

Wilderness areas halve the extinction risk of terrestrial biodiversity

Moreno Di Marco^{1,2*}, Simon Ferrier³, Tom D. Harwood³, Andrew J. Hoskins⁴ & James E. M. Watson^{5,6}

Reducing the rate of global biodiversity loss is a major challenge facing humanity¹, as the consequences of biological annihilation would be irreversible for humankind^{2–4}. Although the ongoing degradation of ecosystems^{5,6} and the extinction of species that comprise them^{7,8} are now well-documented, little is known about the role that remaining wilderness areas have in mitigating the global biodiversity crisis. Here we model the persistence probability of biodiversity, combining habitat condition with spatial variation in species composition, to show that retaining these remaining wilderness areas is essential for the international conservation agenda. Wilderness areas act as a buffer against species loss, as the extinction risk for species within wilderness communities is—on average—less than half that of species in non-wilderness communities. Although all wilderness areas have an intrinsic conservation value^{9,10}, we identify the areas on every continent that make the highest relative contribution to the persistence of biodiversity. Alarming, these areas—in which habitat loss would have a more-marked effect on biodiversity—are poorly protected. Given globally high rates of wilderness loss¹⁰, these areas urgently require targeted protection to ensure the long-term persistence of

biodiversity, alongside efforts to protect and restore more-degraded environments.

Wilderness areas, in which industrial levels of human disturbance are absent or minimal^{9,10}, are the last stronghold of intact ecosystems across Earth. However, their extent has rapidly decreased over past decades; more than 10% of the wilderness that existed in the early 1990s has since been converted to human use^{10,11}. Little is known about the role that wilderness has in supporting the persistence of biodiversity, as reflected in the absence of wilderness targets in the international environmental agenda¹². Here we address this knowledge gap and provide—to our knowledge—the first estimate of the global importance of wilderness areas for the persistence of terrestrial biodiversity. We use communities of vascular plants and invertebrates as surrogates for biodiversity, as these highly diverse and customarily understudied^{13,14} groups represent the largest part of terrestrial biodiversity in terms of both species numbers and biomass (about 60% of species are invertebrates¹⁵ and about 80% of biomass is from plants¹⁶).

We take advantage of an approach¹⁷ that maps the β -diversity of biological communities—that is, the spatial variation in their species composition—on the basis of generalized dissimilarity modelling^{18,19}.

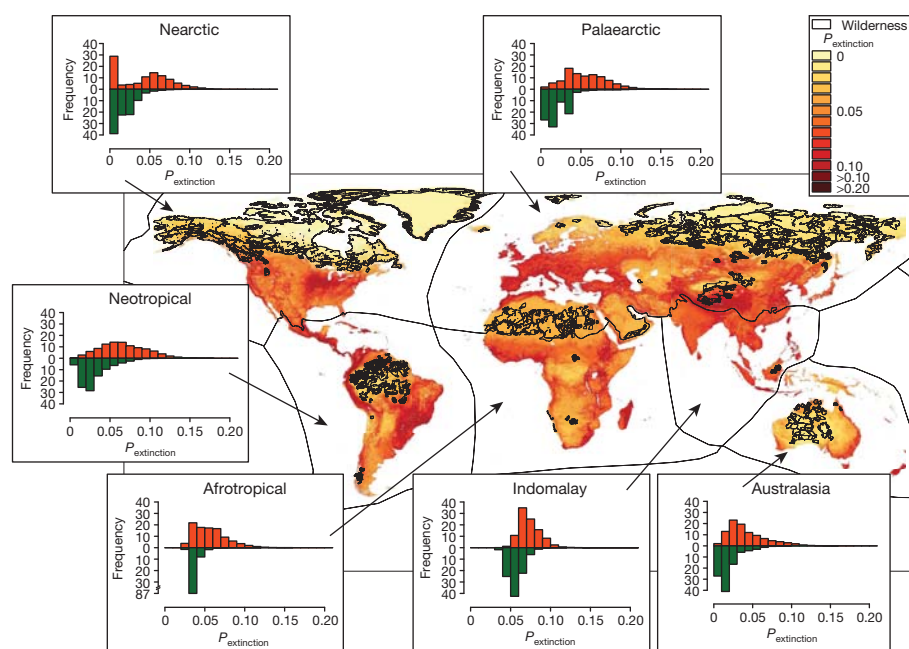


Fig. 1 | Global probabilities of species extinction for communities of invertebrates and vascular plants associated with 1-km² grid cells. The underlying map reports the estimated proportion of native species (originally associated with a particular grid cell) expected to disappear completely from their distribution, owing to the current condition of the

habitats in which they occur. The histogram bars represent the relative frequency distribution of the probability of extinctions ($P_{\text{extinction}}$) that were registered within areas of wilderness (green bars) and non-wilderness (orange bars), for each biogeographical realm.

¹CSIRO Land and Water, Dutton Park EcoSciences Precinct, Brisbane, Queensland, Australia. ²Department of Biology and Biotechnology, Sapienza University of Rome, Rome, Italy. ³CSIRO Land and Water, Black Mountain Laboratories, Canberra, Australian Capital Territory, Australia. ⁴CSIRO Health and Biosecurity, James Cook University, Townsville, Queensland, Australia. ⁵Centre for Biodiversity and Conservation Science, The University of Queensland, Brisbane, Queensland, Australia. ⁶Global Conservation Program, Wildlife Conservation Society, New York, NY, USA. *e-mail: moreno.dimarco@gmail.com

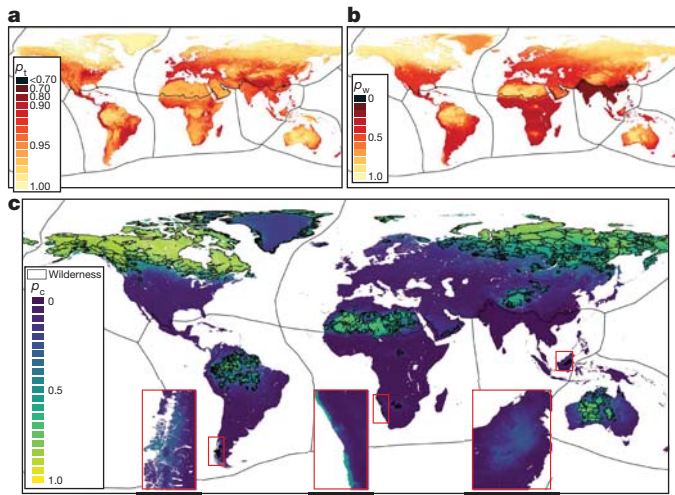


Fig. 2 | Relative contribution of wilderness areas to the persistence of plant and invertebrate communities. **a**, This map reports the total probability of persistence (p_i) of species associated with any given grid cell, accounting for the entire habitat surface of that community. **b**, This map reports the probability of persistence considering only the habitat retained within wilderness (p_w). **c**, This map reports the proportional contribution ($p_c = p_w/p_i$) that wilderness areas make to the probability of persistence of species within each community. Note that **a** and **b** have a different legend scale, to ensure readability.

Instead of delineating discrete types of community, this method assigns each location across the terrestrial surface of the Earth (represented here as a 1-km² grid cell) to a continuum of spatial turnover in biological composition. This approach predicts the proportion of species that any two locations shared when both locations had intact habitat, as a function of the environmental differences and the geographical separation of these locations. Building on such predictions, and the current condition of habitats, we invoke the species–area relationship to estimate the proportion of species in any given community that are expected to persist over the long term across the landscape^{20,21}.

The complement to this estimate represents the proportion of species that are committed to extinction—that is, to disappear from their entire distribution if the habitat condition does not improve. For simplicity, we refer to the set of species represented in a cell from a wilderness area as a ‘wilderness community’ and the set of species represented in a cell that falls outside a wilderness area as a ‘non-wilderness community’. Importantly, the continuous nature of our estimates of β -diversity reflects the reality that a proportion of species in a given wilderness community will also occur in cells found outside wilderness, and vice versa.

We found that wilderness areas act as a buffer against extinction risk. The global probability of species extinction in non-wilderness communities (mean = 5.6%, s.d. = 2.8%) is over twice as high as that of species in wilderness communities (mean = 2.1%, s.d. = 1.6%). The buffering effect that wilderness has on extinction risk was found in every biogeographical realm²², but was higher for realms with larger remaining extents of wilderness such as the Palaearctic (Fig. 1, Extended Data Table 1). Wilderness areas included the vast majority of communities that face low levels of extinction risk in the Nearctic, Palaearctic, Neotropical and Australasian realms, within which wilderness has substantial coverage. The little remaining wilderness of the Afrotropical realm also covered low-risk areas, even if some areas of low-risk were found outside it. Communities in the Indomalayan realm faced the highest overall risk of extinction and had the lowest wilderness coverage of all realms, which provides confirmation of concerns for the biodiversity of this region²³. The buffering effect of wilderness areas on extinction risk was confirmed when looking separately at communities of vascular plants (Extended Data Fig. 1) and invertebrates (Extended Data Fig. 2), with plants showing higher overall values of extinction risk. This result was also confirmed when we accounted for the potential effect of habitat connectivity (Extended Data Fig. 3); after doing so, the average extinction risk for non-wilderness communities (mean = 6.9%, s.d. = 2.9%) was once again twice as high as that of wilderness communities (mean = 3.5%, s.d. = 1.7%).

