PNAS
www.pnas.org

# Supplementary Information for:

***Biodiversity impacts and conservation implications of urban land expansion projected by 2050***

Rohan D. Simkin*
Karen C. Seto
Robert I. McDonald
Walter Jetz

**Corresponding author:**
Rohan D. Simkin
Email: rohan.simkin@yale.edu

**This PDF file includes:**

- Supplementary text
- Figures S1 to S2
- Table S1
- Legends for Datasets S1 & S2
- SI References

**Other supplementary materials for this manuscript include the following:**

Dataset S1: species_appendix_table.xlsx

Dataset S2: cities_appendix_table.xlsx

## Supplementary Information Text

### Matching habitat preferences to land covers

The categorization of species' habitat preferences obtained from MOL follows the International Geosphere-Biosphere Program (IGBP) land cover classification system consisting of 17 classes including natural land-covers and human land-uses (1). Our methodology for refining species range maps (described in Methods) relies on matching these species' habitat preferences to the mapped land-use and land-cover (LULC) classes in the GLOBIO LULC maps. However, the GLOBIO LULC maps use a different LULC categorization system, so the first step in this process is to identify alignment between the two LULC categorization systems.

The GLOBIO model utilizes four broad categories for 'modified' land uses - *urban*, *crop*, *pasture* and *forestry*. Crop and pasture land use categories are further divided into *high-intensity* and *low-intensity* based on estimated rates of fertilizer application (2). Areas not occupied by these modified categories (i.e. 'natural' land covers) are classified using the categories defined in the European Space Agency's CCI Landcover dataset (3). The 2050 GLOBIO LULC map contains an additional category for 'secondary vegetation'. Secondary vegetation arises where modified land uses present in 2015 are abandoned in 2050 due to declining demand (2).

Wherever possible, IGBP LULC classes were matched with the GLOBIO LULC classes using text descriptions. For categories such as crop, urban and some forest types, the text descriptions provided in the product documentation describing the classification schemes enabled us to match LULC classes between the two datasets. To further strengthen our assessment of these matches, and to aid in matching LULC classes where text descriptions were ambiguous, we estimated spatial overlap between LULC classes. We estimate spatial overlap by comparing the 2015 GLOBIO LULC map with the 2015 MODIS landcover product (MCD12Q1), which categorized land covers based on the IGBP LULC classification (4). The percentage overlap for each combination of LULC categories was calculated as the percentage of each MODIS IGBP LULC class that intersects with a given GLOBIO LULC class and vice versa (Figure 1).

The results of the spatial overlap analysis are shown in Figure 1 and the final matchings used in our analysis are shown in Table 1. Categories of forest and crop land covers with similar text descriptors show strong spatial overlap. However, finding a definitive match for grassland, shrubland, wetland and savanna ecosystems, was often less clear (Figure 1). For these landcovers, matches were performed using the best combination of matching text descriptions, and spatial overlap. In many cases multiple GLOBIO LULC classes were matched to each IGBP LULC class. LULC classes were not matched where they showed strong spatial overlap but had inconsistent text descriptions, such as between grassland and crop land covers, for example.

