**Supplementary Materials for**

**Protected area planning to conserve biodiversity in an uncertain world**

Richard Schuster, Rachel Buxton, Jeffrey O. Hanson, Allison D. Binley, Jeremy Pittman, Vivitskaia Tulloch, Frank A. La Sorte, Patrick R. Roehrdanz, Peter H. Verburg, Amanda D. Rodewald, Scott Wilson, Hugh P. Possingham, Joseph R. Bennett

Correspondence to: [richard.schuster@glel.carleton.ca](mailto:richard.schuster@glel.carleton.ca)

**This PDF file includes:**

Materials and Methods

Figs. S1 to S9

Tables S1 to S4

**Materials and Methods**

We used a multi-objective optimization approach that incorporated governance, land use and climate constraints to prioritize the conservation of 29,350 vertebrate species. All scenarios we investigated assumed the current global protected area portfolio is locked in. We further created representation targets for each species on the basis area of habitat maps. Targets for the unpartitioned habitat maps were set following standard practices for global gap analyses and prioritizations (*41-43*), except that instead of using the range sizes of species to set the targets, here we used the total extent of suitable habitat for each species following (*26*).

Biodiversity Data

Species AOH ranges were produced for 10,774 species of birds, 5,219 mammals, 4,462 reptiles and 6,254 amphibians with available IUCN range polygon data following the procedure outlined in (*27*). Species range polygons obtained from the IUCN Red List spatial data portal (*44*) and the Birdlife International spatial data zone (*45*) were first filtered for ‘extant’ range then rasterized to a global 1 km grid in the Eckert IV equal area projection. Individual species range rasters were then modified to only include land cover classes that match the habitat associations for each species. Habitat associations were obtained from the IUCN Red List species habitat classification scheme and were matched to ESA land cover classes for the year 2018 following the crosswalk table presented in (*46*). ESA land cover classification data was aggregated from its native 300 m resolution to match the global 1 km grid using a majority rule. Species ranges were additionally filtered so that only areas within a species accepted elevational range were included. Global elevation data derived from SRTM was obtained from WorldClim v. 2 (*47*). For bird species, seasonal range codes 1-3 (1=year-round; 2=breeding range; 3=non-breeding range) were processed individually and stored as separate range files where applicable.

Basic administrative delineations

National boundaries were derived from the Global Administrative Areas database (*48*). We obtained protected area boundaries from the World Database on Protected Areas (*49*). Following standard procedures for cleaning the protected area dataset (*41,50*), we (i) projected the data to an equal-area coordinate system (World Behrman), (ii) excluded reserves with unknown or proposed designations, (iii) excluded UNESCO Biosphere Reserves (*51*), (iv) buffered sites represented as point localities to their reported area, (v) dissolved boundaries to prevent issues with overlapping areas, and (vi) removed slivers (code available at https://github.com/jeffreyhanson/global-protected-areas). After the protected area data were modified as described above, we overlaid the protected area boundaries with a 10 x 10 km grid covering the Earth. These spatial data procedures were implemented using ArcMap (version 10.3.1) and python (version 2.7.8).

Governance risk

Conservation risk due to governance can affect the outcomes of strategies, and effective governance can promote the resilience of conservation in the face of sociopolitical and economic shocks. We used worldwide governance indicators from the World Bank (*15*) to capture these pressures. The indicators include six scaled measures: voice and accountability; political stability and absence of violence; government effectiveness; regulatory quality; rule of law; and control of corruption (see Table S4 for definitions). We chose these indicators because evidence suggests that they reliably predict protected area effectiveness (*52*) and state investment and efforts for biodiversity conservation (*51*). For each country, we used a mean of annual averages of all six measures (*53*) (Fig. S1).

Land use risk

We used a recently developed global land systems map produced by (*16*) to incorporate the risk of land use change. This map is based on a global land systems map for the year 2000 (*54*) at a 9.25 km2 spatial resolution but is refined based on recent land-cover and land use datasets to a spatial resolution of 1 km2. (*16*) further estimated the impact of land use and land use intensity on biodiversity, with data originating from the PREDICTS project (*53*). They first matched their land-systems classes to varying intensity levels for each land use type (for detailed conversion table, see ref (*54*)). This allowed (*16*) to calculate average biodiversity loss per land system (relative to an unimpacted baseline) by taking the mean model estimates of biodiversity loss per land use intensity class from previous work (*55*). The result gives average relative biodiversity gain or loss per land-system class. Here, we used their modelled mean estimates (following (*56*)) of relative percent biodiversity change for each land-system class for species abundance as a measure of the land use pressure (Fig. S2).

Climate risk

Velocity of climate change in an instantaneous measurement of how projected temperature increases translate to horizontal velocity on the landscape (*17*). It is an integration of both the rate of change in average climate and landscape properties that govern how bands of similar temperature redistribute spatially as climate changes. For example, in a region with high topographic diversity, a species may be able to track its climatic niche through relatively small dispersal distances (e.g. 10s or 100s of meters) upslope or downslope. By contrast, keeping pace with preferred climate under the same magnitude of temperature rise in the plains may require much larger dispersal distances – 100s or 1000s of kilometers. The local velocity of climate change, although a purely physical property, has biological relevance when linked with individual species-level dispersal capacity. Velocity of historical climate change since the last glacial maximum has been suggested as a major driver of patterns of species endemism (e.g. (*57*)) and is commonly used as a biologically scaled metric of climate exposure (e.g. (*19*)). Velocity of future temperature change used here follows the method of (*17*) – and is essentially the ratio of the projected temporal rate of change (C/year) to the spatial rate of change (C/km). Projected temporal rate of change is based on the 20 year mean (2040-2060) projection for mean annual temperature from the HadGEM2-ES model (CMIP5) and the baseline (1960-1990) temperature available from Worldclim v1.4. Spatial rate of change was derived from 30 arc second elevation data and calculated with the ‘terrain’ function from the R ‘raster’ package.