Given the continuous nature of our β -diversity predictions, wilderness habitat made a relative contribution to the persistence (p_c) of species in both wilderness and non-wilderness communities (Fig. 2).

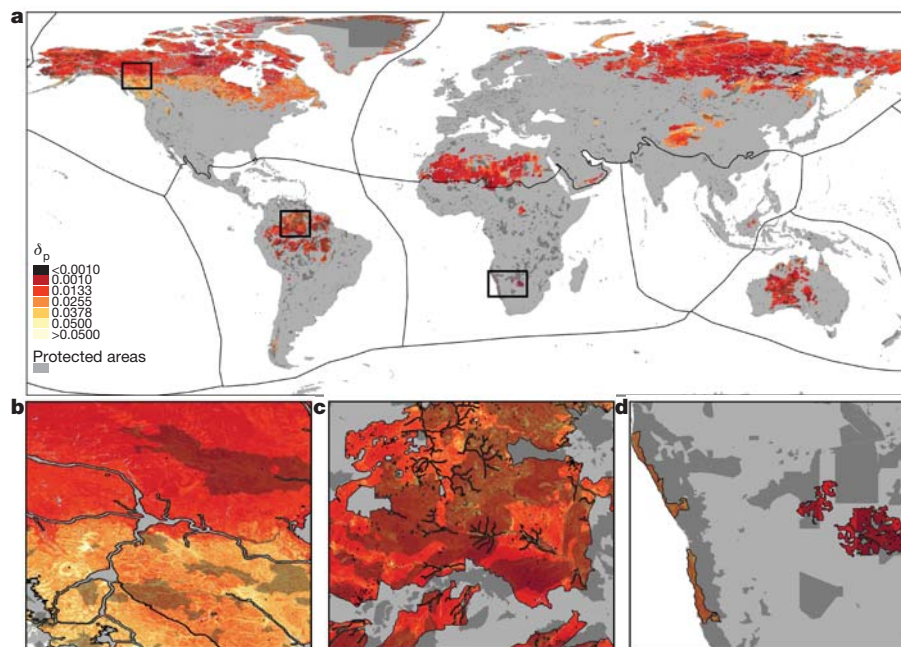


Fig. 3 | Relative contribution of each wilderness grid cell to the estimated probability of persistence of species within invertebrate and vascular plant communities. **a**, Map reporting the estimated effect of loss of a given 1-km² wilderness pixel, in terms of the consequent probability of reduction in global species persistence (δ_p). Polygons overlain with a dark-

grey shade represent terrestrial protected areas. Left, middle and right black boxes denote areas shown in **b**, **c** and **d**, respectively. **b–d**, The inset maps report details of examples of wilderness areas in the Nearctic (**b**), Neotropical (**c**) and Afrotropical (**d**) realms.

Table 1 | Contribution of wilderness areas to overall species persistence within biological communities in each biogeographical realm

Realm	Mean p_c	Maximum p_c	Block ID	Biome	Area (km ²)	Mean δ_p	Maximum δ_p
Australasian	0.239	0.792	26267	TSGSS	12,835	0.034	0.120
			25285	TSGSS	41,426	0.020	0.060
			26185	TSGSS	38,742	0.019	0.050
			25429	TSGSS	25,120	0.018	0.055
			25865	TSGSS	40,790	0.018	0.052
Afrotropical	0.031	0.717	27623	DXS	20,510	0.024	0.053
			29333	DXS	20,548	0.023	0.055
			20550	TSGSS	83,161	0.012	0.043
			19928	TSGSS	44,205	0.005	0.035
			27743	TSGSS	20,673	0.004	0.019
Indomalayan	0.007	0.455	21258	TSMBF	27,837	0.013	0.073
			21094	TSMBF	43,918	0.012	0.078
Nearctic	0.481	0.900	9218	TCF	36,061	0.036	0.102
			12514	TCF	18,704	0.036	0.083
			8926	Tundra	34,585	0.036	0.101
			7597	BFT	24,821	0.036	0.091
			12141	TCF	273,538	0.033	0.091
Neotropical	0.162	0.796	33835	TBMF	78,296	0.026	0.139
			33404	TBMF	22,240	0.023	0.078
			20311	TSMBF	13,449	0.021	0.052
			24334	TSMBF	18,693	0.018	0.053
			24997	TSMBF	20,306	0.018	0.045
Palaeartic	0.309	0.865	16393	DXS	19,778	0.035	0.057
			15588	DXS	23,076	0.034	0.052
			7356	Tundra	43,756	0.032	0.091
			8102	BFT	26,615	0.032	0.075
			16476	DXS	51,814	0.032	0.061

Mean and maximum contribution to persistence (p_c) observed for communities in each biogeographical realm, and mean and maximum reduction in persistence (δ_p) that would be associated with habitat loss in individual locations for each wilderness block. Only the five blocks with highest mean δ_p values are reported for each biogeographical realm (full dataset is in Supplementary Table 1). Only two wilderness blocks remain in the Indomalayan realm; both are listed. The Oceanian and Antarctic realms were excluded from analyses. BFT, boreal forests and taiga; DXS, deserts and xeric shrublands; TBMF, temperate broadleaf and mixed forests; TCF, temperate conifer forests; TSMBF, tropical and subtropical moist broadleaf forests; TSGSS, tropical and subtropical grasslands, savannahs and shrublands.

As expected, species persistence in wilderness communities was highly dependent on wilderness habitat (global mean p_c across wilderness communities was 68%), but many non-wilderness communities also had some degree of dependency on habitats found within wilderness areas (global mean p_c across non-wilderness communities was 13%). This was especially the case for communities in the Amazon basin and communities close to the southern border of the Palaeartic and Nearctic wilderness; in these areas, the survival of species was largely dependent (up to 90%) on habitat in good condition inside wilderness areas. Biogeographical realms characterized by larger extents of wilderness, such as the Nearctic and the Palaeartic, hosted communities with a higher dependency on wilderness habitat (mean p_c was 48% in the Nearctic and 31% in the Palaeartic). In particular, a high dependency on wilderness habitat was found for communities in northern America, northern Asia, the Amazon basin, and arid and semi-arid areas in northern Africa and central Australia. Realms with limited remaining wilderness, such as the Afrotropical realm, showed contrasting patterns. In some areas (such as the Kalahari), the remaining wilderness areas made a generally limited contribution to the persistence of biodiversity. In other cases (such as the Namib Desert), the remaining wilderness areas made high contributions to persistence and acted as habitat refugia for the biota found in the area.

We assessed the effect that the direct loss of a given wilderness location would have on the persistence of biodiversity (the 'delta persistence' value, δ_p) (Fig. 3). We found that the potential losses in the probability of persistence were typically in the range of 0.19 to 3.65% worldwide (95% range of δ_p values for wilderness locations). The extent

to which a wilderness block (that is, an individual patch of contiguous wilderness area) represents the biological diversity of a particular region was reflected in the estimated reduction in species persistence that would result from habitat degradation. The loss of wilderness areas that are characterized by a more-unique biota (that is, areas that exhibit high endemism) and/or that represent the last remaining good-quality habitat for a particular biota had a far greater effect on species persistence. For example, our analysis predicted that the loss of individual grid cells from wilderness areas in the Kalahari savannahs would have relatively little effect, whereas a greater effect was predicted for the loss of cells from wilderness areas in the Namib Desert. This relates to the different levels of endemism that characterize these two areas, and to the fact that biodiversity living in the wilderness areas of the Kalahari are surrounded by communities that face a relatively low risk of extinction, whereas biodiversity living in the wilderness areas of the Namib Desert are surrounded by communities that face a higher risk of extinction (Fig. 1).