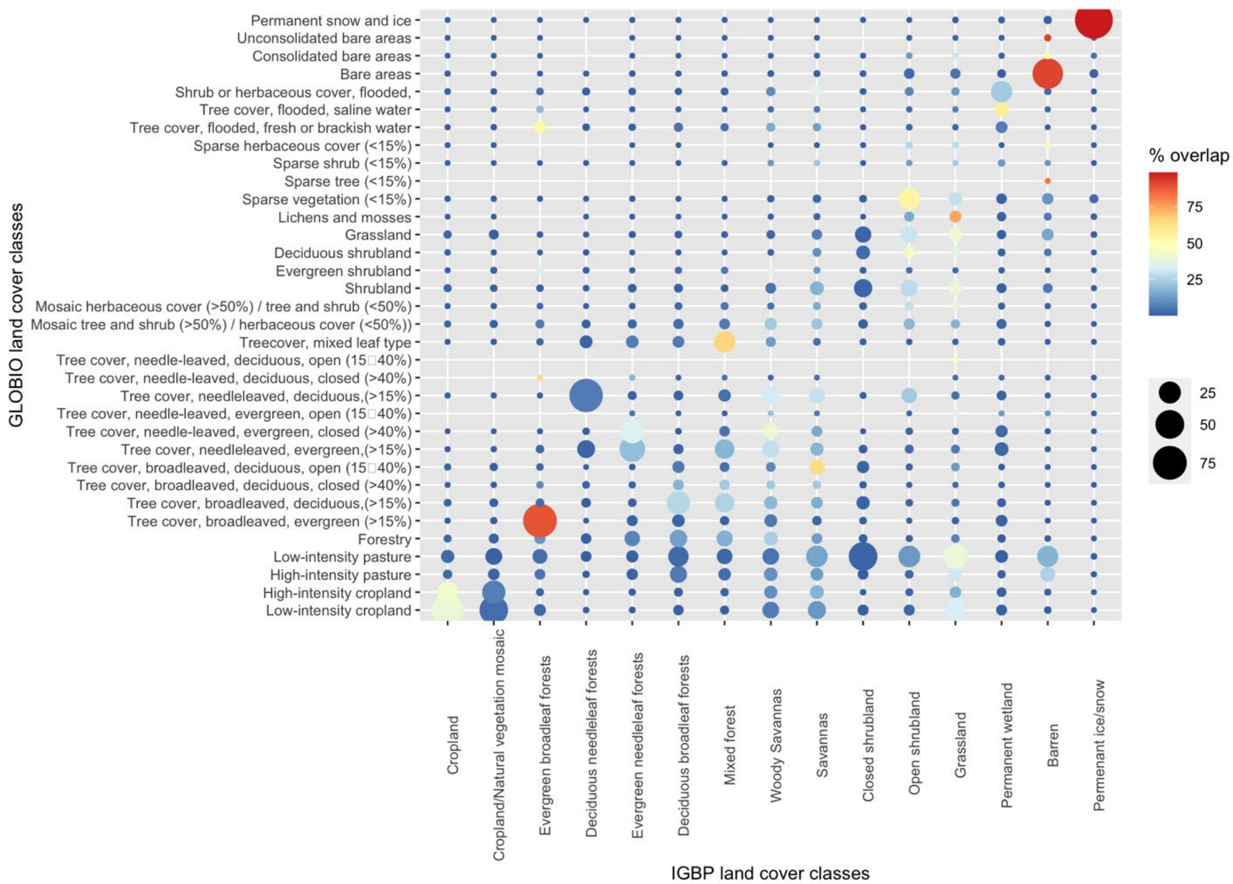


*Figure S1: Spatial overlap of land cover categories in 2015 from the MODIS IGBP land cover classification (MCD12Q1 LC_Type1) and the GLOBIO LULC map for 2015. The size of points indicates the percentage of an IGBP land cover category that intersects with a GLOBIO land cover category. Colors represent the percentage of each GLOBIO land cover category which intersects with each IGBP land cover category.*

*Table S1: Species habitat preferences, based on the IGBP land cover classification categories, matched to GLOBIO LULC categories.*

