et al. The last frontiers of wilderness: Tracking loss of intact forest
269 landscapes from 2000 to 2013. *Sci. Adv.* 3, e1600821 (2017).
- 270 26. Gorenflo, L. J., Romaine, S., Mittermeier, R. A. & Walker-Painemilla, K. Co-
271 occurrence of linguistic and biological diversity in biodiversity hotspots and high
272 biodiversity wilderness areas. *Proc. Natl. Acad. Sci. USA* 109, 8032–8037 (2012).
- 273 27. Pimm, S. L., Jenkins, C. N. & Li, B. V. How to protect half of Earth to ensure it protects
274 sufficient biodiversity. *Sci. Adv.* 4, eaat2616 (2018).
- 275 28. CBD. Strategic plan for Biodiversity 2011-2020. (2010).
- 276 29. Mappin, B. et al. Restoration priorities to achieve the global protected area target.
277 *Conserv. Lett.* e12646 (2019). doi:10.1111/conl.12646

278 30. United Nations General Assembly. Transforming our world: the 2030 Agenda for
279 Sustainable Development. (2015).

280

281 **Methods**

282

283 **Modelling compositional variation in biological communities**

284 Our analyses build on global models of compositional turnover (β -diversity) in biological
285 communities. This approach uses generalised dissimilarity modelling (GDM) to predict the
286 difference in species composition between pairs of sites, as a function of environmental
287 differences between, and spatial separation of, those sites^{18–20}. Modelled relationships
288 between spatial turnover in community composition and environmental gradients are used to
289 generate continuous predictions of β -diversity patterns within a region of interest, without
290 having to delineate communities as discrete entities (Extended Data Figure 6).

291 We employed compositional-turnover models for vascular plant and invertebrate
292 communities generated by Hoskins et al.¹⁷ using the global biodiversity modelling
293 infrastructure BILBI, as recently applied to projecting biodiversity trends under future
294 scenarios of socio-economic development^{21,31,32}. This infrastructure relies on a GDM
295 approach to predict spatial turnover in species composition between any pair of 30 arc-
296 seconds grid cells across the terrestrial surface of the planet (~ 1 km² at the equator). The
297 infrastructure uses a specially modified form of GDM, which corrects for biases introduced
298 into predictions when models are fitted to incomplete survey inventories. This is achieved by
299 replacing the response variable normally employed in GDM fitting (i.e. compositional
300 dissimilarity between pairs of sites) with the probability that a pair of observations drawn
301 randomly from two sites refer to the same or different species. This modelled probability is
302 then back-transformed to a measure of proportional dissimilarity in species composition
303 between communities. This modification of the standard GDM approach minimises the risk
304 that incompleteness and biases in survey inventories result in inflated estimates of turnover.

305 In the BILBI infrastructure, separate GDMs were built for each of the 61 biome-realm
306 combinations of the terrestrial globe²², with models fitted separately for invertebrates and
307 plants. Each model was fitted to species location records derived from the Global
308 Biodiversity Information Facility (GBIF; as detailed in Hoskins et al.¹⁷). A total of 132,761
309 species of invertebrates (with 13,244,784 location records) and 254,145 species of vascular
310 plants (with 52,489,096 location records) were employed globally. The selection of a reduced
311 subset of GBIF records followed both an extensive data cleaning and name matching process,
312 and a selection of plant and invertebrate taxa for which there were consistent collection
313 methodologies and communities of practice and relatively complete coverage, so as to
314 minimise the number of “single specimen” records. While GBIF data present inherent
315 limitations, especially in terms of the variation in sampling intensity for different parts of the
316 globe, the innovative enhancement to GDM modelling employed in our study reduces the
317 bias introduced by incomplete sampling^{17,33}. In fact, comparing “observation pairs” (as
318 opposed to site pairs) in the BILBI modelling infrastructure ensured that variation in
319 sampling intensity was effectively accounted for during model fitting, because the probability
320 that two observations in two sites refer to the same species is independent of the number of
321 other species observed. In doing so, our approach relies on the assumption that range-
322 restricted species are less likely to be found within the dataset being sampled, hence species
323 that are less likely to be sampled (due to their natural rarity) will increase our estimates of
324 dissimilarity in the areas they exist. Furthermore, by focusing on spatial patterns in a
325 collective property of biodiversity – i.e. compositional turnover – rather than modelling
326 distributions of individual species, the BILBI infrastructure is expected to achieve relatively
327 robust extrapolation of patterns across poorly-sampled regions, even when species occurring
328 in these regions are unsurveyed.

329 The proportional compositional dissimilarity between grid cells was predicted as a
330 function of the following environmental variables¹⁷: Minimum Monthly Temperature³⁴,
331 Maximum Monthly Temperature³⁴, Maximum Diurnal Temperature Range³⁴, Annual
332 Precipitation³⁴, Actual Evaporation³⁴, Potential Evaporation³⁴, Minimum Monthly Water
333 Deficit³⁴, Maximum Monthly Water Deficit³⁴, Soil pH³⁵, Soil Clay Proportion³⁵, Soil Silt
334 Proportion³⁵, Soil Bulk Density³⁵, Soil Depth³⁵, Ruggedness Index³⁶, Topographic Wetness
335 Index³⁵. All temperature, evaporation, and water deficit surfaces were adjusted for the effects
336 of topographic aspect and shading^{37,38}.