In every biogeographical realm (except the Indomalayan realm), the persistence of a subset of communities depended mostly ($p_c > 70\%$) on habitat found within wilderness areas (Supplementary Table 1). For these communities, the loss of even a single grid cell of wilderness area can have a large effect on the persistence of species (a δ_p value of up to 14% in the Neotropical realm). Although the highest average δ_p values were found in the Nearctic and Palaeartic realms, there were at least some highly valued blocks in every realm (Table 1, Extended Data Fig. 4). Overall, these high-value blocks of wilderness were spread across biome types from arid environments to tropical moist forests, but

were more common in tropical and subtropical forests and shrublands. However, we found that the level of formal protection²⁴ for wilderness areas that make the highest contributions to species persistence differed very little from that of other wilderness areas (Extended Data Fig. 5). Although the average δ_p value across grid cells from protected wilderness areas was slightly higher than random in all realms (with the exception of the Neotropical realm), the difference in terms of effect size was small in the Afrotropical realm (Cohen's $d = 0.33$) and negligible elsewhere (Cohen's $d < 0.2$) (Extended Data Table 2). This means that wilderness areas in which habitat loss would have the greatest effect on biodiversity are not better-protected than other wilderness areas (18.45% protection as a global average).

The remaining intact ecosystems of Earth—which are increasingly seen as essential for providing ecosystem services on which humanity relies²⁵ and maintaining the bio-cultural connections of indigenous communities²⁶—have been neglected in efforts to conserve biodiversity. This is largely due to a belief that these areas are less vulnerable to threatening processes and less rich in threatened biodiversity, thereby having lower conservation value¹². For example, recent analyses of vertebrate taxa²⁷ found that areas with low human impact host fewer restricted-range species than it would be expected by chance. However, these species might have lost part of their original distribution as a consequence of rapid loss of wilderness areas¹⁰. Our research shows that, today, many wilderness areas are critical in reducing extinction risk of terrestrial biodiversity. These areas are important because they host highly unique biological communities and/or represent the majority of remaining natural habitats for biological communities that have suffered high levels of habitat loss elsewhere. Alarming, these invaluable areas are not better protected than average. Our findings point to the need for a targeted retention of the remaining areas of wilderness to be coupled with efforts that aim at protecting and restoring important habitats in degraded environments²⁸. We believe it is vital that these two aims are viewed as highly complementary and non-substitutable components of a truly integrated approach to promoting the overall persistence of the biodiversity of our planet. A strategic expansion of the global protected-area estate is needed to preserve the irreplaceable wilderness areas that are most at risk, alongside national land-use legislation and the enforcement of business standards for reducing industrial footprints in intact ecosystems^{8,12}. In addition, regions that have already lost the largest part of their wilderness (such as the Indomalayan and Afrotropical realms) require conservation strategies that focus on the restoration of ecosystem integrity²⁹. The value of wilderness in the international biodiversity agenda can be no longer understated if nations are truly committed to achieving the Sustainable Development Goals³⁰.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1567-7>.

Received: 12 December 2018; Accepted: 20 August 2019;

Published online 18 September 2019.

1. Johnson, C. N. et al. Biodiversity losses and conservation responses in the Anthropocene. *Science* **356**, 270–275 (2017).

2. Steffen, W. et al. Planetary boundaries: guiding human development on a changing planet. *Science* **347**, 1259855 (2015).
3. Cardinale, B. J. et al. Biodiversity loss and its impact on humanity. *Nature* **486**, 59–67 (2012).
4. Ceballos, G., Ehrlich, P. R. & Dirzo, R. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proc. Natl Acad. Sci. USA* **114**, E6089–E6096 (2017).
5. Watson, J. E. M. et al. Persistent disparities between recent rates of habitat conversion and protection and implications for future global conservation targets. *Conserv. Lett.* **9**, 413–421 (2016).
6. Venter, O. et al. Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nat. Commun.* **7**, 12558 (2016).
7. Newbold, T. et al. Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment. *Science* **353**, 288–291 (2016).
8. Di Marco, M., Venter, O., Possingham, H. P. & Watson, J. E. M. Changes in human footprint drive changes in species extinction risk. *Nat. Commun.* **9**, 4621 (2018).
9. Mittermeier, R. A. et al. Wilderness and biodiversity conservation. *Proc. Natl Acad. Sci. USA* **100**, 10309–10313 (2003).
10. Watson, J. E. M. et al. Catastrophic declines in wilderness areas undermine global environment targets. *Curr. Biol.* **26**, 2929–2934 (2016).
11. Allan, J. R., Venter, O. & Watson, J. E. M. Temporally inter-comparable maps of terrestrial wilderness and the last of the wild. *Sci. Data* **4**, 170187 (2017).
12. Watson, J. E. M. et al. Protect the last of the wild. *Nature* **563**, 27–30 (2018).
13. Clark, J. A. & May, R. M. Taxonomic bias in conservation research. *Science* **297**, 191–192 (2002).
14. Di Marco, M. et al. Changing trends and persisting biases in three decades of conservation science. *Glob. Ecol. Conserv.* **10**, 32–42 (2017).
15. Chapman, A. D. *Numbers of Living Species in Australia and the World*, <http://www.environment.gov.au/biodiversity/abrs/publications/other/species-numbers/2009/06-references.html> (Report for the Australian Biological Resources Study, Canberra, 2009).
16. Bar-On, Y. M., Phillips, R. & Milo, R. The biomass distribution on Earth. *Proc. Natl Acad. Sci. USA* **115**, 6506–6511 (2018).
17. Hoskins, A. J. et al. Supporting global biodiversity assessment through high-resolution macroecological modelling: methodological underpinnings of the BILBI framework. Preprint at <https://www.biorxiv.org/content/10.1101/309377v3> (2019).
18. Ferrer, S., Manion, G., Elith, J. & Richardson, K. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Divers. Distrib.* **13**, 252–264 (2007).
19. Ferrer, S. et al. Mapping more of terrestrial biodiversity for global conservation assessment. *Bioscience* **54**, 1101–1109 (2004).
20. Allnutt, T. F. et al. A method for quantifying biodiversity loss and its application to a 50-year record of deforestation across Madagascar. *Conserv. Lett.* **1**, 173–181 (2008).
21. Di Marco, M. et al. Projecting impacts of global climate and land-use scenarios on plant biodiversity using compositional-turnover modelling. *Glob. Chang. Biol.* **25**, 2763–2778 (2019).
22. Olson, D. M. et al. Terrestrial ecoregions of the world: a new map of life on earth. *Bioscience* **51**, 933–938 (2001).
23. Sodhi, N. S., Koh, L. P., Brook, B. W. & Ng, P. K. L. Southeast Asian biodiversity: an impending disaster. *Trends Ecol. Evol.* **19**, 654–660 (2004).
24. IUCN & UNEP–WCMC. The World Database on Protected Areas (WDPA) version July/2018, www.protectedplanet.net (UNEP–WCMC, Cambridge, 2018).
25. Potapov, P. et al. The last frontiers of wilderness: tracking loss of intact forest landscapes from 2000 to 2013. *Sci. Adv.* **3**, e1600821 (2017).
26. Gorenflo, L. J., Romaine, S., Mittermeier, R. A. & Walker-Painemilla, K. Co-occurrence of linguistic and biological diversity in biodiversity hotspots and high biodiversity wilderness areas. *Proc. Natl Acad. Sci. USA* **109**, 8032–8037 (2012).
27. Pimm, S. L., Jenkins, C. N. & Li, B. V. How to protect half of Earth to ensure it protects sufficient biodiversity. *Sci. Adv.* **4**, eaat2616 (2018).
28. CBD. Strategic Plan for Biodiversity 2011–2020 (CBD, 2010).
29. Mappin, B. et al. Restoration priorities to achieve the global protected area target. *Conserv. Lett.* **12**, e12646 (2019).
30. United Nations General Assembly. Transforming our World: the 2030 Agenda for Sustainable Development, A/RES/70/1 (United Nations General Assembly, 2015).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

METHODS

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

We used compositional-turnover models for vascular plant and invertebrate communities that were previously generated¹⁷ using the 'Biogeographic Infrastructure for Large-scaled Biodiversity Indicators' (BILBI), a global biodiversity modelling infrastructure that has recently been applied to projecting biodiversity trends under future scenarios of socio-economic development^{21,31,32}. This infrastructure relies on a GDM approach to predict spatial turnover in species composition between any pair of 30-arc-second grid cells across the terrestrial surface of the planet (about 1 km² at the equator). The infrastructure uses a specially modified form of GDM, which corrects for biases that are introduced into predictions when models are fitted to incomplete survey inventories. This is achieved by replacing the response variable that is normally used in GDM fitting (compositional dissimilarity between pairs of sites) with the probability that a pair of observations drawn randomly from two sites refer to the same or different species. This modelled probability is then back-transformed to a measure of proportional dissimilarity in species composition between communities. This modification of the standard GDM approach minimizes the risk that incompleteness and biases in survey inventories result in inflated estimates of turnover.