| **Habitat category (IGBP)** | | | **Matching GLOBIO land cover category** | |
|---|---|---|---|---|
| ***Value*** | ***Name*** | ***Description*** | ***Value*** | ***Description*** |
| 1 | Evergreen needleleaf forests | Dominated by evergreen conifer trees (canopy >2m). Tree cover >60%. | 70 | Tree cover, needle-leaved, evergreen, closed to open (>15%) |
| | | | 71 | Tree cover, needle-leaved, evergreen, closed (>40%) |
| 2 | Evergreen broadleaf forests | Dominated by evergreen broadleaf and palmate trees (canopy >2m). Tree cover >60%. | 50 | Tree cover, broadleaved, evergreen, closed to open (>15%) |
| 3 | Deciduous needleleaf forests | Dominated by deciduous needleleaf (larch) trees (canopy >2m). Tree cover >60%. | 80 | Tree cover, needle-leaved, deciduous, closed to open (>15%) |
| | | | 81 | Tree cover, needle-leaved, deciduous, closed (>40%) |
| | | | 82 | Tree cover, needle-leaved, deciduous, open (15-40%) |
| 4 | Deciduous broadleaf forests | Dominated by deciduous broadleaf trees (canopy >2m). Tree cover >60%. | 60 | Tree cover, broadleaved, deciduous, closed to open (>15%) |
| | | | 61 | Tree cover, broadleaved, deciduous, closed (>40%) |
| 5 | Mixed forest | Dominated by neither deciduous nor evergreen (40-60% of each) tree type (canopy >2m). Tree cover >60%. | 90 | Tree cover, mixed leaf type (broadleaved and needle-leaved) |
| 6 | Closed shrubland | Dominated by woody perennials (1-2m height) >60% cover. | 120 | Shrubland |
| 7 | Open shrubland | Dominated by woody perennials (1-2m height) 10-60% cover. | 150 | Sparse vegetation (tree, shrub, herbaceous cover) (<15%) |
| | | | 120 | Shrubland |
| | | | 130 | Grassland |
| | | | 122 | Deciduous shrubland |
| | | | 100 | Mosaic tree and shrub (>50%) / herbaceous cover (<50%) |
| | | | 110 | Mosaic herbaceous cover (>50%) / tree and shrub (<50%) |
| 8 | Woody Savanna | Tree cover 30-60% (canopy >2m). | 70 | Tree cover, needle-leaved, evergreen, closed to open (>15%) |
| | | | 80 | Tree cover, needle-leaved, deciduous, closed to open (>15%) |
| | | | 71 | Tree cover, needle-leaved, evergreen, closed (>40%) |
| | | | 60 | Tree cover, broadleaved, deciduous, closed to open (>15%) |
| | | | 100 | Mosaic tree and shrub (>50%) / herbaceous cover (<50%) |
| | | | 72 | Tree cover, needle-leaved, evergreen, open (15-40%) |
| | | | 121 | Evergreen shrubland |
| 9 | Savanna | Tree cover 10-30% (canopy >2m). | 80 | Tree cover, needle-leaved, deciduous, closed to open (>15%) |
| | | | 62 | Tree cover, broadleaved, deciduous, open (15-40%) |
| | | | 120 | Shrubland |

| Habitat category (IGBP) | | | Matching GLOBIO land cover category | |
|---|---|---|---|---|
| ***Value*** | ***Name*** | ***Description*** | ***Value*** | ***Description*** |
| | | | 70 | Tree cover, needle-leaved, evergreen, closed to open (>15%) |
| | | | 60 | Tree cover, broadleaved, deciduous, closed to open (>15%) |
| | | | 100 | Mosaic tree and shrub (>50%) / herbaceous cover (<50%) |
| | | | 110 | Mosaic herbaceous cover (>50%) / tree and shrub (<50%) |
| 10 | Grassland | Dominated by herbaceous annuals (<2m). | 130 | Grassland |
| | | | 120 | Shrubland |
| | | | 140 | Lichens and mosses |
| | | | 110 | Mosaic herbaceous cover (>50%) / tree and shrub (<50%) |
| 11 | Permanent wetland | Permanently inundated lands with 30-60% water cover and >10% vegetated cover. | 180 | Shrub or herbaceous cover, flooded, fresh/saline/brackish water |
| | | | 170 | Tree cover, flooded, saline water |
| | | | 160 | Tree cover, flooded, fresh or brackish water |
| 12 | Cropland | At least 60% of area is cultivated cropland. | 230 | Low-intensity cropland |
| | | | 231 | High-intensity cropland |
| 14 | Cropland/Natural vegetation mosaic | Mosaics of small-scale cultivation 40-60% with natural tree, shrub, or herbaceous vegetation. | 230 | Low-intensity cropland |
| | | | 231 | High-intensity cropland |
| 15 | Permanent ice/snow | At least 60% of area is covered by snow and ice for at least 10 months of the year. | 220 | Permanent snow and ice |
| 16 | Barren | At least 60% of area is non-vegetated barren (sand, rock, soil) areas with less than 10% vegetation. | 200 | Bare areas |
| | | | 201 | Consolidated bare areas |
| | | | 202 | Unconsolidated bare areas |
| | | | 153 | Sparse herbaceous cover (<15%) |
| | | | 151 | Sparse tree (<15%) |
| | | | 152 | Sparse shrub (<15%) |