We also explored an alternative measure of climate risk: exposure to extreme events. Anthropogenic climate change is affecting the frequency and duration of extreme heat events (*58,59*). Exposure to these events can adversely affect human populations (*60-62*) and natural systems (*10,63*). For species in natural systems, these events can further the decline and extirpation of populations, increasing the chances of extinction (*10*,*64*). Extreme heat events and extreme cold events can also promote the formation of novel ecosystems (*63*), generate enhanced selection pressures (*65,66*), and change the phenology of life history events (*67,68*). There are a number of climate indices that have been used to estimate the occurrence of these events (*69,70*). These indices are often context specific and there is little consensus on the most appropriate technique (*71*).

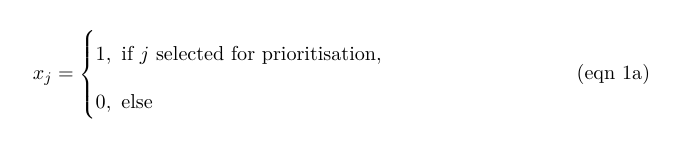
For this alternative measure, we estimated climatic risk based on the estimated trend in the annual proportion of days containing extreme heat events from 1979 to 2019 (*18*). Extreme heat events were estimated using hourly air temperature at 2 m above the surface and gridded at a 31 km (0.28125° at the equator) spatial resolution (*72*). The temperature data was acquired from the European Centre for Medium-Range Weather Forecasts (ECMWF) fifth generation atmospheric reanalysis of the global climate (ERA5) (*73,74*). The approach first extracted daily minimum and maximum temperature for each grid cell over the 41-year period. To reduce the influence of warming trends, the daily minimum and maximum temperature was then detrended across years for each day and grid cell using empirical mode decomposition (EMD) (*75,76*). The occurrence of extreme heat events was estimated using the following approach: The detrended minimum and maximum temperature data was treated as normally distributed across years for each day and grid cell. The probability density function for the detrended minimum and maximum temperature was then estimated using the mean and standard deviation calculated across years for each day and grid cell. Extreme heat events occurred when the probabilities for both minimum and maximum temperature on a given day and grid cell were within the 0.95-1.00 quartile of the probability density function. The trend in the annual proportion of days containing extreme heat events for each year was calculated for each grid cell using beta regression with a logit link function and an identity function in the precision model (*77,78*) (Fig. S7 – S9). See (*18*) for additional details.

Multi-objective optimization of pressure reduction

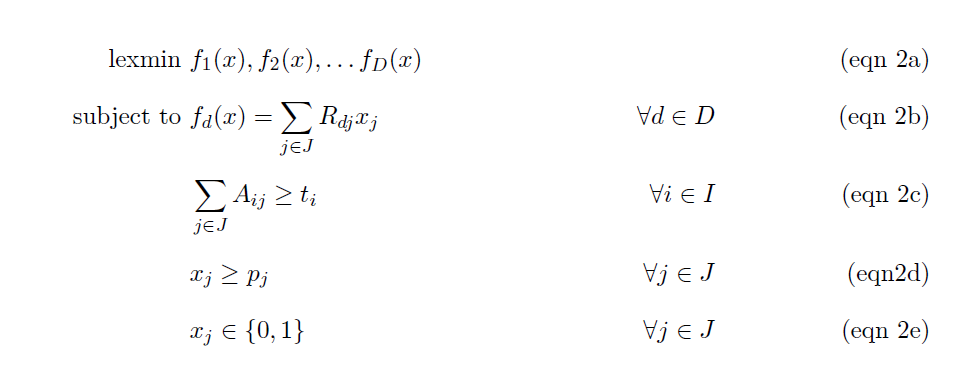
We processed all data described previously to a 10 x 10 km resolution and clipped data to the extent of land based on the global administrative areas database. We then developed an extension on the minimum set problem, which has the goal to identify a set of sites within a planning area that represents all conservation targets in the fewest number of sites (*22*). Instead of including a single objective in the problem formulation, we expanded it to include multiple objectives. Specifically, we used a hierarchical (lexicographic) approach that assigns a priority to each objective, and sequentially optimizes for the objectives in order of decreasing priority. At each step, it finds the best solution for the current objective, but only from among those that would not degrade the solution quality for higher-priority objectives. We considered up to three objectives in our prioritization scenarios, i) governance risk, ii) land use risk, and iii) climate risk. To compare different scenarios, we calculated solutions for each unique objective combination (n = 15), as well as one where we use a constant objective function as the null scenario, as the order of the hierarchy can influence the results.