In the BILBI infrastructure, separate GDMs were built for each of the 61 biome–realm combinations of the terrestrial globe²², with models fitted separately for invertebrates and plants. Each model was fitted to species–location records derived from the Global Biodiversity Information Facility (GBIF), as previously detailed¹⁷. A total of 132,761 species of invertebrates (with 13,244,784 location records) and 254,145 species of vascular plants (with 52,489,096 location records) were used globally. The selection of a reduced subset of GBIF records followed both an extensive data-cleaning and name-matching process, and the selection of plant and invertebrate taxa for which there were consistent collection methodologies, consistent communities of practice and relatively complete coverage; this aimed at minimizing the number of 'single specimen' records. Although GBIF data present inherent limitations (especially in terms of the variation in sampling intensity for different parts of the globe), the enhancement to GDM modelling used in our study reduces the bias that is introduced by incomplete sampling^{17,33}. Comparing 'observation pairs' (rather than site pairs) in the BILBI modelling infrastructure ensured that variation in sampling intensity was effectively accounted for during model fitting, because the probability that two observations in two sites refer to the same species is independent of the number of other species observed. In doing so, our approach relies on the assumption that range-restricted species are less likely to be found within the dataset being sampled; species that are less likely to be sampled (owing to their natural rarity) will therefore increase our estimates of dissimilarity in the areas they do exist. Furthermore, by focusing on spatial patterns in a collective property of biodiversity (compositional turnover) rather than modelling distributions of individual species, the BILBI infrastructure is expected to achieve relatively robust extrapolation of patterns across poorly sampled regions—even when the species that occur in these regions have not been surveyed.

The proportional compositional dissimilarity between grid cells was predicted as a function of the following environmental variables¹⁷: minimum monthly temperature³⁴, maximum monthly temperature³⁴, maximum diurnal temperature range³⁴, annual precipitation³⁴, actual evaporation³⁴, potential evaporation³⁴, minimum monthly water deficit³⁴, maximum monthly water deficit³⁴, soil pH³⁵, soil clay proportion³⁵, soil silt proportion³⁵, soil bulk density³⁵, soil depth³⁵, ruggedness index³⁶ and the topographical wetness index³⁵. All temperature, evaporation and water-deficit surfaces were adjusted for the effects of topographical aspect and shading^{37,38}.

Measuring the condition of habitats. We estimated the current condition of habitats using land-use maps for the year 2015, which were derived from the latest update of the land-use harmonisation project³⁹ (LUH2). These maps represent the percentage coverage—for each 0.25° grid cell of the globe—of 12 classes of land use: forested land (primary or secondary), non-forested land (primary or secondary), managed pasture, rangeland, urban land, land under C₃ crops (annual, perennial or nitrogen-fixing) and land under C₄ crops (annual or perennial). Estimates of the proportional coverage for each land-use class were downscaled from the original 0.25° resolution to a resolution of 30 arc-seconds (approximately 1 km² at the equator) following a previously described approach³⁸, to match the scale of biological communities and wilderness areas. Our approach differed slightly from the previous approach³⁸ to accommodate the added computational complexity

of fitting to 12 land-use classes instead of 5. Our approach also used more-recent datasets^{40–42} during the fitting process (Supplementary Methods). Following recent analyses^{21,31,32}, values for the 12 LUH2 classes were combined into a cumulative habitat-condition score by multiplying each percentage land-use value by a coefficient that represents the proportional native-species richness (or 'α diversity') that is expected to be retained under each land-use class; this coefficient was derived from the 'Projecting Responses of Ecological Diversity In Changing Terrestrial Systems' (PREDICTS) database^{7,21,43,44}. The coefficients were estimated from a hierarchical mixed-effects model to assess how natural species richness responds to land-use change⁴³.

We also used the habitat-condition surface as the basis for a sensitivity analysis of the potential effect of habitat connectivity. Connectivity was calculated following a previous approach⁴⁵, assuming cellwise permeability as a function of relative habitat condition. Because this calculation multiplies the connectivity of a cell by its current condition, the resultant surface is—by definition—lower than that measured by condition alone.

Estimating biodiversity persistence and the risk of species extinctions. We estimated the proportion of species associated with each grid cell i expected to persist in the long term anywhere within their range (p_i). We followed a previous study²⁰ in using the species–area relationship to translate the ratio between the remaining area and the original (pre-degradation) area of habitat across similar ecological environments (relative to the biological community in a given cell i) into the proportion of persisting species (p_i). This value was derived as a function of the modelled similarity (s_{ij}) in species composition between the focal cell (i) and other grid cells (j) found in the same biome–realm, derived using the GDM approach described in 'Modelling compositional variation in biological communities', as well as the condition of habitat in each of those cells (c_j), expressed as

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

in which the numerator represents the condition-weighted area of habitat remaining across ecological environments similar to grid cell i (the remaining extent of the biological community that comprises species that were originally present in cell i) and the denominator represents the original area of similar ecological environments (the extent of that biological community if all habitats were intact). The parameter z is the coefficient of the species–area relationship, set to 0.25 as per previous studies^{18–20,31}. After estimating p_i for each grid cell, we derived values for extinction risk (e_i) that represent the proportion of species associated with each grid cell i expected to be lost from their range, as a simple complement of persistence

$$e_i = 1 - p_i \quad (2)$$

We made separate estimates of persistence and extinction for vascular plant communities and invertebrate communities, and then averaged the values across the two groups to report aggregated biodiversity results. It is important to clarify that this method (as for any other method built on species–area relationship theory) does not estimate the precise timing of extinction. Rather, it estimates the proportion of species that are expected to become extinct over the long term, as a consequence of the habitat conditions observed at the present time. We thus invoke the concept of 'species committed to extinction' (for an example of this, see ref. ⁴⁶) as those species that were originally present in an area and that are estimated to disappear from their entire range, given deterioration of habitat condition. Some of these extinctions might already have been realized at the time of assessment, whereas others are expected to be realized over longer time periods into the future (as an extinction debt) unless habitat condition improves.

Estimating the contribution of wilderness areas to biodiversity persistence. We represented the distribution of wilderness areas using a previous map of terrestrial wilderness¹¹, at a global resolution of 1 km². The distribution of wilderness was derived by identifying all areas that are free of human pressure and that cover a contiguous area of $\geq 10,000$ km². The estimate of human pressure was, in turn, derived from the Human Footprint map⁴⁷, which represents cumulative human pressure on the environment. As wilderness encompasses regions with very diverse biological characteristics in terms of species diversity, levels of endemism and spatial turnover in species composition, we quantified the role of wilderness in promoting biodiversity persistence across different locations and across taxa. We did this by estimating the extinction risk within wilderness communities versus that within non-wilderness communities. A 'wilderness community' is defined here as the set of species that are associated with a cell found inside a wilderness area; the extinction risk for this community is therefore calculated by making this the focal cell i in Eqs. (1) and (2). Extinction risk for each 'non-wilderness community' is calculated in a similar manner, by making a particular cell falling outside wilderness the focal cell i in Eqs. (1) and (2).

We assessed the relative contribution that habitat found within wilderness areas makes to the persistence of terrestrial biodiversity, both globally and within each biogeographical realm²². To do so, we started from the estimate of the proportion of species ($p_{i,t}$) associated with each grid cell i expected to persist considering any available habitat (inside and outside wilderness). We then repeated this calculation considering only habitat found inside wilderness grid cells. By re-running the BILBI infrastructure using this ‘filtered’ habitat-condition map, we estimated the proportion of species associated with each grid cell i expected to persist if wilderness were the only habitat remaining ($p_{i,w}$). By comparing this latter value (which is based on only wilderness habitat) to the former value (which is based on all habitat), we were able to measure the relative contribution ($p_{i,c}$) that wilderness areas make to the total persistence of biodiversity associated with each grid cell, as

$$p_{i,c} = \frac{p_{i,w}}{p_{i,t}} \quad (3)$$

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

We also estimated the potential reduction in biodiversity persistence (δ_p) that would result from the loss of habitat in any given wilderness grid cell to identify those areas in which the effect of habitat loss would be highest. This value was calculated from the slope of the species–area curve (using Eq. (1)) for the grid cell in question, as

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

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

We report the mean and maximum wilderness contribution values ($p_{i,c}$) observed across grid cells within in each biogeographical realm. In addition, we report the mean and maximum δ_p values observed across grid cells within each block of wilderness (defined as individual patches of contiguous wilderness areas). Separate analyses were run for vascular plant communities and invertebrate communities, and values were then averaged to report aggregated biodiversity results. **Measuring the protection level of wilderness areas with different biodiversity value.** We measured the relationship between δ_p values and protection status for each wilderness pixel of the globe, using the World Database on Protected Areas²⁴. Following a previous description⁴⁸, we excluded internationally designated sites that are not considered as protected areas, excluded ‘proposed’ sites and sites with an unknown status, represented sites without a defined shape as geodetic buffers of the appropriate area, and excluded marine-only sites as well as the marine portion of coastal sites.