337

338 **Measuring the condition of habitats**

339 We estimated the current condition of habitats, using land-use maps for the year 2015 derived
340 from the latest update of the land-use harmonisation project³⁹ (LUH2). These maps represent
341 the percentage coverage, for each 0.25° grid cell of the globe, of 12 classes of land use:
342 forested land (primary or secondary), non-forested land (primary or secondary), managed
343 pasture, rangeland, urban land, C3 crops (annual, perennial, or nitrogen fixing), C4 crops
344 (annual, or perennial). Estimates of the proportional coverage for each land-use class was
345 downscaled from the original 0.25° resolution to a resolution of 30 arc-seconds
346 (approximately 1 km at the equator) following the approach described in Hoskins et al.³⁸, to
347 match the scale of biological communities and wilderness areas. Our approach differed
348 slightly from the original approach of Hoskins et al., to accommodate the added
349 computational complexity of fitting to 12 land-use classes instead of 5. It also differed in the
350 use of more recent datasets^{40–42} during the fitting process (see Supplementary Methods for
351 additional details). Following recent analyses^{21,31,32}, values for the 12 LUH2 classes were
352 combined into a cumulative habitat condition score, by multiplying each percentage land-use
353 value for a coefficient representing the proportional native species richness (or “ α diversity”)

354 expected to be retained under each land-use class, derived from the PREDICTS
355 database^{7,21,43,44}. The coefficients were estimated from a hierarchical mixed-effects model to
356 assess how natural species richness responds to land use change⁴³.

357 We also used the habitat condition surface as the basis for a sensitivity analysis on the
358 potential effect of habitat connectivity. Connectivity was calculated following Drielsma et
359 al.⁴⁵, assuming cell-wise permeability as a function of relative habitat condition. Since this
360 calculation multiplies the connectivity of a cell by its current condition, the resultant surface
361 is by definition lower than that measured by condition alone.

362

363 **Estimating biodiversity persistence and the risk of species extinctions**

364 We estimated the proportion (p_i) of species associated with each grid cell i expected to persist
365 anywhere within their range. We followed Allnutt et al.²⁰ in employing the Species-Area
366 relationship (SAR) to translate the ratio between the remaining area and the original (pre-
367 degradation) area of habitat across similar ecological environments (relative to the biological
368 community in a given cell i) into the proportion of species (p_i) expected to persist over the
369 long term. This value was derived as a function of the modelled similarity (s_{ij}) in species
370 composition between the focal cell i and other grid cells j found in the same biome-realm,
371 derived using the GDM approach described above, as well as the condition of habitat in each
372 of those cells (c_j):

373

$$374 \quad p_i = \left[\frac{\sum_{j=1}^n s_{ij} c_j}{\sum_{j=1}^n s_{ij}} \right]^z \quad [1]$$

375

376 where the numerator represents the condition-weighted area of habitat remaining across
377 similar ecological environments to grid cell i (i.e. the remaining extent of the biological
378 community comprising species originally present in cell i) and the denominator represents the

379 original area of similar ecological environments (i.e. the extent of that biological community
380 if all habitats were intact). The parameter z is the coefficient of the SAR, set to 0.25 as per
381 previous studies^{18–20,31}. After estimating persistence values p_i for each grid cell, we derived
382 extinction risk values (e_i), representing the proportion of species associated with each grid
383 cell i expected to be lost from their range, as a simple complement of persistence:

384

$$385 \quad e_i = 1 - p_i \quad [2]$$

386

387 Separate estimates of persistence and extinction were made for vascular plant
388 communities and invertebrate communities, and values were then averaged across the two
389 groups to report aggregated biodiversity results. It is important to clarify that this method (as
390 for any other method built on SAR theory) does not estimate the precise timing of extinction.
391 Rather, it estimates the proportion of species which are expected to become extinct over the
392 long term, as a consequence of the habitat conditions observed in the present time. We thus
393 invoke the concept of “species committed to extinction” (eg see Chaudhary & Mooers⁴⁶), as
394 those species originally present in an area that are estimated to disappear from their entire
395 range, given deterioration of habitat condition. Some of these extinction might have already
396 been realised at the time of assessment, while others are expected to be realised over longer
397 time periods into the future (as an extinction debt), unless habitat condition improves.

398

399 **Estimating the contribution of wilderness areas to biodiversity persistence**

400 We represented the distribution of wilderness areas using the map of terrestrial wilderness by
401 Allan et al.¹¹, at a global resolution of 1 km². The distribution of wilderness was derived by
402 identifying all areas free of human pressure and covering a contiguous area of $\geq 10,000$ km².
403 The estimate of human pressure was in turn derived from the Human Footprint map⁴⁷, a

404 representation of the cumulative human pressure on the environment. As wilderness
405 encompasses different regions with very diverse biological characteristics in terms of species
406 diversity, levels of endemism, and spatial turnover in species composition, we quantified its
407 role in promoting biodiversity persistence across different locations and across taxa. We did
408 this by estimating the extinction risk within wilderness communities versus that within non-
409 wilderness communities. A “wilderness community” is defined here as the set of species
410 associated with a cell found inside wilderness; the extinction risk for this community is
411 therefore calculated by making this the focal cell i in equations 1 and 2. Extinction risk for
412 each “non-wilderness community” is calculated in a similar manner, by making a particular
413 cell falling outside wilderness the focal cell i in equations 1 and 2.

414 We assessed the relative contribution that habitat found within wilderness areas
415 makes to the persistence of terrestrial biodiversity, both globally and within each
416 biogeographic realm²². To do so, we started from the estimate of the proportion of species
417 ($p_{i,t}$) associated with each grid cell i expected to persist considering any available habitat
418 (inside and outside wilderness). We then repeated this calculation considering only habitat
419 found inside wilderness grid cells. By re-running the BILBI infrastructure using this
420 “filtered” habitat condition map, we estimated the proportion of species associated with each
421 grid cell i expected to persist if wilderness were the only habitat remaining ($p_{i,w}$). By
422 comparing this latter value, based only on wilderness habitat, to the former value, based on
423 all habitat, we were able to measure the relative contribution ($p_{i,c}$) that wilderness areas make
424 to the total persistence of biodiversity associated to each grid cell:

425

$$426 \quad p_{i,c} = \frac{p_{i,w}}{p_{i,t}} \quad [3]$$

427

428 where $p_{i,c}$ values are by definition in the range 0 to 1, given $p_{i,w} \leq p_{i,t}$. This value represents an
429 estimate of the contribution that wilderness, as a whole, gives to the persistence of species in
430 any given biological community.