In systematic conservation planning, conservation features describe the biodiversity units (e.g., species, communities, habitat types) that are used to inform protected area establishment. Planning units describe the candidate areas for protected area establishment (e.g., cadastral units). Each planning unit contains an amount of each feature (e.g., presence/absence, number of individuals). A prioritization describes a candidate set of planning units selected for protected establishment. Each feature has a representation target indicating the minimum amount of each feature that ideally should be held in the prioritization (e.g., 50 presences, 200 individuals). To minimize risk, we have a set of datasets describing the relative risk associated with selecting each planning unit for protected area establishment. Thus, we wish to identify a prioritization that meets the representation targets for all of the conservation features, with minimal risk.

Let I denote the set of conservation features (indexed by i), and J denote the set of planning units (indexed by j). To describe existing conservation efforts, let pj indicate (i.e., using zeros and ones) if each planning unit j ∈ J is already part of the global protected area system. To describe the spatial distribution of the features, let Aij denote (i.e., using zeros and ones) if each feature is present or absent from each planning unit. To ensure the features are adequately represented by the solution, let ti denote the conservation target for each feature i ∈ I. Next, let D denote the set of risk datasets (indexed by d). To describe the relative risk associated with each planning unit, let Rdj denote the risk for planning units j ∈ J according to risk datasets d ∈ D.

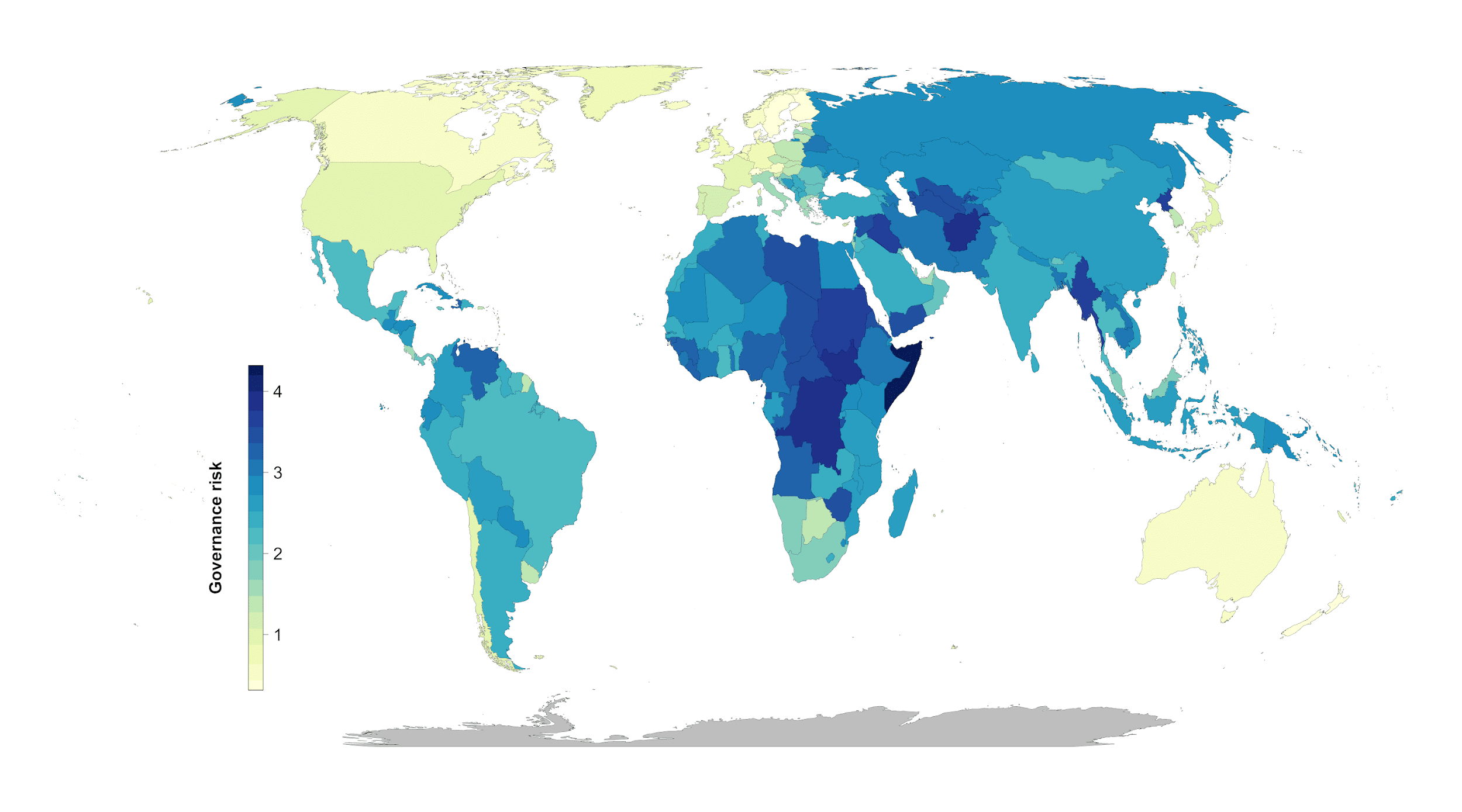
The problem contains the binary decision variables xj for planning units j ∈ J.