We assessed the effect size of the difference in mean δ_p values across protected and non-protected wilderness grid cells in each realm using Cohen’s d statistic⁴⁹. We also evaluated the difference between the observed mean δ_p value in protected wilderness areas and the value associated with 1,000 random samples of wilderness grid cells, (each of the same size as the number of protected cells). We measured how many times the observed mean δ_p was higher than the random mean δ_p .

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

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

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

Code availability

R code for deriving estimates of compositional dissimilarity and the proportion of persisting species is available from ref. 17.

31. Kim, H. et al. A protocol for an intercomparison of biodiversity and ecosystem services models using harmonized land-use and climate scenarios. *Geosci. Model Dev.* **11**, 4537–4562 (2018).
32. Leclère, D. et al. Towards Pathways Bending the Curve of Terrestrial Biodiversity Trends Within the 21st Century (IIASA, 2018).
33. Ware, C. et al. Improving biodiversity surrogates for conservation assessment: a test of methods and the value of targeted biological surveys. *Divers. Distrib.* **24**, 1333–1346 (2018).
34. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
35. Hengl, T. et al. SoilGrids1km—global soil information based on automated mapping. *PLoS ONE* **9**, e105992 (2014).
36. Amatulli, G. et al. A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Sci. Data* **5**, 180040 (2018).
37. Ferrier, S., Harwood, T., Williams, K. J. & Dunlop, M. Using Generalised Dissimilarity Modelling to Assess Potential Impacts of Climate Change on Biodiversity Composition in Australia, and on the Representativeness of the National Reserve System (CSIRO Climate Adaption Flagship Working Paper Series 13E) (CSIRO, Canberra, 2012).
38. Hoskins, A. J. et al. Downscaling land-use data to provide global 30'' estimates of five land-use classes. *Ecol. Evol.* **6**, 3040–3055 (2016).
39. Hurtt, G. C. et al. Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Clim. Change* **109**, 117–161 (2011).
40. European Commission Joint Research Centre & Columbia University Center for International Earth Science Information Network. GHS Population Grid, derived from GPW4, Multitemporal (1975, 1990, 2000, 2015). http://data.europa.eu/89h/jrc-ghsl-ghs_pop_gpw4_globe_r2015a (2015).
41. Pesaresi, M. et al. GHS built-up grid, derived from Landsat, multitemporal (1975, 1990, 2000, 2014) <https://ec.europa.eu/jrc/en/publication/ghs-built-grid-derived-landsat-multitemporal-1975-1990-2000-2014-ir2017-v10> (2015).
42. DiMiceli, C. M. et al. Annual global automated MODIS vegetation continuous fields (MOD44B) at 250 m spatial resolution for data years beginning day 2000–2010 (2011).
43. Hudson, L. N. et al. The database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) project. *Ecol. Evol.* **7**, 145–188 (2017).
44. Hill, S. L. L. et al. Worldwide impacts of past and projected future land-use change on local species richness and the Biodiversity Intactness Index. Preprint at <https://www.biorxiv.org/content/10.1101/311787v1> (2018).
45. Drielsma, M., Ferrier, S. & Manion, G. A raster-based technique for analysing habitat configuration: the cost-benefit approach. *Ecol. Modell.* **202**, 324–332 (2007).
46. Chaudhary, A. & Mooers, A. Terrestrial vertebrate biodiversity loss under future global land use change scenarios. *Sustainability* **10**, 2764 (2018).
47. Venter, O. et al. Global terrestrial human footprint maps for 1993 and 2009. *Sci. Data* **3**, 160067 (2016).
48. Butchart, S. H. M. et al. Shortfalls and solutions for meeting national and global conservation area targets. *Conserv. Lett.* **8**, 329–337 (2015).
49. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. (Academic, 1988).
50. GRASS Development Team. Geographic resources analysis support system (GRASS GIS) software, version 7.2. <http://grass.osgeo.org> (2017).
51. QGIS Development Team. QGIS geographic information system, <http://qgis.osgeo.org> (2017).
52. R Core Team. R: A Language and Environment for Statistical Computing. <https://www.r-project.org/> (2018).

Acknowledgements This work was funded by Research Agreement no. 2017113325 between CSIRO and the University of Queensland. M.D.M. acknowledges support from the European Union’s Horizon 2020 research and innovation programme (Marie Skłodowska-Curie grant agreement no. 793212).

Author contributions M.D.M., S.F. and J.E.M.W. framed the study. M.D.M., T.D.H. and A.J.H. carried out the analyses. M.D.M., S.F., T.D.H., A.J.H. and J.E.M.W. discussed and interpreted the results. M.D.M., S.F. and J.E.M.W. wrote the manuscript with support from T.D.H. and A.J.H.

Competing interests The authors declare no competing interests.

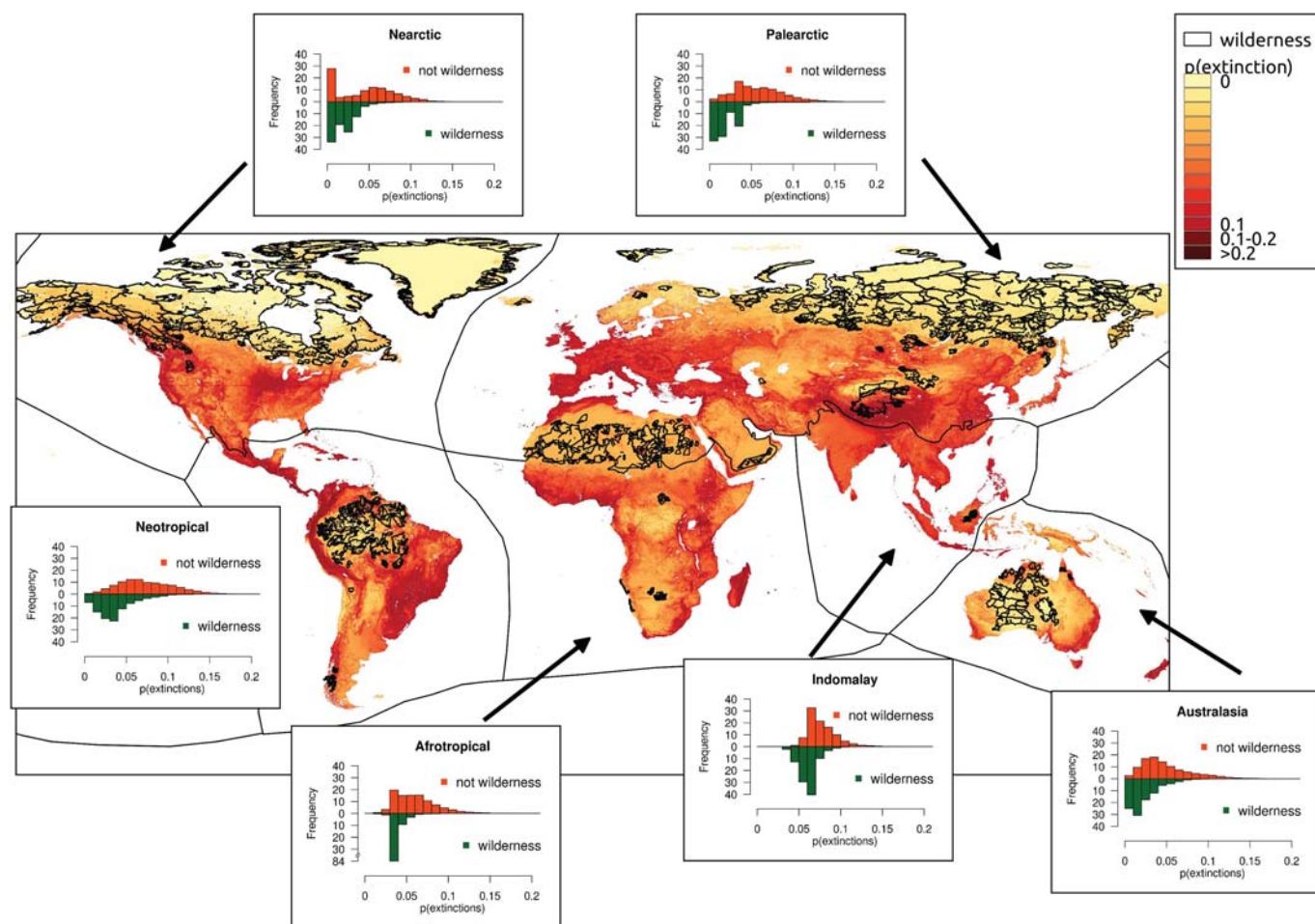
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-019-1567-7>.

Correspondence and requests for materials should be addressed to M.D.