***Unmatched land covers***

There is no equivalent IGBP LULC class for the GLOBIO LULC classes *pasture* or *forestry*. The following sections describe the procedure for matching species habitat preferences to these categories.

*Pasture*

In the GLOBIO 2015 LULC map, the distribution of land used for pasture is determined based on land suitability defined by the Gridded Livestock of the World (GLW) dataset – a global gridded dataset that models livestock density based on vegetation, topography, climate and demography (5). The use of the GLW dataset provides a means of identifying pasture/rangeland, which can be difficult to distinguish from grassland or other similar land covers using satellite imagery alone. The GLOBIO land allocation algorithm allocates all pasture land for 2015 onto areas identified as natural land-cover classes in 2015 by the 2015 CCI land-cover image.

We match species habitat preferences to high and low intensity pasture in the GLOBIO based on assumptions about the potential differences in habitat value of these categories. For high intensity pasture, we assume that the habitat present is highly modified and offers similar habitat value to cropped land. As such, we match species with a habitat preference for cropped land to high-intensity pasture in both 2015 and 2050. We recognize that low-intensity pasture is more likely to retain some characteristics of natural habitat. We reclassify low-intensity pasture pixels in the 2015 GLOBIO LULC map to the underlying CCI 2015 natural land-cover category and match species habitat preferences accordingly.

For low-intensity pasture pixels in the 2050 GLOBIO LULC map that transition from natural land covers or low-intensity pasture in 2015, we recognize that the low-intensity pasture categorization is likely to retain some habitat value. As such, we reclassify the 2050 low-intensity pasture pixels back to the 2015 natural land-cover classification and match species habitat preferences accordingly. Land that transitions from a modified land cover class in the 2015 GLOBIO LULC map (i.e. urban, crop, forestry or high-intensity pasture) to low-intensity pasture in 2050 is unlikely to retain significant habitat value. As such, only species with a habitat preference for cropped landscapes are matched to such land.

In taking this approach we aim to acknowledge the importance of the expansion of land managed for livestock as a driver of habitat declines, while also acknowledging the difficulty in accurately determining the ability of a species to persist within a land-cover category for which habitat values may vary from heavily modified to relatively intact.

*Forestry*

We take a conservative approach and assume that species do not persist in land allocated to forestry. We take this approach because forest operations often pose a threat to rare and threatened species. In doing so, we aim to put focus on these species while acknowledging that this approach fails to consider the habitat value of commercially managed forests. However, given the relatively small global footprint of forestry land in the GLOBIO model, we do not believe that this is a significant limitation.

## Secondary vegetation sensitivity analysis

The GLOBIO LULC maps used in our analysis include substantial areas of *secondary vegetation* in 2050. The GLOBIO LULC map identifies secondary vegetation where a modified landcover (i.e. urban, crop, pasture or forestry) is abandoned due to declining demand (2). The habitat value provided by secondary vegetation, and the capacity of species to utilize this habitat is a significant source of uncertainty. Whether a species is able to recolonize abandoned lands will depend on the species' particular habitat requirements, dispersal ability and distance to nearest source population, along with the prior land-use history.