431 We also estimated the potential reduction in biodiversity persistence (δp) that would
432 result from the loss of habitat in any given wilderness grid cell, so as to identify those areas
433 where the impact of habitat loss would be highest. This value was calculated from the slope
434 of the species area curve (Eq.1) for the grid cell in question as:

435

$$436 \quad \delta p_i = \left[\frac{(\sum_{j=1}^n s_{ij} c_j) + 0.5}{\sum_{j=1}^n s_{ij}} \right]^Z - \left[\frac{(\sum_{j=1}^n s_{ij} c_j) - 0.5}{\sum_{j=1}^n s_{ij}} \right]^Z \quad [4],$$

437

438 which represents the potential impact of the removal of cell i in intact condition. This value
439 can be interpreted as the relative global change in the persistence of a given biological
440 community, which comprises all species found within a grid cell i , which would be expected
441 to result from the loss of habitat in that grid cell.

442 We reported the mean and maximum wilderness contribution values ($p_{i,c}$) observed
443 across grid cells within in each biogeographic realm. In addition, we reported the mean and
444 maximum delta persistence (δp) values observed across grid cells within each block of
445 wilderness (defined as individual patches of contiguous wilderness land). Once again,
446 separate analyses were run for vascular plant communities and invertebrate communities, and
447 values were then averaged to report aggregated biodiversity results.

448

449 **Measuring the protection level of wilderness areas with different biodiversity value**

450 We measured the relationship between δp values and protection status for each wilderness
451 pixel of the globe, using the World Database on Protected Areas²⁴. Following the description
452 in Butchart et al.⁴⁸, we excluded those internationally designated sites not considered as

453 protected areas, excluded ‘proposed’ sites and those with an unknown status, represented
454 sites without a defined shape as geodetic buffers of the appropriate area, and excluded
455 marine-only sites as well as the marine portion of coastal sites.

456 We assessed the effect size of the difference in mean δp values across protected and
457 non-protected wilderness grid cells in each realm using *Cohen’s d* statistic⁴⁹. We also
458 evaluated the difference between the observed mean δp value in protected wilderness areas
459 and that associated to 1,000 random samples of wilderness grid cells, each being of the same
460 size as the number of protected cells. We measured how many times the observed mean δp
461 was higher than the random mean δp .

462 Spatial data preparation was done in the GrassGIS⁵⁰, map outputs and layouts were
463 prepared in QGIS⁵¹, statistical analyses were performed in R⁵².

464

465 **Additional Methods references**

- 466 31. Kim, H. et al. A protocol for an intercomparison of biodiversity and ecosystem services
467 models using harmonized land-use and climate scenarios. *Geosci. Model Dev.* 11,
468 4537–4562 (2018).
- 469 32. Leclère, D. et al. Towards pathways bending the curve of terrestrial biodiversity trends
470 within the 21 st century. IIASA (2018). doi:10.22022/ESM/04-2018.15241
- 471 33. Ware, C. et al. Improving biodiversity surrogates for conservation assessment: A test
472 of methods and the value of targeted biological surveys. *Divers. Distrib.* 24, 1333–1346
473 (2018).
- 474 34. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high
475 resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25, 1965–
476 1978 (2005).
- 477 35. Hengl, T. et al. SoilGrids1km - Global soil information based on automated mapping.
478 *PLoS One* 9, e105992 (2014).
- 479 36. Amatulli, G. et al. Data Descriptor: A suite of global, cross-scale topographic variables
480 for environmental and biodiversity modeling. *Sci. Data* 5, 1–15 (2018).
- 481 37. Ferrier, S., Harwood, T., Williams, K. J. & Dunlop, M. Using Generalised Dissimilarity
482 Modelling to assess potential impacts of climate change on biodiversity composition in
483 Australia, and on the representativeness of the National Reserve System. (2012).
- 484 38. Hoskins, A. J. et al. Downscaling land-use data to provide global 30" estimates of five
485 land-use classes. *Ecol. Evol.* 6, 3040–3055 (2016).
- 486 39. Hurtt, G. C. et al. Harmonization of land-use scenarios for the period 1500-2100: 600
487 years of global gridded annual land-use transitions, wood harvest, and resulting
488 secondary lands. *Clim. Change* 109, 117–161 (2011).

- 489 40. European Commission, J. & Columbia University, C. GHS Population Grid, derived
490 from GPW4, Multitemporal (1975, 1990, 2000, 2015). (2015). Available at:
491 http://data.europa.eu/89h/jrc-ghsl-ghs_pop_gpw4_globe_r2015a.
- 492 41. Pesaresi, M. et al. GHS built-up grid, derived from Landsat, multitemporal (1975, 1990,
493 2000, 2014). Eur. Comm. Jt. Res. Centre, JRC Data Cat. (2015).
- 494 42. DiMiceli, C. M. et al. Annual global automated MODIS vegetation continuous fields
495 (MOD44B) at 250 m spatial resolution for data years beginning day 65, 2000--2010,
496 collection 5 percent tree cover. Univ. Maryland, Coll. Park. MD, USA (2011).
- 497 43. Hudson, L. N. et al. The database of the PREDICTS (Projecting Responses of
498 Ecological Diversity In Changing Terrestrial Systems) project. *Ecol. Evol.* 7, 145–188
499 (2017).
- 500 44. Hill, S. L. L. et al. Worldwide impacts of past and projected future land-use change on
501 local species richness and the Biodiversity Intactness Index. *bioRxiv* 311787 (2018).
502 doi:10.1101/311787
- 503 45. Drielsma, M., Ferrier, S. & Manion, G. A raster-based technique for analysing habitat
504 configuration: The cost-benefit approach. *Ecol. Modell.* 202, 324–332 (2007).
- 505 46. Chaudhary, A. & Mooers, A. Terrestrial Vertebrate Biodiversity Loss under Future
506 Global Land Use Change Scenarios. *Sustainability* 10, 2764 (2018).
- 507 47. Venter, O. et al. Global terrestrial Human Footprint maps for 1993 and 2009. *Sci. Data*
508 3, 273–281 (2016).
- 509 48. Butchart, S. H. M. et al. Shortfalls and Solutions for Meeting National and Global
510 Conservation Area Targets. *Conserv. Lett.* 8, 329–337 (2015).