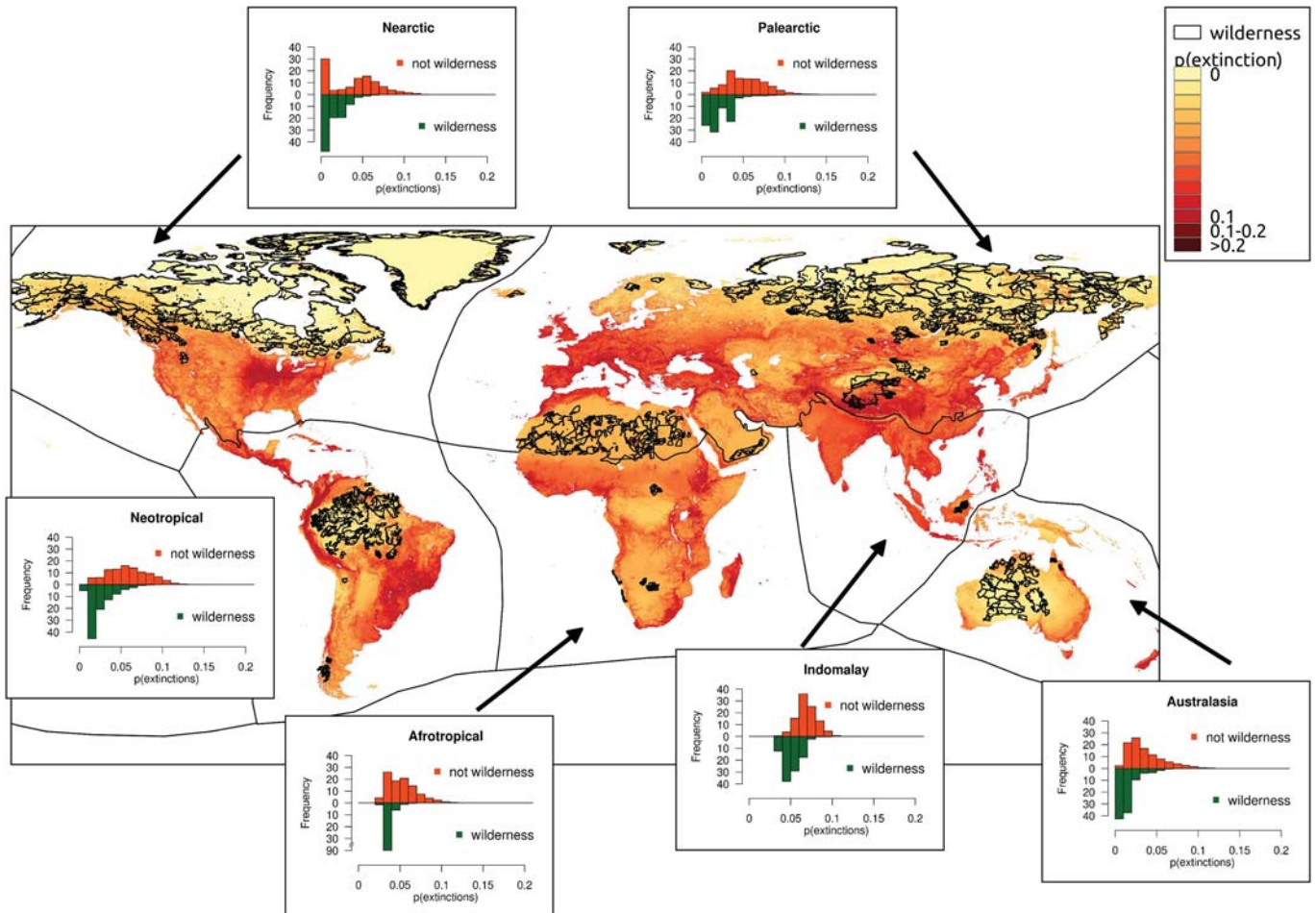
Peer review information *Nature* thanks Elizabeth Boakes, Samantha Hill and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



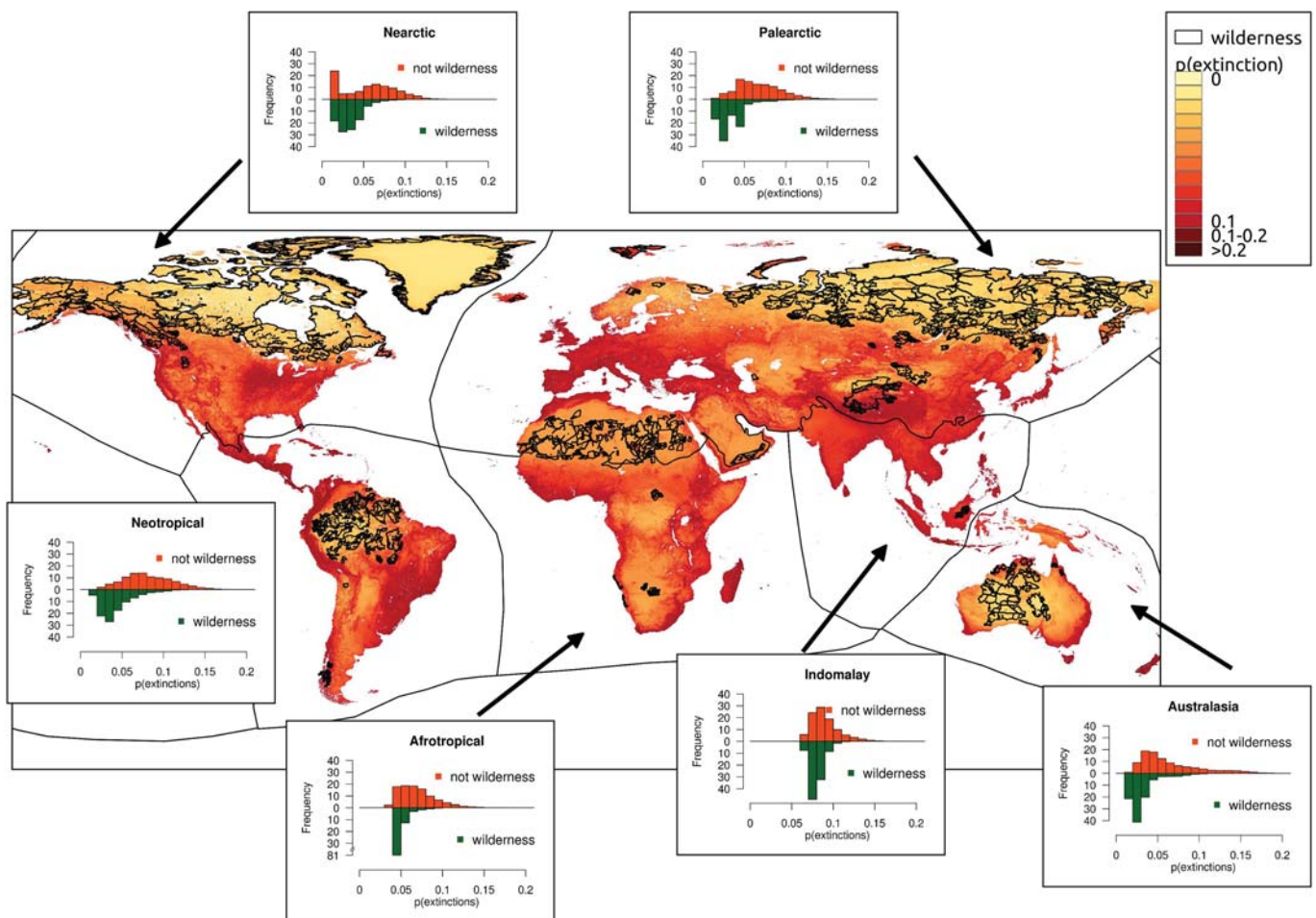
Extended Data Fig. 1 | Global-scale probabilities of species extinction for communities of vascular plants associated with each grid cell. The underlying map reports the estimated proportion of native species—

originally associated with a particular grid cell—that are expected to disappear from their distribution, owing to the current condition of the habitats in which they occur.