Results presented in our analysis assume that species do not recolonize secondary habitat. We term this the 'no regain' scenario. To account for uncertainty in this assumption we test an additional 'regain' scenario, in which species are assumed to recolonize secondary habitat, if it is of a suitable type.

We match species habitat preferences to secondary vegetation based on assumptions about how species will utilize secondary vegetation under the regain and no regain scenarios. This approach is described below.

### ***Habitat preference matching***

*No regain scenario*

Under a no regain scenario, we assume that species do not recolonize secondary vegetation, but we assume that species present in the 2015 modified LULC category are able to persist as this land transitions to secondary vegetation. To do this, secondary vegetation pixels in the 2050 GLOBIO LULC map are reclassified back to their 2015 value (i.e. either urban, crop, pasture or forestry), and species habitat preferences matched accordingly.

*Regain scenario*

Under a regain scenario, species are able to utilize secondary vegetation where the recolonizing vegetation type matches the species' habitat preferences. Secondary vegetation is not broken down by vegetation type in the GLOBIO LULC maps. So, we determine the category of vegetation most likely to occur at each secondary vegetation pixel by utilizing the 2015 CCI land-cover layer, from which the GLOBIO LULC maps were partially derived (3). We assume that the remotely sensed LULC classes in the 2015 CCI land-cover map are a reasonable approximation of the type of vegetation likely to recolonize pixels identified as secondary vegetation in the GLOBIO 2050 LULC map.

Pixels that transition from pasture or forestry in the GLOBIO 2015 map to secondary vegetation in the 2050 GLOBIO map all intersect with natural land-cover classes in the 2015 CCI land-cover image. For such pixels we reclassify the 2050 secondary vegetation pixel to the 2015 CCI land-cover natural LULC category and match species habitat preferences to this category (Table 3).

There is no directly intersecting natural land-cover in the 2015 CCI land-cover map for pixels that transition from urban or crop in the 2015 GLOBIO LULC map to secondary vegetation in 2050. This is because urban and crop pixels in the GLOBIO 2015 LULC map are derived directly from the CCI 2015 land-cover map (2). We estimate the most likely vegetation type for each such pixel by creating a duplicate of the 2015 CCI land cover map in which all urban and crop cells are converted to the natural land cover occurring most commonly in a 1000 m circular neighborhood, repeated iteratively until there are no remaining urban or crop values in the image. Secondary vegetation pixels in the 2050 GLOBIO LULC map that transition from urban or crop in 2015 are then reclassified to this intersecting interpolated natural land-cover class and species habitat preferences matched accordingly (Table 3).

### ***Sensitivity analysis results***

For some species, the availability of secondary habitat is predicted to offset the loss of HSR driven by expansion of non-habitat land-uses. For many species this means that the net loss of HSR is lower under a regain scenario (Figure 2). Fewer species are predicted to be heavily impacted (≥10% net HSR loss of which ≥25% is caused by urban expansion) under a regain scenario, compared to a no regain scenario (Figure S2).

This result demonstrates that the impact of habitat loss is lessened, but still substantial under a scenario where species are able to recolonize secondary habitats. The results presented in this sensitivity analysis demonstrate the range of impacts that could be expected, given uncertainties regarding the type of secondary habitats likely to become available in 2050, and the capacity of species to use these habitats. Outcomes are likely to lie somewhere between the results presented in our regain and no regain scenarios, and will vary between species and regions.