The reserve selection problem is formulated following:

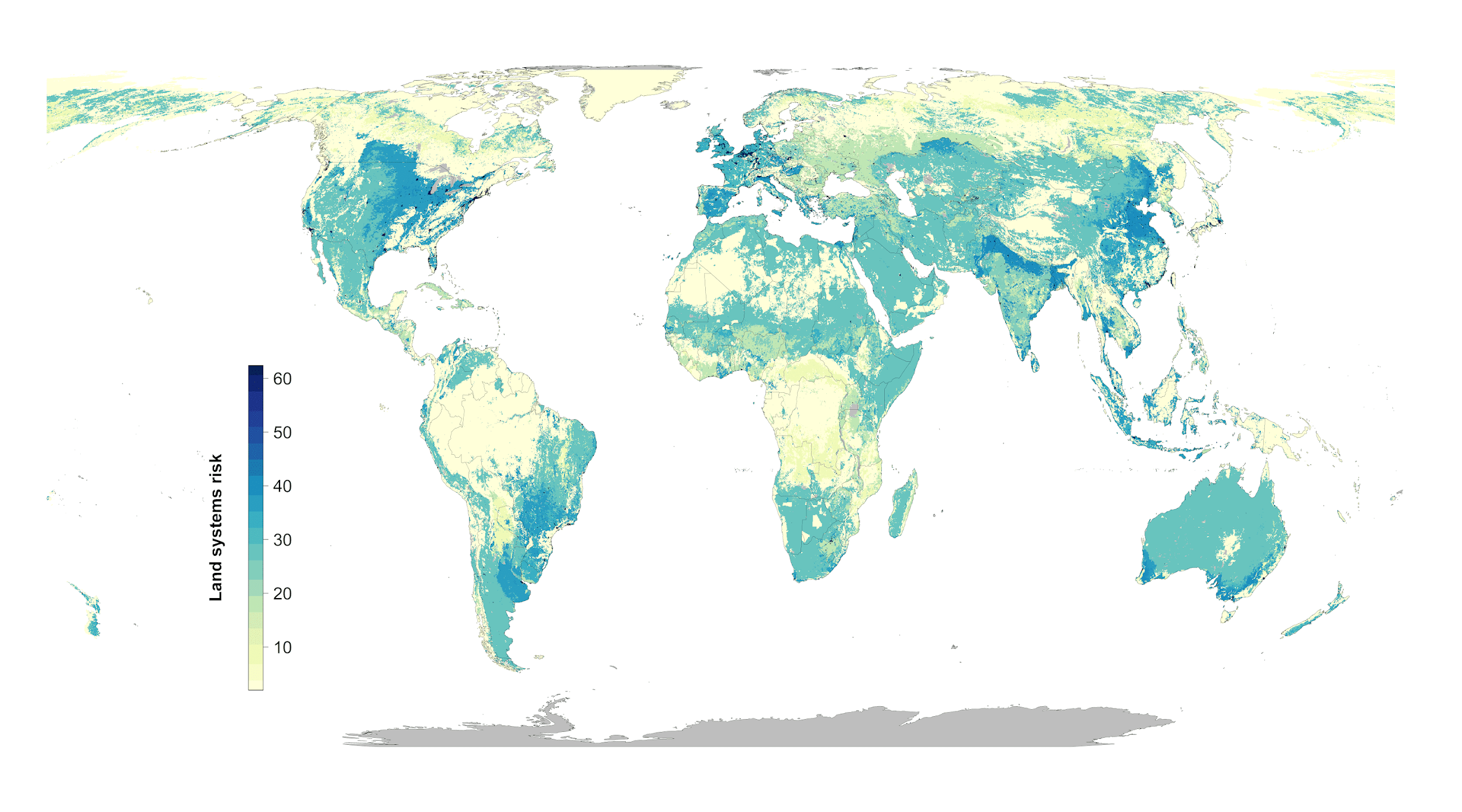


The objective function (eqn 2a) is to hierarchically (lexicographically) minimize multiple functions. Constraints (eqn 2b) define each of these functions as the total risk encompassed by selected planning units given each risk dataset. Constraints (eqn 2c) ensure that the representation targets (ti) are met for all features. Constraints (eqn 2d) ensure that the existing protected areas are selected in the solution. Finally, constraints (eqns 2e) ensure that the decision variables xj contain zeros or ones.