- 511 49. Cohen, J. Statistical power analysis for the behavioral sciences (2nd ed.). (Academic
512 Press, 1988).
- 513 50. GRASS Development Team. Geographic Resources Analysis Support System (GRASS
514 GIS) Software, Version 7.2. (2017). Available at: <http://grass.osgeo.org>
- 515 51. QGIS Development Team. QGIS Geographic Information System. Open Source
516 Geospatial Foundation Project (2017). Available at: <http://qgis.osgeo.org>.
- 517 52. R Core Team. R: A Language and Environment for Statistical Computing. (2018).
518 Available at: <https://www.r-project.org/>.

519

520 **Data availability**

521 All input data used in these analyses derive from published sources cited in the methods
522 section. Supplementary Table S1 and Extended Data Tables 1 and 2 report the results for each
523 realm and each wilderness block. Any other datasets generated in the current study are available
524 from the corresponding author upon reasonable request.

525

526 **Computer code**

527 An R code to derive estimates of compositional dissimilarity and the proportion of persisting
528 species is available from <https://doi.org/10.1101/309377>.

529

530 **Acknowledgements**

531 This work was funded by Research Agreement n. 2017113325 between CSIRO and The
532 University of Queensland. MDM acknowledges support from the European Union's H2020
533 research and innovation programme (Marie Skłodowska-Curie grant agreement n° 793212).

534

535 **Author contributions**

536 MDM, SF, and JEMW framed the study. MDM, TDH and AJH carried out the analyses.

537 MDM, SF, TDH, AJH, and JEMW discussed and interpreted the results. MDM, SF, and

538 JEMW wrote the manuscript with support from TDH, and AJH.

539

540 **Competing interests**

541 Authors declare no competing interests.

542

543 **Supplementary Information**

544 Supplementary Methods describe the land-use downscaling process. Supplementary Table 1

545 reports the contribution that wilderness areas make to the overall persistence of communities,

546 aggregated by individual wilderness block.