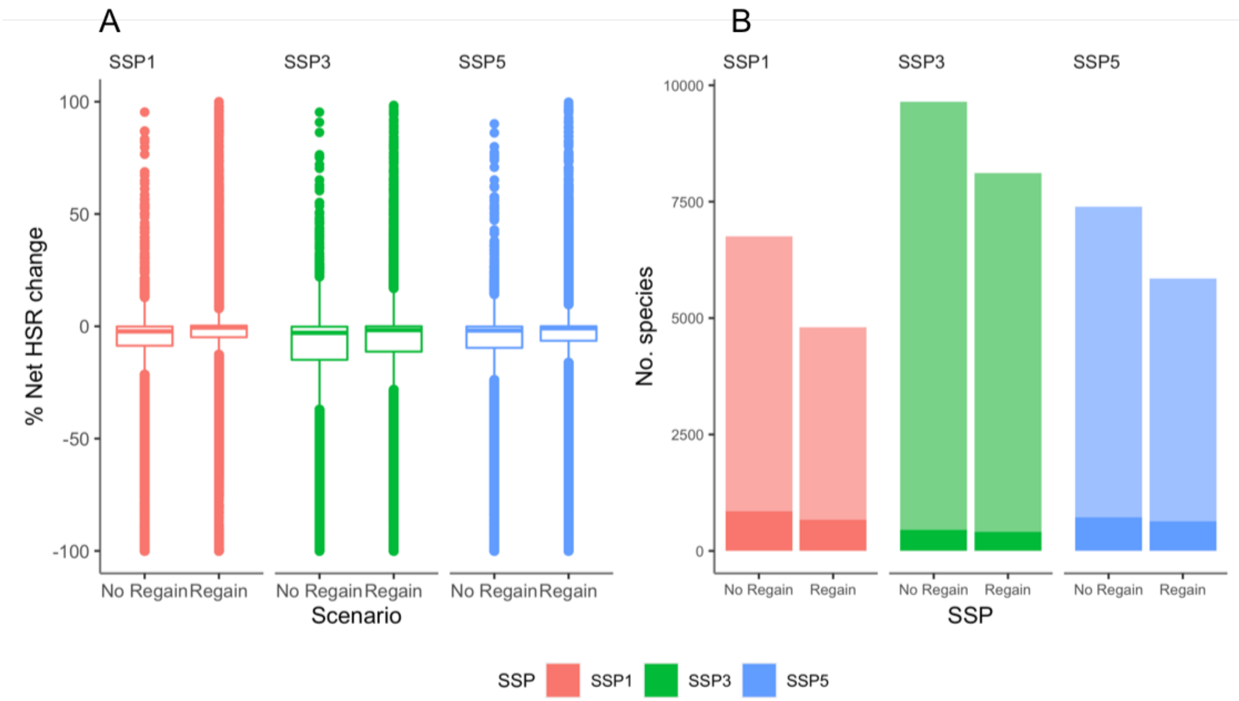


*Figure S2: Sensitivity analysis results. A) The distribution of percent net Habitat Suitable Range (HSR) change across all species under regain and no regain scenarios. Species that gain > 100% HSR are not shown (<1% of all). B) Number of heavily impacted species by SSP and Regain/No Regain scenario. Height of bars indicates the number of species that loose ≥ 10% of 2015 HSR by 2050. Darker stacked segments indicate the proportion of these species for which urban land expansion drives ≥ 25% of this habitat loss.*

***Dataset S1 (separate file).***

Species_appendix_table.xlsx – Summary of species predicted to be 'heavily impacted' by our analysis (i.e. $\geq$ 10% net HSR loss, of which urban land causes $\geq$ 25%).

***Dataset S2 (separate file).***

Cities_appendix_table.xlsx – Summary of urban impact hotspot clusters.

***References***


1. T. R. Loveland, A. S. Belward, The IGBP-DIS global 1km land cover data set, DISCover: First results. *International Journal of Remote Sensing* **18**, 3289-3295 (1997).
2. A. M. Schipper *et al.*, Projecting terrestrial biodiversity intactness with GLOBIO 4. *Global Change Biology* **26**, 760-771 (2020).
3. European Space Agency (2017) Land Cover CCI Product User Guide Version 2.0. (UCL-Geomatics, Belgium).
4. M. Friedl, D. Sulla-Menashe, (2019), MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MCD12Q1.006 Accessed via Google Earth Engine.
5. T. P. Robinson *et al.*, Mapping the Global Distribution of Livestock. *PLOS ONE* **9**, e96084 (2014).