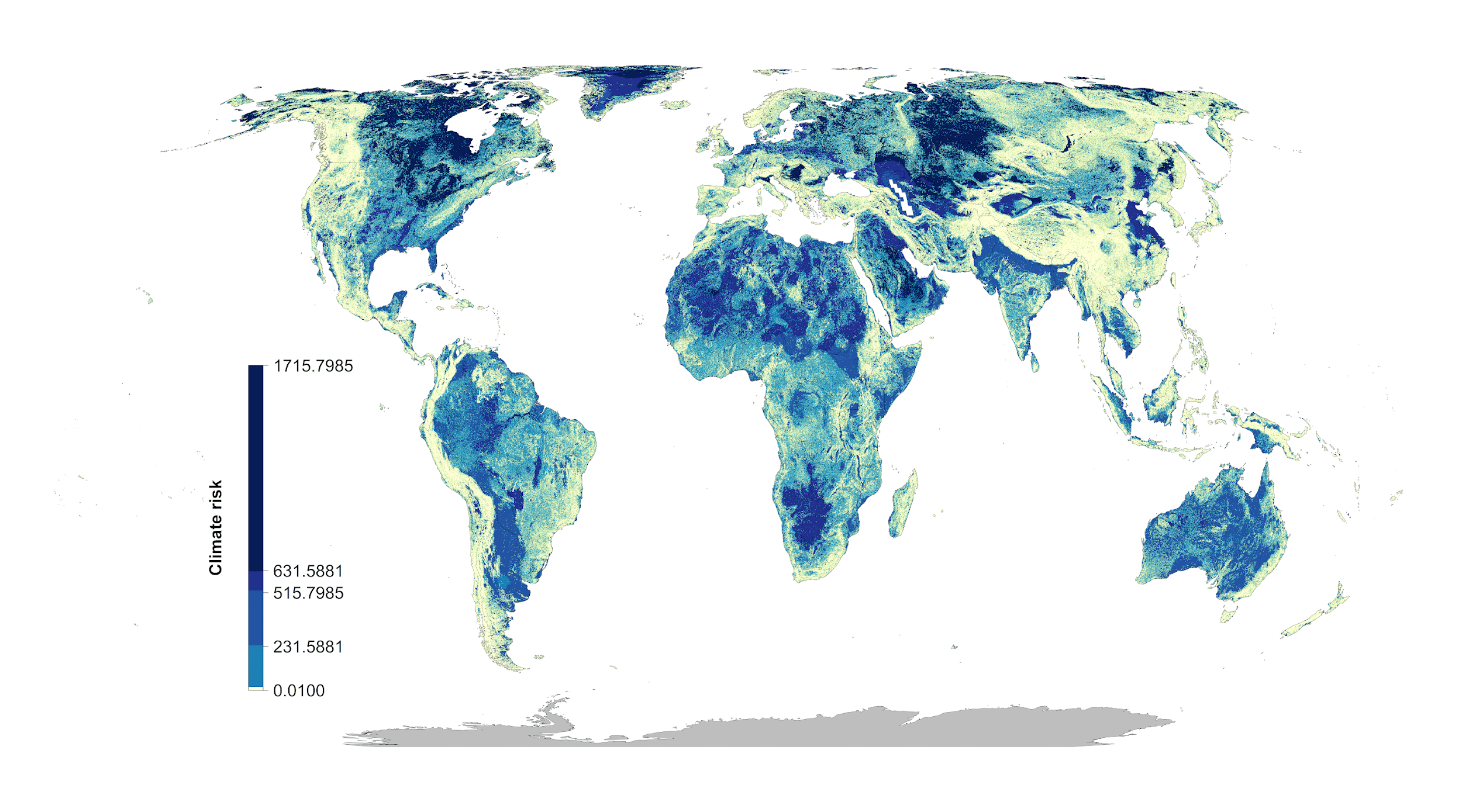
For all scenarios we locked in current protected areas. Following (*26*), we used flexible targets for suitable habitat based on species’ ranges. Species with less than 1,000 km2 of suitable habitat were assigned a 100% target (1,802 amphibians, 893 avian and 645 mammalian species), species with more than 250,000 km2 of suitable habitat were assigned a 10% target (712 amphibians, 4,518 avian and 1,868 mammalian species) and species with an intermediate amount of suitable habitat were assigned a target by log-linearly interpolating values between the previous two thresholds (2,683 amphibians, 5,190 avian and 2,557 mammalian species). Migratory bird species were assigned targets for each seasonal distribution separately). Additionally, to prevent species with very large suitable habitats from requiring excessively large amounts of area to be protected, the targets for species’ distributions larger than 10,000,000 km2 were capped at 1,000,000 km2. This upper limit affected only 206 (1%) species, and sensitivity analyses showed that it had little effect on our results. We acknowledge that these targets are arbitrary; however, they are more precise than previous targets based on species’ ranges (which can contain a large amount of unsuitable habitat), and accounts for the increased vulnerability of species with smaller range sizes (*79*), as well as the difficulty in conserving all habitat for species that occur over large areas.



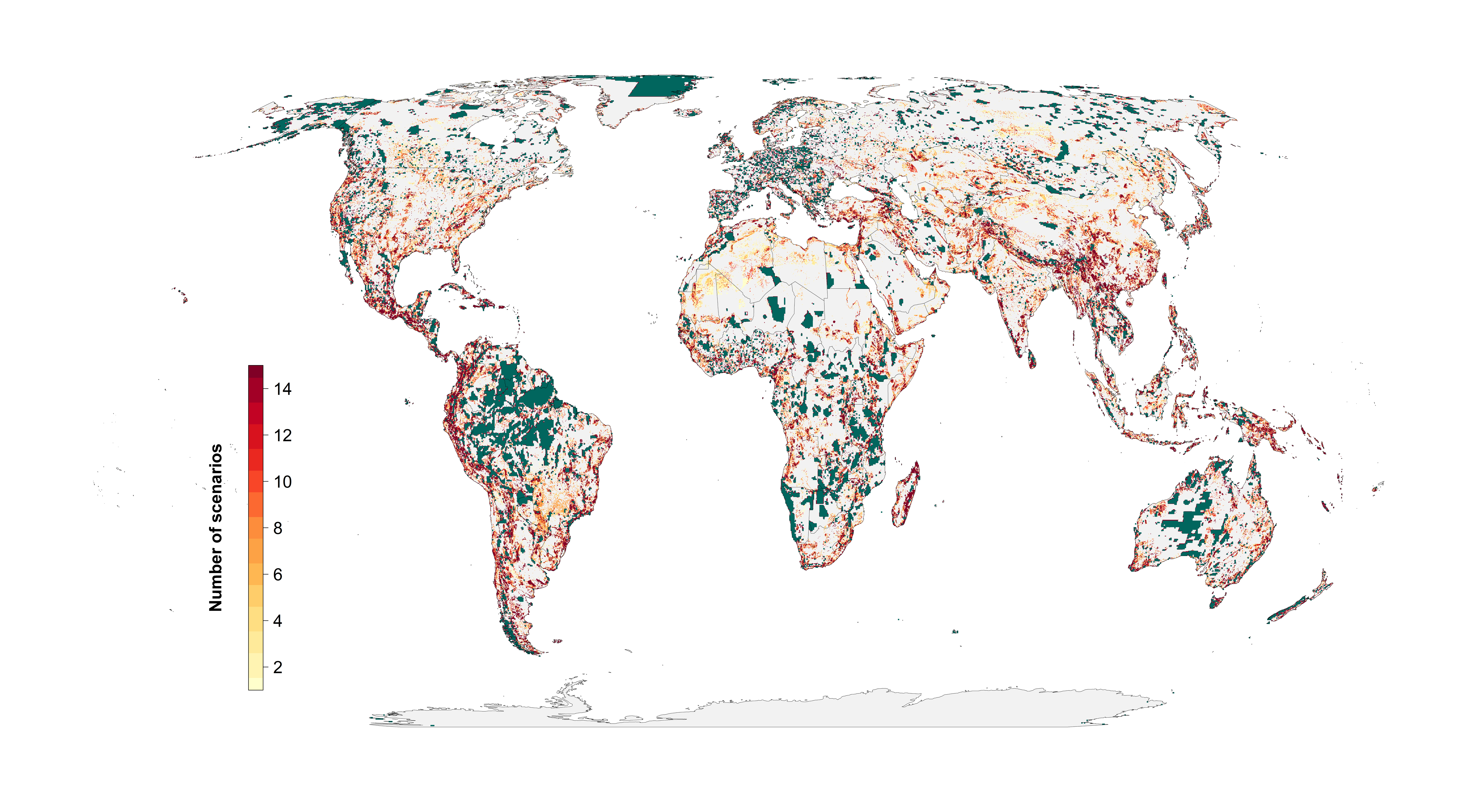
**Fig. S1. Governance risk (yellow = low, blue= high)**



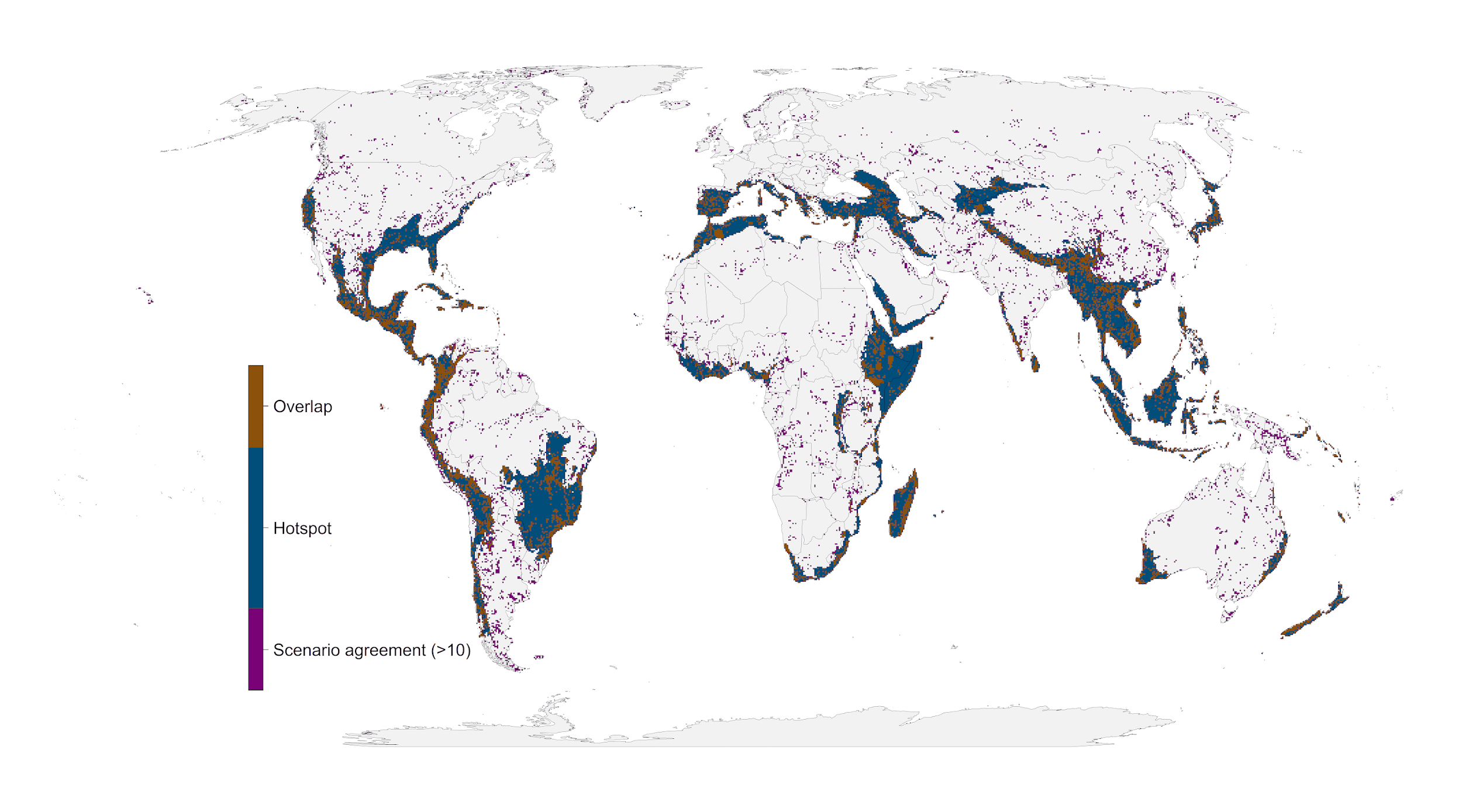
**Fig. S2. Land systems risk (yellow = low, blue= high)**