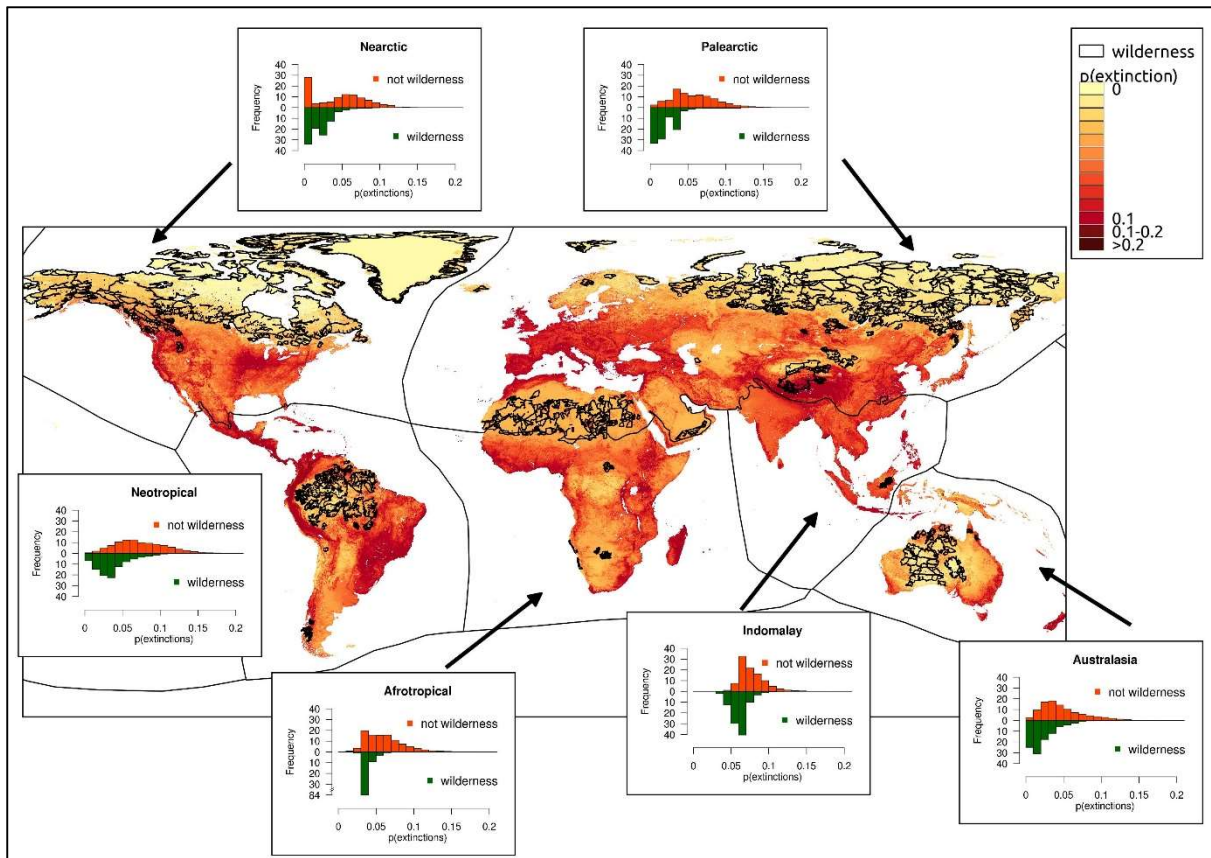
547

548 **Corresponding author**

549 Moreno Di Marco, Dept. of Biology and Biotechnology, Sapienza University of Rome, viale

550 dell'Università 32, 00185 Rome, Italy. E: moreno.dimarco@gmail.com T: +39 06 4991 4759.

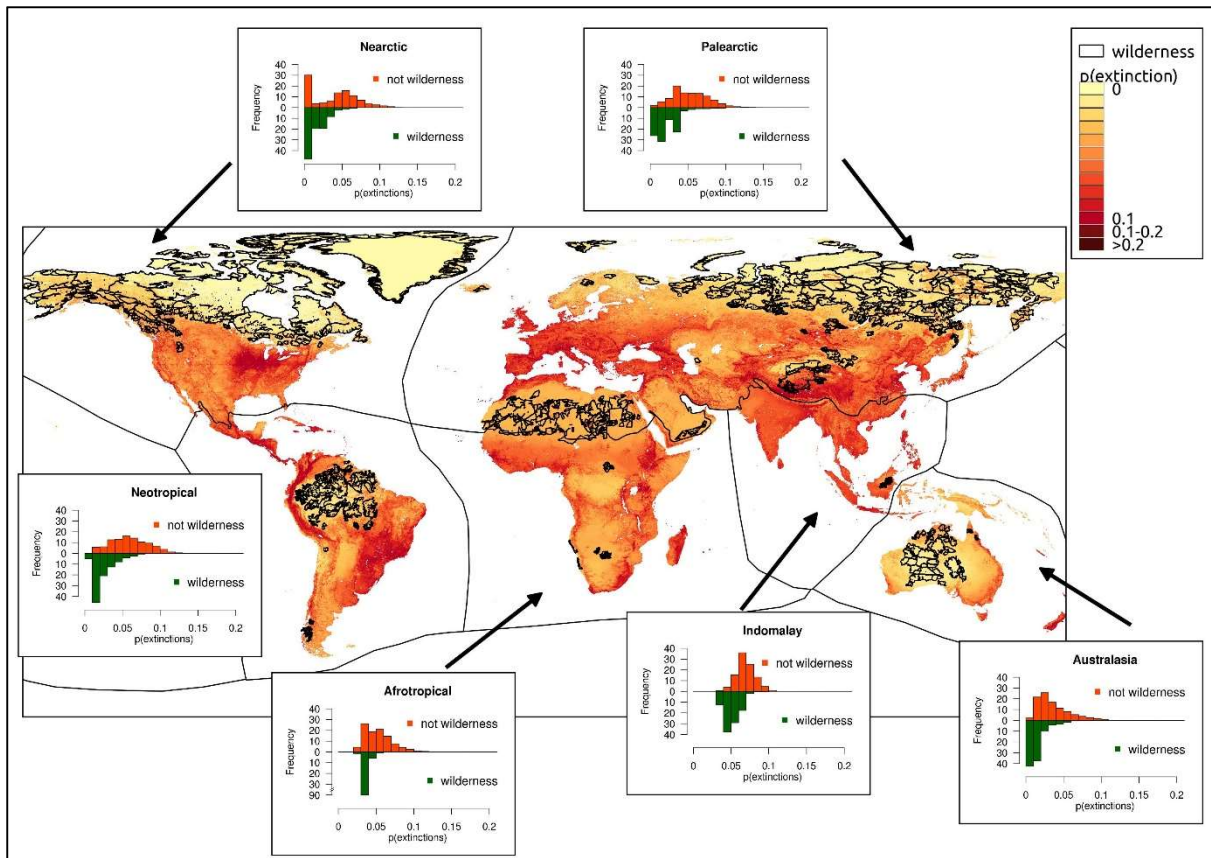
551



552

553 **Extended Data Figure 1. Global-scale probabilities of species extinction for communities**
 554 **of vascular plants associated with each grid cell.**

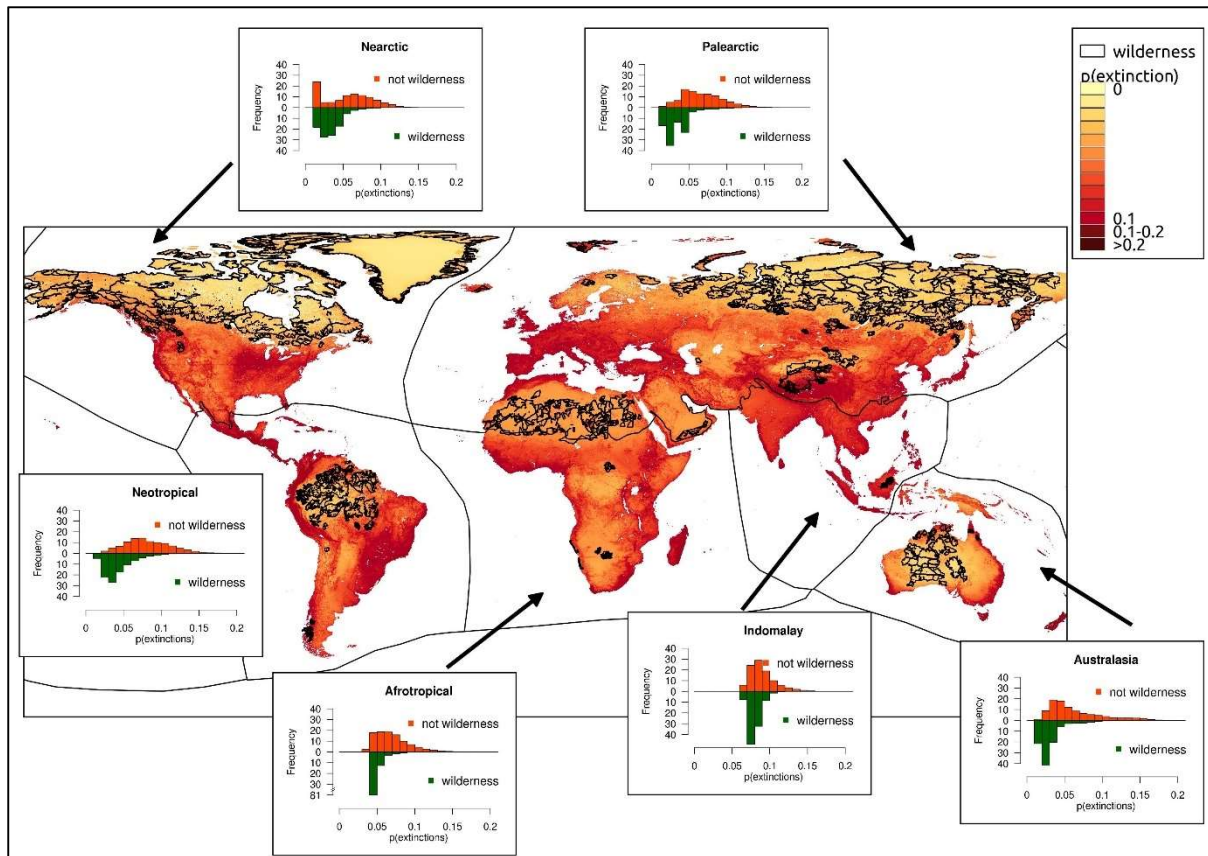
555 The underlying map reports the estimated proportion of native species, originally associated
 556 with a particular grid cell, expected to disappear from their distribution due to the current
 557 condition of the habitats where they occur.



558

559 **Extended Data Figure 2. Global-scale probabilities of species extinction for communities**
 560 **of invertebrates associated with each grid cell.**

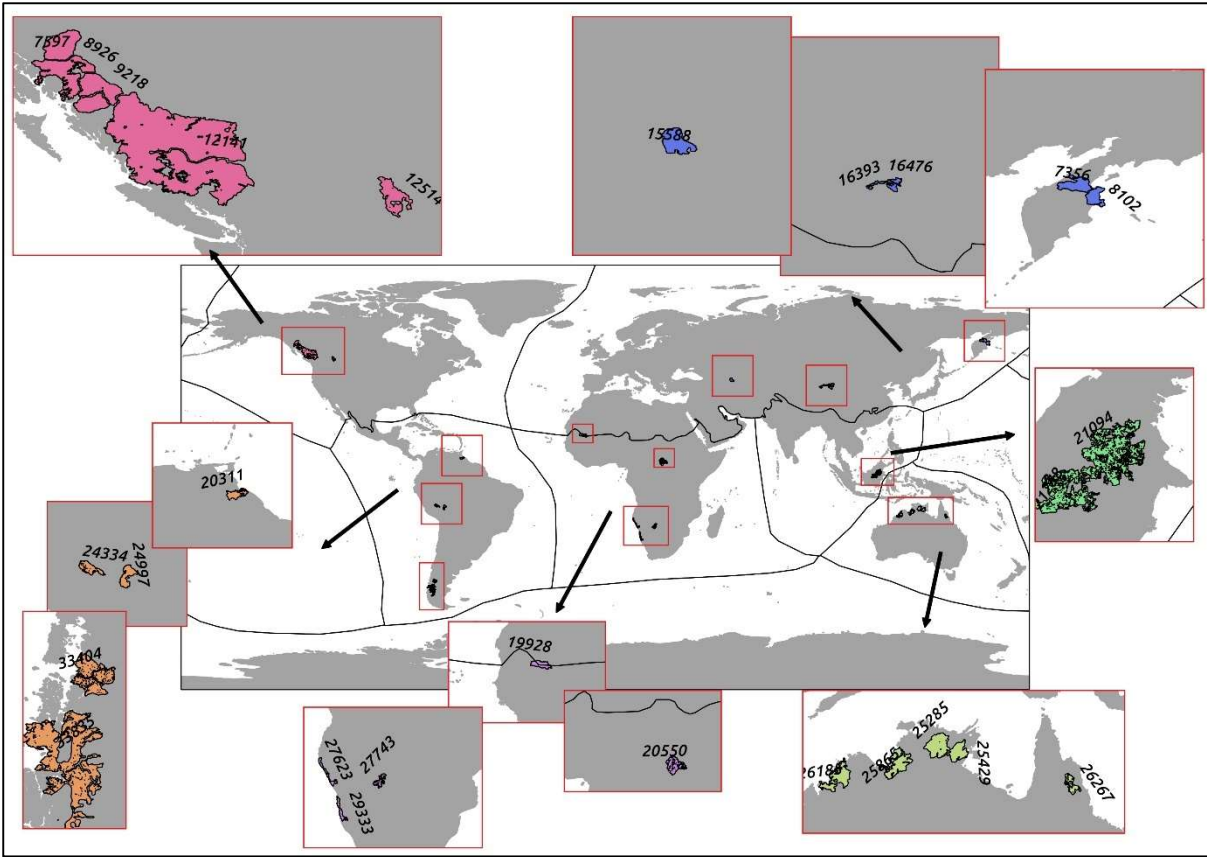
561 The underlying map reports the estimated proportion of native species, originally associated
 562 with a particular grid cell, expected to disappear from their distribution due to the current
 563 condition of the habitats where they occur.



