# Supplementary material

Creation of Species Rarity Layer (2018)

The Species Rarity Layer (2018) used in this research was created for the Global Safety Net (Dinerstein et al., 2020) to capture unprotected habitats hosting narrow range endemic and threatened species.

Six datasets were used to create the Species Rarity Layer (see SM Table 1); most of these data layers are generated using optimization approaches to conserve the maximum number of species in the smallest area possible. First, we overlaid each of these biodiversity data layers with all terrestrial protected areas (UNEP-WCMC and IUCN, 2018) to remove areas already set aside for conservation. To remove double counting, we subtracted any overlapping areas with previous datasets. (For example, all Alliance for Zero Extinction sites (AZE) are included within Key Biodiversity Areas (KBAs) and for this step-wise analysis we removed AZE sites from the KBA coverage). We ingested resulting layers into Google Earth Engine to derive remaining habitat in each layer using percent tree-cover maps (Hansen et al., 2013) in forested ecoregions (except boreal forests) and excluded globally significant patterns of human land use and populations (“anthromes”) in non-forested ecoregions (Ellis et al., 2010) (see (Dinerstein et al., 2017) for detailed methods]. We selected all non-overlapping unprotected areas within each of layers 1 through 4 and only the remaining habitat for layers 5 and 6 as contributions toward target 1. For layers 1 to 5 within “species rarity,” we added a 1-km buffer around all unprotected sites except for layer 6, rare plant species, as the size of a “rare plant pixel” was ~10,000 km2.

We partitioned two of the above datasets to calculate a median pixel values: IUCN range- size rarity raster (median = 0.006) (Hill et al., 2019) and small-range vertebrates raster (median = 24) (Pimm et al., 2018). For both datasets, only pixels greater than or equal to the median values were used. In the case of rare plant species, to be conservative, we excluded pixels containing only one to two rare plant species. The rationale here is that some of these are known from one to a few specimens. All raster data were converted to vector data (polygon) for further analysis.

Caveats and Sources of Error

*Fractional Analysis*

Our analysis is based on the Copernicus Global Land Service: Land cover (version 3) product at 100 m resolution (Buchhorn et al., 2020). The overall accuracy of these maps is 80.6% +/-0.4% at 95% confidence level. Class-specific accuracy varies from >85% for forest, grass, and permanent water (inland water bodies) to 65–85% for shrub, crop, and urban. Croplands are generally overestimated at the cost of shrub class (Tsendbazar et al., 2020). Therefore, grassland habitats might be underestimated in our analysis. In arid regions bare land can represent both natural habitat for species and developed land. We chose to be conservative in our estimate of remaining land and to focus on tropical regions where bare land is most predominantly representative of urban, over-grazed, and cultivated areas.

*Cost Assessment*

There has been an urgent need for empirical land cost data to plan and prioritize conservation needs (Coomes et al., 2018). To our knowledge this is the first study to estimate land costs for conservation based on real acquisition data from purchase and lease projects and provides a benchmark for funding needed in the next three years. However, several caveats exist that should be considered. Our estimates are at a coarse spatial scale limited by the amount of data we could compile and the accuracy of site locations. Location descriptions for many projects referenced the nearest town or large city, which limited our ability to incorporate more specific local spatial attributes (i.e. human footprint). The large effect of acquisition land size and acquisition type on cost estimates contributed strongly to the uncertainty around those estimates. The Monte Carlo simulation using different acquisition scenarios proved effective at capturing this variation, but more data from a diversity of funding mechanisms such as redesignations could improve estimates. To account for this, we suggest that our cost estimates are best interpreted as land valuations or ‘land rents’ based on empirical data, rather than actual purchase/lease costs. Management costs are also a major factor that should be considered when looking at the total price tag of protecting sites, but these costs depend heavily on site-level factors such as restoration needs, site security, and the type of conservation stewardship set up at the time of acquisition. We therefore chose to focus only on acquisition costs.

**SM Table 1. The six biodiversity datasets comprising the Species Rarity Layer and their terrestrial area.** First, terrestrial protected areas (UNEP-WCMC and IUCN, 2018) were overlaid to remove areas already set aside for conservation. Each dataset was then sequentially overlaid from 1–6, subtracting any overlapping areas with previous datasets to remove double-counting.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Biodiversity Dataset** | **Area (km2, before fractional analysis)** | **Reference** |
| 1 | Alliance for Zero Extinction Sites  | 616,306 | Birdlife International, 2018 (Birdlife International, 2018) |
| 2 | IUCN Range Rarity Sites | 161,282 | Hill et al., 2019 (Hill et al., 2019) |
| 3 | IUCN Threatened Species Sites | 34,631 | Hill et al., 2019 (Hill et al., 2019) |
| 4 | Key Biodiversity Areas | 1,757,798 | Birdlife International, 2018 (Birdlife International, 2018) |
| 5 | Small-range Vertebrate Sites | 279,540 | Pimm et al., 2018 (Pimm et al., 2018) |
| 6 | Rare Plant Species | 198,231 | Enquist et al., 2019 (Enquist et al., 2019) |

**SM Table 2.** **Distribution of Conservation Imperative sites by ecoregion.** (See Supplementary Spreadsheet)

**SM Table 3. Model selection table with AIC and R2 values.** GDP = per capita GDP, pop = population density

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **df** | **AICc** | **deltaAICc** | **R2** |
| Realm + Purchase Size + Purchase Type + GDP + pop | 11 | 2250.82 | 0.00 | 0.73 |
| Purchase Size + Purchase Type + GDP  | 12 | 2252.56 | 1.74 | 0.68 |
| Realm + Purchase Size + Purchase Type + GDP | 10 | 2256.59 | 5.78 | 0.72 |
| Realm + Purchase Size + Purchase Type | 9 | 2273.33 | 22.51 | 0.7 |
| Purchase Size + Purchase Type + GDP + pop | 6 | 2318.86 | 68.04 | 0.68 |
| Purchase Size + Purchase Type | 4 | 2327.95 | 77.13 | 0.65 |

**SM Table 4. Model estimates for the top candidate model using realm, purchase size, purchase type, per capita GDP, and population size.** Estimates are shown with Indomalaya as the reference level for realm and lease as the reference level for purchase type. 95% confidence interval values in bold do not cross zero.

|  |  |  |
| --- | --- | --- |
| **Predictor** | **Estimate** | **95% CI** |
| (Intercept) | 7.59 | **[6.33, 8.84]** |
| Realm [Afrotropic] | -0.47 | **[-0.73, -0.21]** |
| Realm [Australasia] | -1.57 | **[-2.45, -0.68]** |
| Realm [Nearctic] | -1.17 | **[-1.50, -0.83]** |
| Realm [Neotropic] | -0.58 | **[-0.79, -0.37]** |
| Realm [Palearctic] | -0.74 | **[-1.13, -0.35]** |
| Purchase size | -0.68 | **[-0.71, -0.64]** |
| Acquisition type [purchase] | 0.97 | **[0.66, 1.28]** |
| Per capita GDP | 0.18 | **[0.07, 0.28]** |
| Population size | 0.03 | **[0.02, 0.08]** |

****

**SM Figure 1. Locations of project cost data.** Darker reds indicate more projects.

**SM File.** **R Code for land cost model fitting and simulation**

A

B

**SM Figure 2. Probability distributions for the predicted mean cost per hectare and total land costs.** **A**) Shows the probability distribution of mean cost per hectare for each realm. Green indicates leases (10-year average) and orange indicates purchases. **B**) Shows the predicted total cost for land acquisition in each realm. Red vertical lines indicate the mean value and blue dashed lines indicate the 50% probability range. Note that the x axis is adjusted for each realm.