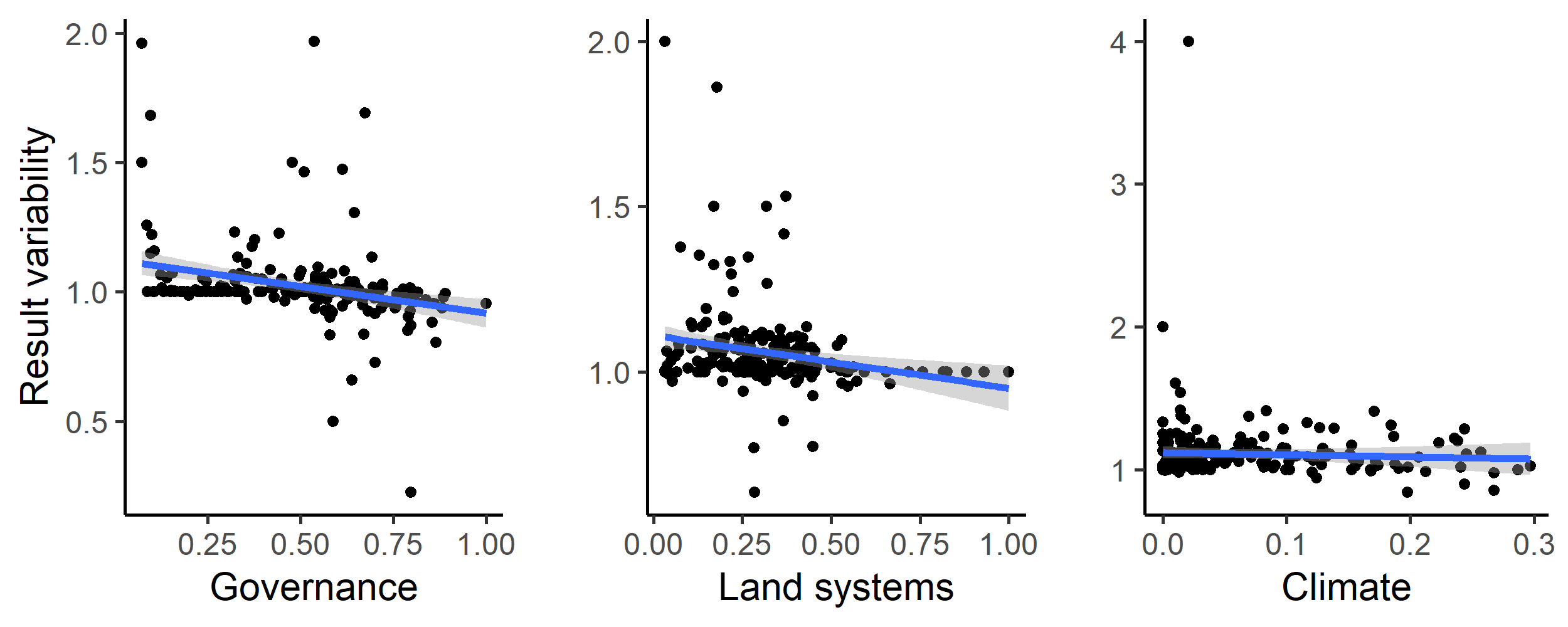
**Fig. S3. Climate risk (climate velocity) (yellow = low, blue= high)**



**Fig. S4: Scenario overlap. green = protected areas. Color gradient from yellow (one scenario) to red (15 scenarios) = overlap.**



**Fig. S5. Areas of high scenario overlap (>10 scenarios, green) compared to biodiversity hotspots (*28*) (blue).**



**Fig. S6: Influence of average country specific risk factors on the optimization outcomes compared between null scenario and the scenarios including one of the risk factors. Each data point represents the results for one country. The fitted blue lines and 95% confidence bands are from ordinary least-squares regression.**

Map

Description automatically generated

**Fig. S7. Alternative climate risk metric (extreme heat events) (yellow = low, blue= high)**

Chart, scatter chart

Description automatically generated

**Fig. S8: Alternative climate risk scenario “No regrets” areas that were identified as priority habitat for protection regardless of the risks included in our analysis.**

Map

Description automatically generated

**Fig. S9: Alternative climate risk scenarios percent country-level variation between the null scenario and the 15 scenarios including risk. Countries whose results are consistent across the 15 scenarios (e.g., Brazil) have low variation, while countries whose results are less consistent across the 15 scenarios (e.g., Sweden) have high variation. The kmeans method (*37*) was used to generate class intervals for visualization.**

|  |  |  |
| --- | --- | --- |
| **Scenario** | **Risk factors included** | **Global land area protected [%]** |
| **null** | - | 21.27 |
| **1** | G | 21.35 |
| **2** | L | 22.31 |
| **3** | C | 23.79 |
| **4** | G > L | 21.93 |
| **5** | L > G | 22.18 |
| **6** | G > C | 23.78 |
| **7** | C > G | 23.31 |
| **8** | L > C | 23.52 |
| **9** | C > L | 22.99 |
| **10** | G > L > C | 23.52 |
| **11** | G > C > L | 23 |
| **12** | L > G > C | 23.5 |
| **13** | L > C > G | 23.08 |
| **14** | C > G > L | 22.3 |
| **15** | C > L > G | 22.99 |

**Table S1.** Scenarios explored and global protection results. The risk factor order represents the order risk factors were included in the hierarchical prioritization. (G = governance, L = land use, C = Climate).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Afghanistan | Åland | Albania | Algeria |  |
| N | 15.95 | 57.14 | 38.46 | 10.62 |  |
| G | 14.95 | 85.71 | 35.66 | 7.71 |  |
| L | 17.03 | 85.71 | 43.71 | 10.32 |  |
| C | 19.25 | 57.14 | 46.15 | 13.69 |  |
| GL | 15.87 | 85.71 | 37.41 | 8.94 |  |
| LG | 16.55 | 100 | 38.11 | 11.59 |  |
| GC | 19.3 | 57.14 | 46.5 | 13.71 |  |
| CG | 17.89 | 71.43 | 39.16 | 12.74 |  |
| LC | 17.8 | 71.43 | 44.06 | 13.07 |  |
| CL | 19.52 | 57.14 | 40.56 | 13.36 |  |
| GLC | 17.8 | 57.14 | 43.71 | 13.15 |  |
| GCL | 19.44 | 57.14 | 41.96 | 13.38 |  |
| LGC | 17.81 | 57.14 | 44.06 | 13.05 |  |
| LCG | 16.58 | 85.71 | 38.11 | 12.36 |  |
| CGL | 17.52 | 85.71 | 43.36 | 12.4 |  |
| CLG | 19.52 | 57.14 | 40.56 | 13.36 |  |

**Table S2**. Country specific results for the 15 scenarios investigated. Numbers represent % of land area of a country selected (including existing protected areas).  
(As an example 5 countries included here, full list in csv. N = null, G = governance, L = land use, C = Climate)   
<https://drive.google.com/file/d/1eD4y4K8XG4nxnRL5fNtiTqzuqfIJ_DfB/view?usp=sharing>

|  |  |  |  |
| --- | --- | --- | --- |
| Country.Name | Country.Code | MeanIndex | SDIndex |
| Afghanistan | AFG | -1.65038 | 0.16074 |
| Albania | ALB | -0.28043 | 0.219515 |
| Algeria | DZA | -0.86838 | 0.121774 |
| American Samoa | ASM | 0.747997 | 0.127264 |
| Andorra | AND | 1.359029 | 0.04054 |
| Angola | AGO | -1.16429 | 0.217384 |
| Anguilla | AIA | 1.138708 | 0.225908 |
| Antigua and Barbuda | ATG | 0.687351 | 0.143042 |
| Argentina | ARG | -0.19472 | 0.196541 |
| Armenia | ARM | -0.29545 | 0.091655 |
| Aruba | ABW | 1.181311 | 0.090913 |
| Australia | AUS | 1.591282 | 0.033469 |
| Austria | AUT | 1.559385 | 0.080972 |
| Azerbaijan | AZE | -0.84662 | 0.123512 |
| Bahamas, The | BHS | 0.991142 | 0.212122 |
| Bahrain | BHR | 0.067606 | 0.189151 |
| Bangladesh | BGD | -0.8678 | 0.131258 |
| Barbados | BRB | 1.154432 | 0.145899 |

**Table S3. Governance risk score table (see csv)**

(As an example Afghanistan – Barbados are included below)<https://drive.google.com/file/d/1g_LePBfCbphXzTiCOXCzQtNLSSYoV6me/view?usp=sharing>

|  |  |
| --- | --- |
| **Indicator** | **Definition**  Source: World Bank, 2020 (<https://datacatalog.worldbank.org/dataset/worldwide-governance-indicators>) |
| Voice and accountability | “Voice and accountability captures perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media.” |
| Political stability and absence of violence | “Political Stability and Absence of Violence/Terrorism measures perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism.” |
| Government effectiveness | “Government effectiveness captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.” |
| Regulatory quality | “Regulatory quality captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development.” |
| Rule of law | “Rule of law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.” |
| Control of corruption | “Control of corruption captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests.” |
|  |  |

**Table S4**. Worldwide governance indicator definitions from the World Bank (15).