564

565 **Extended Data Figure 3. Global-scale probabilities of species extinction for communities**
 566 **of invertebrates and vascular plants associated with each grid cell, accounting for habitat**
 567 **connectivity.**

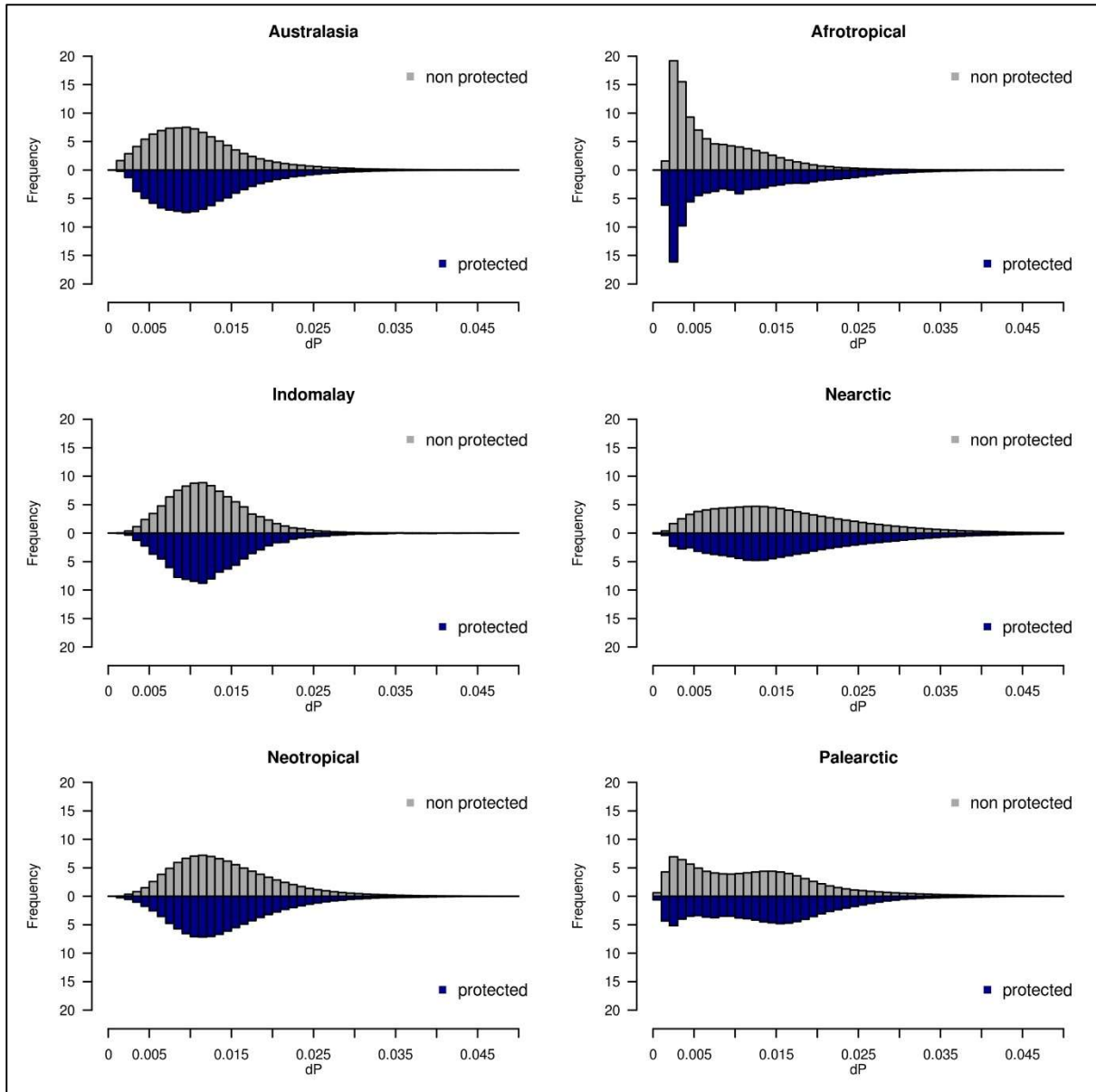
568 The underlying map reports the estimated proportion of native species, originally associated
 569 with a particular grid cell, expected to disappear from their distribution due to the current
 570 condition of the habitats where they occur, as well as the level of connectivity between habitats.



571

572 **Extended Data Figure 4. Distribution of the top-5 wilderness blocks identified for each**

573 **realm.** Numbers in the map report ID codes for the block (corresponding to Table S1).