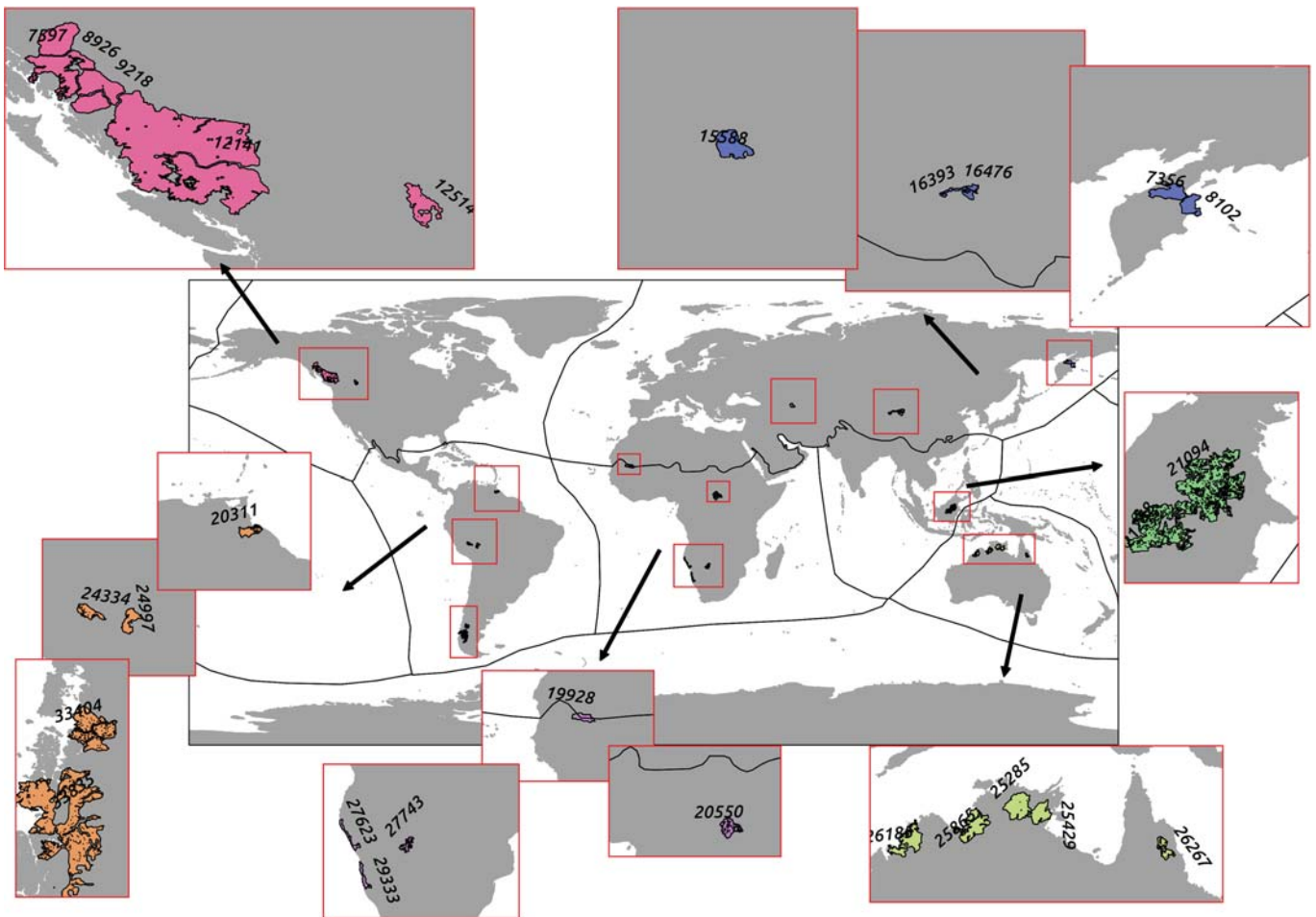
Extended Data Fig. 2 | Global-scale probabilities of species extinction for communities of invertebrates associated with each grid cell. The underlying map reports the estimated proportion of native species—

originally associated with a particular grid cell—that are expected to disappear from their distribution, owing to the current condition of the habitats in which they occur.

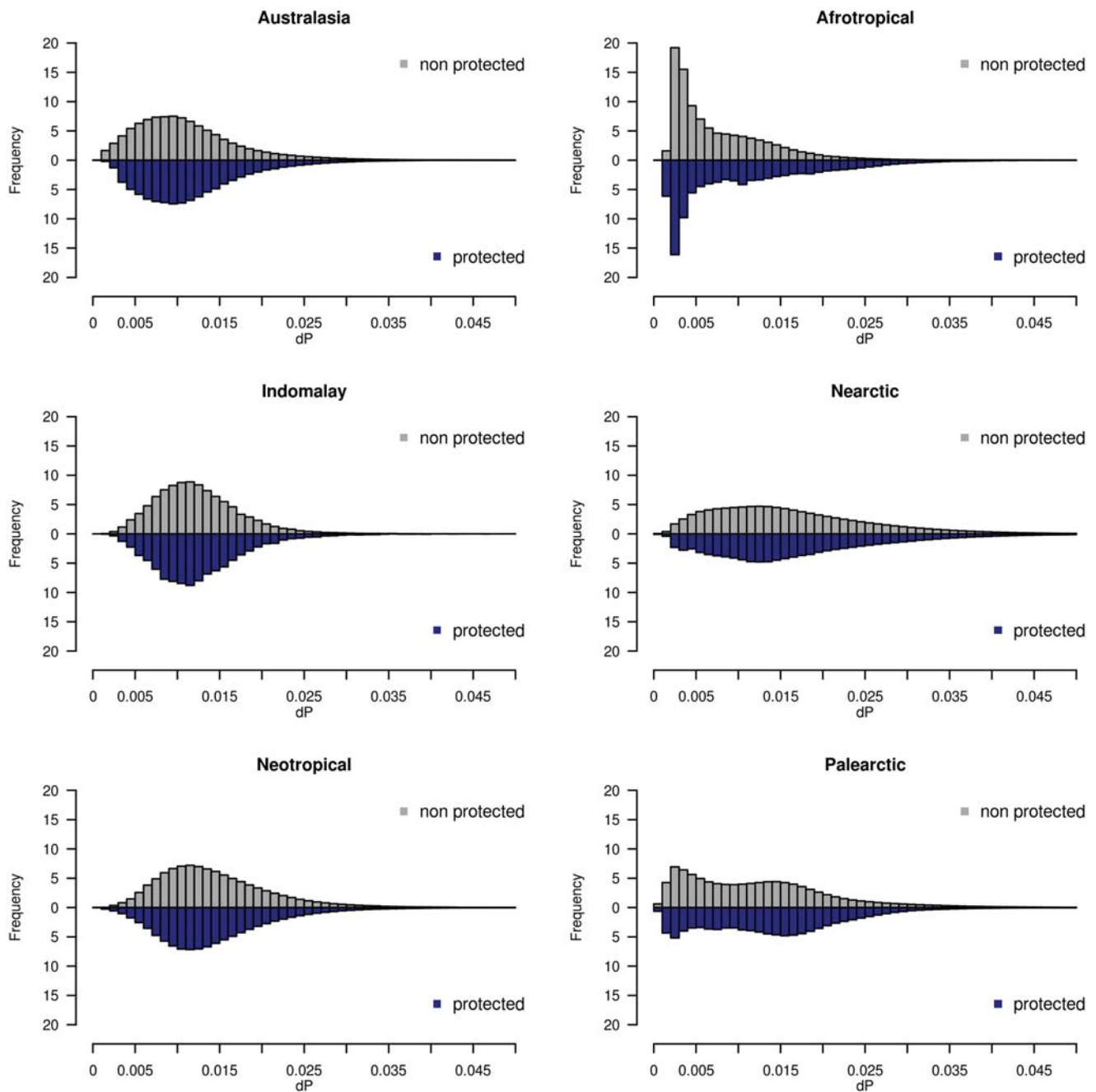


Extended Data Fig. 3 | Global-scale probabilities of species extinction for communities of invertebrates and vascular plants associated with each grid cell, accounting for habitat connectivity. The underlying map reports the estimated proportion of native species—originally associated

with a particular grid cell—that are expected to disappear from their distribution (owing to the current condition of the habitats in which they occur, as well as the level of connectivity between habitats).