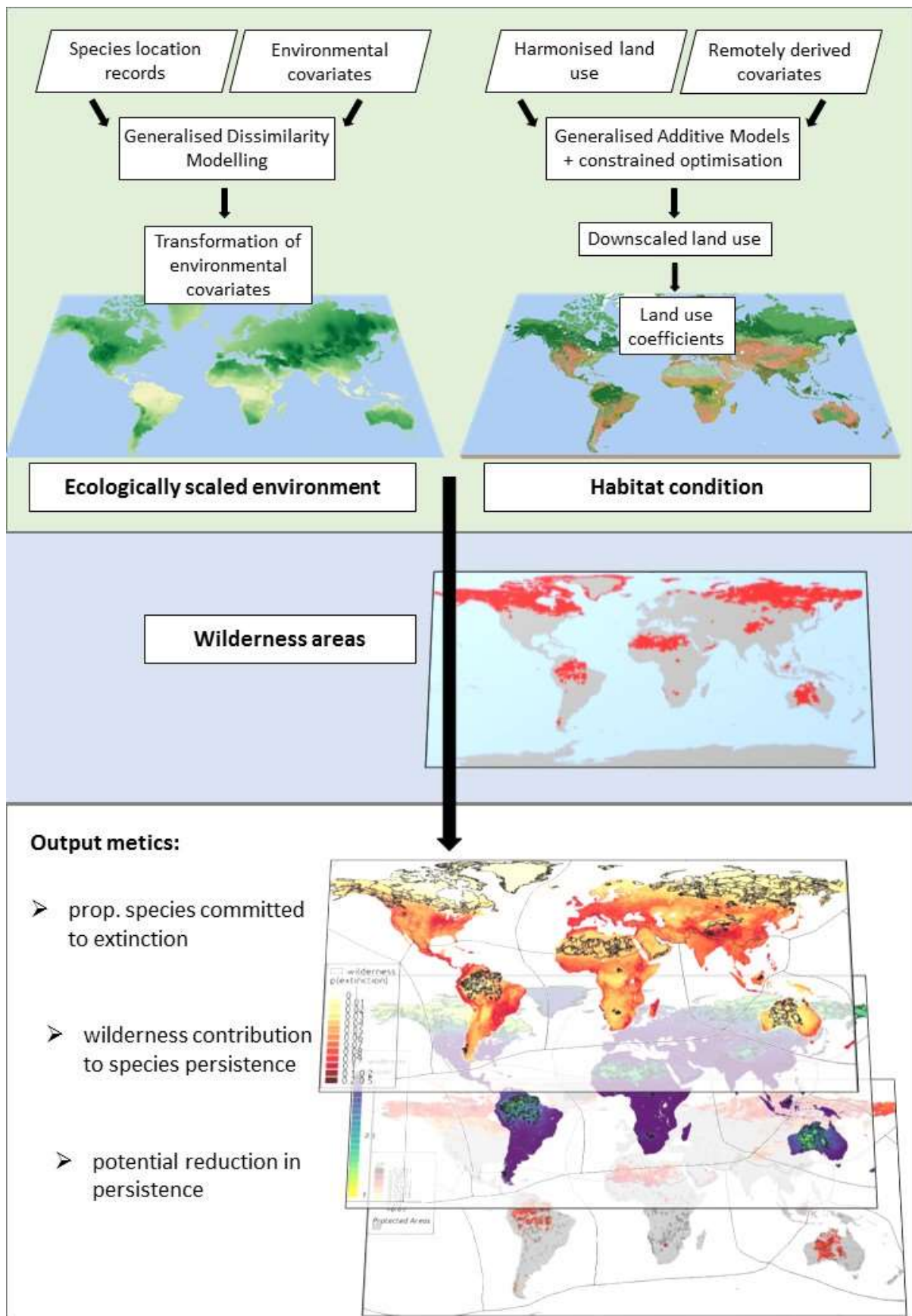


574

575 **Extended Data Figure 5. Frequency distribution of the contribution that individual**
 576 **wilderness grid cells make to the probability of persistence of invertebrate and vascular**
 577 **plant communities (δp).**

578 The histogram bars represent the relative frequency distribution of the δp values for wilderness
 579 pixels inside (blue bars) and outside (grey bars) protected areas, in each biogeographic realm.

580



582 **Extended Data Figure 6. Analytical framework used to estimate the probability of**
583 **persistence of biological communities.**

584 The framework combines estimates of spatial turnover in species composition, from which
585 ecologically scaled environments are derived, and estimates of habitat condition. The
586 framework produces a spatially explicit (1km) estimate of biodiversity persistence, from which
587 a number of metrics are derived: proportion of species committed to extinction, contribution of
588 wilderness areas to global species persistence, and potential reduction in persistence in case of
589 wilderness degradation.

590

Realm	Inside wilderness	Outside wilderness
Australasia	0.019 (0.014)	0.044 (0.025)
Afrotropical	0.038 (0.006)	0.054 (0.021)
IndoMalay	0.057 (0.013)	0.077 (0.018)
Nearctic	0.017 (.015)	0.042 (0.032)
Neotropics	0.031 (0.019)	0.066 (0.028)
Palaearctic	0.021 (0.017)	0.057 (0.026)

591

592 **Extended Data Table 1** Mean extinction risk (with standard deviation in parentheses)

593 observed across communities of invertebrates and vascular plants in each biogeographic realm,

594 inside and outside wilderness areas.

595

Metric \ Realm	AA	AT	IM	NA	NT	PA
mean δp in protected areas	0.0119	0.0105	0.0126	0.0171	0.0144	0.0142
average random mean δp	0.0115	0.0083	0.0124	0.0164	0.0145	0.0126
observed > random	100%	100%	100%	100%	0%	100%
Cohen's d	0.1024	0.3300	0.0412	0.0870	-0.0058	0.1868

596

597

598 **Extended Data Table 2** Difference in the estimated reduction of global species persistence

599 (δp) associated to the loss of a protected or non-protected wilderness pixel. The first row reports

600 the mean δp values observed across all protected wilderness grid cells of a biogeographic

601 realm; the second row reports the average across 1,000 mean δp values obtained by randomly

602 selecting an equivalent number of wilderness cells; the third column reports the percentage

603 times in which the observed mean δp was higher than the mean δp from a random sample (out

604 of 1,000 random samples); the last row reports the effect size (Cohen's d statistic) of the

605 difference between δp values in protected and non-protected wilderness cells. Realms

606 acronyms are as follow: AA Australasia, AT Afrotropical, IM Indomalay, NA Nearctic, NT

607 Neotropics, PA Palearctic. Two realms were excluded from analyses: Oceanian, and Antarctic.

608