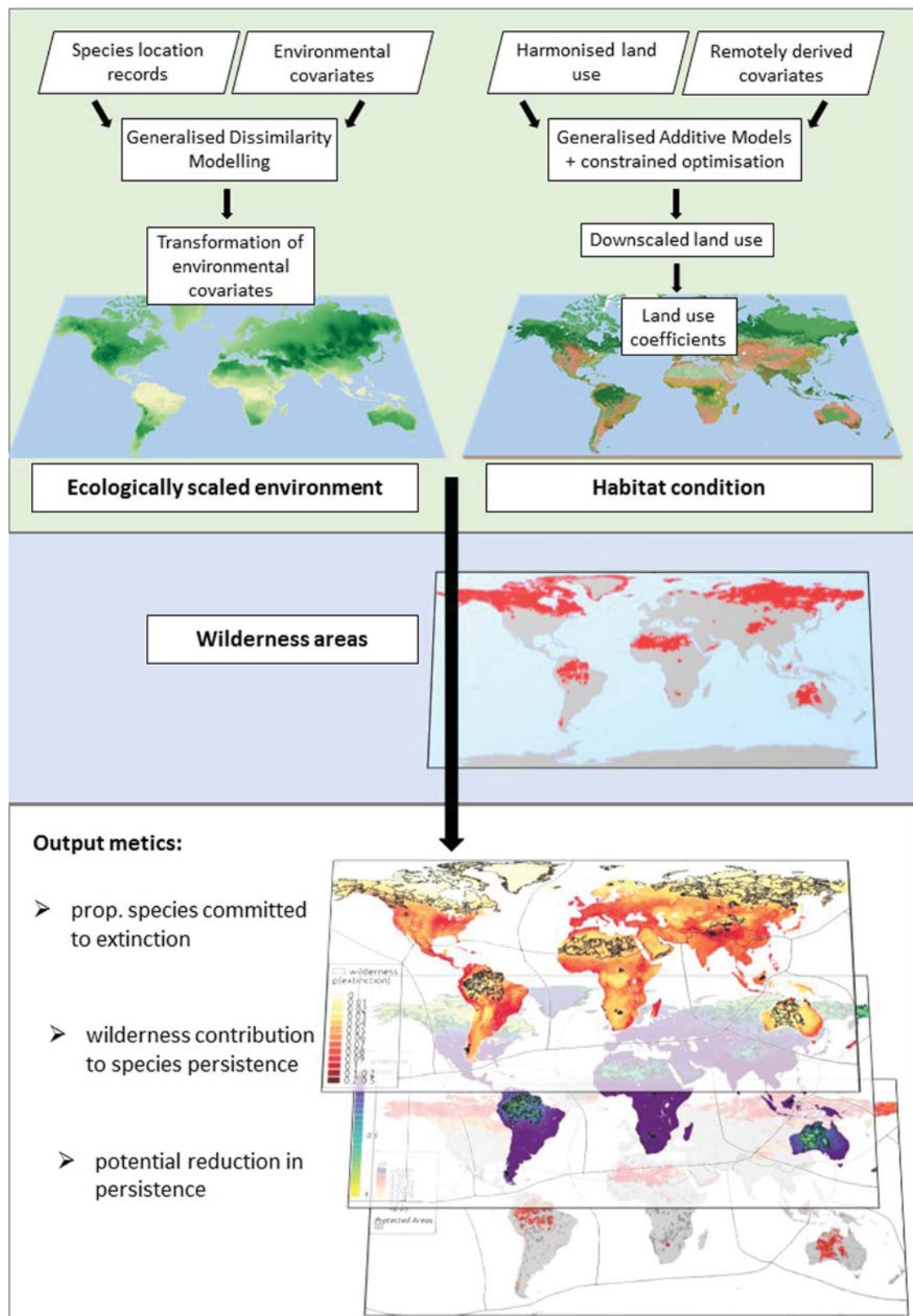


Extended Data Fig. 4 | Distribution of the top-five blocks of wilderness identified for each realm. Numbers in the map report the identifier codes for the block (corresponding to Supplementary Table 1).



Extended Data Fig. 5 | Frequency distribution of the contributions that individual wilderness grid cells make to the probability of persistence of invertebrate and vascular plant communities (δ_p). The histogram bars

represent the relative frequency distribution of the δ_p values for wilderness pixels inside (blue bars) and outside (grey bars) protected areas, in each biogeographical realm.



Extended Data Fig. 6 | Analytical framework used to estimate the probability of persistence of biological communities. The framework combines estimates of spatial turnover in species composition (from which ecologically scaled environments are derived) with estimates of habitat condition. The framework produces a spatially explicit (1 km²) estimate

of biodiversity persistence, from which a number of metrics are derived: the proportion of species committed to extinction, the contribution of wilderness areas to global species persistence, and the potential reduction in persistence in case of wilderness degradation.

Extended Data Table 1 | Mean extinction risk (with s.d. in parentheses) observed across communities of invertebrates and vascular plants in each biogeographical realm, inside and outside wilderness areas

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)

Extended Data Table 2 | Difference in the estimated reduction of global species persistence (δ_p) associated with the loss of a protected or non-protected wilderness pixel

Metric \ Realm	AA	AT	IM	NA	NT	PA
mean δ_p in protected areas	0.0119	0.0105	0.0126	0.0171	0.0144	0.0142
average random mean δ_p	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

The first row reports the mean δ_p values observed across all protected wilderness grid cells of a biogeographical realm. The second row reports the average across 1,000 mean δ_p values obtained by randomly selecting an equivalent number of wilderness cells. The third row reports the percentage of times in which the observed mean δ_p value was higher than the mean δ_p value from a random sample (out of 1,000 random samples). The fourth row reports the effect size (Cohen's d statistic) of the difference between δ_p values in protected and non-protected wilderness cells. AA, Australasia; AT, Afrotropical; IM, Indomalayan; NA, Nearctic; NT, Neotropical; PA, Palaearctic. The Oceanian and Antarctic realms were excluded from analyses.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

no software was used to collect data

Data analysis

Spatial analyses were performed in GRASSGIS v 7.2
Map layouts were produced in QGIS v 2.18.5
Statistical analyses were performed in R v 3.2.3
Custom code was used to estimate compositional dissimilarity (available from Code Availability statement)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

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

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This study builds on an innovative approach to map the spatial variation in species composition of biological communities, based on generalised dissimilarity modelling (ref. 17 in the main text). The method assigns any given location across Earth (represented as a 1 km grid cell) to a continuum of spatial turnover in biological composition. By utilising this approach, we predict the proportion of species within each biological community that are expected to persist, or go extinct, in the long term, given the current condition of habitats across the landscape where that community originally occurred. We express extinction risk as the proportion of species, originally associated with a given grid cell, that are expected to disappear (in the long term) from within their original range. We evaluate the contribution of terrestrial wilderness areas to the persistence of biodiversity.
Research sample	Species location records used in the generalised dissimilarity models derive from the Global Biodiversity Information Facility (GBIF; as described in ref 17 in the main text). A total of 132,761 species of invertebrates (with 13,244,784 location records) and 254,145 species of vascular plants (with 52,489,096 location records) were represented globally. The location of wilderness areas derive from the map of terrestrial wilderness of the world (ref 11 in the main text), representing all land grid cells (1 km resolution) classified as wilderness. Land-use coverage values (used to derive habitat conditions) derive from the LUH2 dataset (ref 39 in the main text).
Sampling strategy	No sample size calculation performed. For occurrence records, the complete GBIF database for vascular plants and invertebrates of the world was downloaded. For wilderness areas, all wilderness grid cells were considered. Likewise, land-use was evaluated for all terrestrial grid cells of the world.
Data collection	Occurrence data were downloaded from the the Global Biodiversity Information Facility (described in ref 17 in the main text). Wilderness areas were downloaded from Dryad, (ref 11 in the main text). Land-use data were downloaded from the Land Use Harmonisation website (ref 39 in the main text).
Timing and spatial scale	All data were retrieved from online repositories accessed during the period July/2017-November/2018. The generalised dissimilarity models, from which our estimates of compositional turnover derive, were built on terrain, climatic, and soil variables referring to the period 1960-2010 (described in full details in ref. 17 cited in the main text). The distribution of wilderness areas refers to year 2009. Land-use data refer to year 2015. The spatial scale of analysis has a global extent and a resolution of 30 arc-seconds (approx. 1km at the equator).
Data exclusions	For occurrence points, used to build GDM models: GBIF data were filtered to remove erroneous points falling outside land areas, as well as marine taxa (full details in ref. 17 in the main text). For spatial dataset, all analyses were limited to land areas above of the latitude of 60°S (i.e. excluding Antarctica).
Reproducibility	We created a spatial database in GRASSGIS, where all spatial maps were stored at the same spatial projection, extent, and resolution. All spatial data associations were performed in this environment, ensure consistency in all input maps used to estimate extinction risk for biological communities and the reproducibility of all calculations. We repeated our analysis after separating plant from invertebrate communities (ie different taxa), and after accounting for connectivity in natural habitats (ie different habitat conditions). In all occasions we verified that the average extinction risk for non-wilderness communities was twice as high as that of wilderness communities, confirming general validity and reproducibility of our findings.
Randomization	Not relevant. Our study analyses the global contribution of terrestrial wilderness areas to the persistence of biodiversity. All terrestrial wilderness areas are considered, as well as the modelled turnover in species composition from generalised dissimilarity models (GDMs). GDMs were built on all available occurrence records fro invertebrate and vascular plant species.
Blinding	Not relevant. All available data were used (species occurrences of the world's invertebrates and plants, land use in all terrestrial grid cells, and spatial location of the world's wilderness areas).
Did the study involve field work?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- | n/a | Involvement in the study |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |

Methods

- | n/a | Involvement in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |