Detailed Methods

Models of climate impacts

To assess climate impacts on species, I use species distribution models, which relate estimates of species presence or absence in different locations to variables describing the environment [1]. Most applications of species distribution models to date have used point occurrence records of species. However, such records are strongly biased taxonomically and geographically (toward Western Europe, North America and Australasia in particular) [2,3]. Global estimates of climate-change impacts have therefore tended to use expert-drawn extent-of-occurrence distribution maps instead [4–6], and this is the approach that I adopt here. I show that predicted changes in the distributional area of species are consistent with previous global studies that have used point occurrence data (by comparison of Figure S6 with Figure 2 in ref. [7]). Distribution maps for amphibians, reptiles and mammals were obtained from IUCN [8], and for birds from BirdLife International [9]. I excluded from these raw range maps areas where species are considered vagrant, and areas where species are present only in the non-breeding season or during migration. I then further refined the range maps to exclude areas outside the known elevational limits of species, where available. Elevational limits were obtained from the website of the IUCN Red List (www.iucnredlist.org; accessed on 19th September 2016) for amphibians, reptiles and mammals, and from BirdLife International’s World Bird Database (http://datazone.birdlife.org) for birds.

I used four climate variables, commonly reported to show a good association with vertebrate distributions [10]: minimum temperature of the coldest month, total annual precipitation, growing degree days, and water balance. I initially considered annual mean temperature as well, which has been used previously in applications of distribution models [10], but it showed a strong correlation globally with minimum temperature and growing degree days, which would have caused statistical problems for some of the modelling methods used (e.g. generalized linear models). Minimum temperature of the coldest month and total annual precipitation were taken directly from the WorldClim Version 1.4 dataset [11]. Growing degree days and water balance were calculated based on other climatic variables from WorldClim. Growing degree days was calculated using previously published methods [12] as follows. First, daily mean temperature values were interpolated from monthly values using cubic spline interpolation, implemented with the ‘spline’ function in R Version 3.3.2. Second, growing degree days was calculated as [12]:

\[ GDD = \sum \max\{0, (T_m - T_t)\} \]

where \( T_m \) is the daily mean temperature and \( T_t \) is the threshold temperature above which growth is possible (assumed to be 5°C). I also calculated water balance using previously published methods, as the difference between total annual precipitation and total annual potential evapotranspiration (PET) [12]. Monthly PET values were calculated as [12]:

\[ PET = 58.93 \times T_{(above \, 0^\circ C)} \frac{12}{12}, \]

where \( T \) is monthly average temperature. Monthly values were then summed to obtain total annual PET.

I resampled all of the distribution maps and climate data onto a 10-km × 10-km equal-area grid prior to analysis. All grid cells which intersected some part of a species’ distribution polygon were considered to be occupied, to avoid very narrowly distributed species being discounted from the analysis. Climatic variables were resampled using bilinear interpolation.
I fit distribution models using five methods commonly used in other studies [1,13]: Maxent, Generalized Linear Models (GLM), Random Forests (RF), BIOCLIM and DOMAIN. I chose methods that allow the fitting of relatively simple functional forms, to avoid as much as possible overfitting the coarse-scale distribution maps used. Initially, I also explored the use of Generalized Additive Models (GAMs), which have been shown to produce an accurate representation of species’ distributions [1], but these models proved to be computationally infeasible for generating global-scale models for all vertebrate species.

For all modelling approaches, I discounted all species that occupied fewer than 10 of the analysis grid cells. Upper limits on the number of occupied grid cells were also imposed for some of the modelling approaches, to maintain computational tractability (in terms of runtime and memory requirements): 400,000 for Maxent and GLM, 2,000 for RF, and 5,000 for BIOCLIM and DOMAIN. These limits were determined by experimentation of the time and memory requirements for the different models for different numbers of occupied grid cells.

Sampling of unoccupied grid cells was performed differently for the different modelling approaches. BIOCLIM and DOMAIN are presence-only approaches, not requiring information on unoccupied locations. For Maxent, 10,000 ‘background’ points (sometimes fewer for very widespread species) were sampled randomly, without replacement, from among all grid cells (both occupied and unoccupied). For GLM and RF, a set of unoccupied cells, of equal number as the occupied cells (or fewer for very widespread species where the number of unoccupied grid cells available was smaller than the number of occupied grid cells), were sampled randomly without replacement [14]. As in previous studies [5], to avoid model over-fitting [14], the selection of background or pseudo-absence points was limited to realm-biome combinations occupied by the species’ range map.

For Maxent and GLM models, the forms of the functions relating environmental variables to species’ presence can take a variety of complexities. To avoid overfitting, I restricted both models to use only linear and quadratic relationships. Random forests used the Breiman algorithm, with 500 candidate trees. BIOCLIM and DOMAIN are simple envelope approaches that fit species’ distributions based on the area of climatic space encompassed by the occupied locations [1]. All distribution models were fit in R Version 3.3.2 [15]: GLM models were fit using the ‘glm’ function in the standard ‘stats’ package; Maxent, BIOCLIM and DOMAIN models were fit using their respective functions within the ‘dismo’ package Version 1.1-4 [16]; and RF models were fit using the ‘randomForest’ package Version 4.6-12 [17].

To evaluate the accuracy of each distribution model, I initially divided the dataset into 80% for calibrating the model and 20% for evaluating the model. This division was performed just once for each species-model combination, rather than performing a full cross-validation, because of the long time needed to run the models for all species. I evaluated all models using the Area Under the Receiver Operating Characteristic Curve (AUC) statistic. Although there are issues with using AUC values to assess distribution model accuracy [18], AUC is a simple, generally applicable measure of model accuracy, and is widely used in broad-scale applications of distribution models [5]. I considered models with an AUC value greater than 0.8 to be useful for making future projections. Models with this level of accuracy were generated for 20,932 species using Maxent, 20,932 using GLM, 20,938 using RF, 18,184 using BIOCLIM, and 17,876 using DOMAIN.

Models of land-use impacts

The models of land-use impacts followed previously published methods [19]. The models were based on data in the PREDICTS database [20], which contains data on the composition of assemblages in different land uses, drawn from individual publications from all over the
world. I used data for vertebrate species only from the 2016 public release of the PREDICTS database [21]. This subset of the database contained 479,642 records of the occurrence or abundance of 6,184 vertebrate species (20% of all described terrestrial vertebrate species [22]), from 158 publications [23–180], sampled at 7,585 locations spanning all terrestrial biomes except one (flooded grasslands and savannahs).

For this study, I modelled just one measure of local assemblage biodiversity: sampled species richness. Although measuring biodiversity only as species richness may not capture all important facets of biodiversity change [181], species richness remains a widely accepted and widely used metric [19]. Furthermore, there is no simple monotonic relationship between the predicted climatic suitability from species distribution models and other measures of biodiversity, such as abundance [182]. Species richness was modelled as a function of three measures of human pressure at each site: a broad categorization of land use, intensity of human land use and human population density. Land use was classified based on the description of the habitat given in the source publications using published criteria [183] into the following categories: primary vegetation for habitat never previously destroyed by human actions or extreme natural events, secondary vegetation for habitat recovering after prior destruction by human actions or extreme natural events (divided into young, intermediate and mature secondary vegetation depending on stage of recovery), plantation forest for areas planted with tree crops (e.g. fruit or timber), cropland for areas planted with herbaceous crops, pasture for areas grazed regularly or permanently by livestock, or urban for areas with human habitation or buildings. The intensity with which humans use the land was classified into three levels – minimal, light and intense – also using the description of the habitat given in the original source publications. The criteria used for classifying land-use intensity varied by land-use category. For example, intensity in primary vegetation was classified based on disturbances such as hunting or logging, whereas intensity in cropland was classified based on crop diversity, crop rotation, chemical inputs and degree of mechanization. The full set of criteria are published elsewhere [183]. Human population density was estimated by overlaying the sampled sites with published estimates of human population density [184] using Python code developed for the arcpy module of ArcMap Version 10.3 [185].

I modelled species richness as a function of the human pressure variables using generalized linear mixed-effects models [186]. These models were implemented using the glmmADMB package Version 0.8.3.3 [187] in R Version 3.2 [15]. I used a model with a negative binomial error distribution, to account for the over-dispersion present in sampled species richness values. I fit a random intercept of study identity to account for differences in broad location, sampling methods, sampling effort, and taxonomic group sampled among different published studies. All human pressure variables were fitted as fixed effects. For human population density, I considered linear and quadratic polynomial terms. I also considered interactions between land use and land-use intensity, and between land use and human population density. I selected the best-fitting combination of fixed-effects terms and their interactions using backward stepwise model selection. Interaction terms were tested first, and then dropped to test the main effects. Any main effects that were part of significant interaction terms were retained in the final model, regardless of their significance as main effects.

**Climate and land-use scenarios**

I applied my models of the response of biodiversity to climate and land use to the Representative Concentration Pathways (RCP) scenarios [188], the most recent global and widely used set of scenarios for which both climate and land-use estimates are available. There are four RCP scenarios, which make widely differing assumptions about future socio-
economic pathways and thus very different predictions about future climate and land use (see Table 1). The scenarios are labelled in terms of the radiative forcing (the amount by which the energy balance of the Earth is altered, in W/m$^2$) in 2100: 2.6, 4.5, 6.0 and 8.5. RCP 8.5 has been characterized as ‘business as usual’ [19,189] and most closely matches recent trends in greenhouse gas emissions [190].

I obtained climate estimates for the RCP scenarios from WorldClim Version 1.4 [11] at a spatial resolution of 5 arc-minutes. These data consist of averages across two future time periods 2041-2060 and 2061-2080. They were generated by applying predicted anomalies from global climate models (GCMs) to the WorldClim estimates of climate for the reference period (1960-1990). I obtained estimates from the 12 GCMs that produced estimates for all four scenarios: BCC-CSM1-1, CCSM4, GISS-E2-R, HadGEM2-AO, HadGEM2-ES, IPSL-CM5A-LR, MIROC-ESM-CHEM, MIROC-ESM, MIROC5, MRI-CGCM3, and NorESM1-M. The derived bioclimatic variables – growing degree days and water balance – were calculated in the same way as for the reference time period (see above) for each GCM and future time period. I then calculated the average (mean) predicted value for each climate variable and each future time period across all 12 GCMs. Finally, I projected all maps of future climate onto the same 10-km equal area grid as used for fitting the distribution models (see above), using a bilinear interpolation. When applying the distribution models onto the future climate estimates, I used only the multi-model means and not the prediction uncertainty across GCMs, for consistency with the land use projections (uncertainty estimates are not available for the RCP land-use estimates [191]).

Projections of land use were taken from the land-use harmonization project [192]. Land-use predictions are generated by a single model for each RCP scenario, precluding a consideration of model uncertainty in the predictions [191]. The land-use projections consist of estimates of the proportion of each terrestrial grid cell, at a spatial resolution of 0.5°, in each of 6 major land-use classes: primary vegetation, secondary vegetation, plantation forest, cropland, pasture and urban. Predictions are made for a single ‘secondary vegetation’ class. To divide this class according to stage of recovery – into young, intermediate and mature secondary vegetation – I used the estimates of the proportional area of each grid cell that will transition between different land-use classes each time step to calculate the age profile of secondary vegetation in each cell and for each time step, following the methods in ref. [19].

Directly associated projections of human population density are available for the historical land-use estimates, and for the RCP 4.5 and RCP 8.5 future scenarios. In the first and last of these cases, the projections consist of gridded estimates. For RCP 4.5, projections are made only at the country level. To resolve these country-level estimates into gridded spatial estimates, I assumed that spatial patterns within countries were constant and matched present-day patterns (using the population density estimates described above). For RCP 2.6 and RCP 6.0, as in ref. [19], I used country-level projections of human population density from the UN population projections. Specifically I used the ‘medium’ scenario, which most closely matches the global assumptions about human population density in the RCP 2.6 and RCP 6.0 scenarios. These country-level projections were resolved into gridded estimates using the same method as for the human population projections from RCP 4.5.

Estimates of land-use intensity, which is important in determining the effects of land use on biodiversity, are not directly available in the land-use estimates described above. Instead, as in ref. [19], I estimated land-use intensity for the present day, for the future and for the reference period using simple spatial models. To estimate current land-use intensity (in the year 2005), I used the map of ‘Global Land Systems’ [193] at a spatial resolution of 5 arc-minutes, which divides coarse land-use categories into finer divisions based on cropland intensity, pasture livestock density and human population density. Each global land system was assigned to one of the land use and land-use intensity combinations considered in the
models of the response of biodiversity [19]. Then the proportional area of each 0.5° × 0.5° grid cell occupied by each land use and land-use intensity combination was estimated as the proportion of finer 5-arc-minute cells within this coarser grid cell that had a corresponding global land system category. To project land-use intensity for the reference period (1990) and the future periods, I modelled land-use intensity in 2005 (using general linear models) as a function of the proportion of cell area occupied by its corresponding land use, and human population density, allowing relationships to vary among UN sub-regions. UN sub-region data came from the world borders shapefile version 0.3 (http://thematicmapping.org/downloads/world_borders.php), transformed into a raster at 0.5° spatial resolution using ArcMap version 10.3 [185]. These models were then applied onto land-use and human population density estimates for the reference and future periods, as described above, assuming the spatial distribution of UN sub-regions to be constant.

Projections of climate-impact models

All species distribution models were projected using the ‘predict’ function in the ‘dismo’ package Version 1.1-4 [16]. This produces estimates of the relative climatic suitability of each grid cell in the predicted area and for the predicted time period and scenario combination. However, the exact interpretation of this output varies depending on the modelling algorithm: sometimes values represent probability of presence, sometimes just relative suitability [194]. Therefore, it is common practice in projecting distribution models to convert the raw output to a binary prediction of species presence or absence. To do this, I used a threshold that minimizes the difference between model sensitivity and specificity, which has been shown generally to perform well [194].

I projected species distributions for the reference time period (1960-1990) and each of the two future time periods (2041-2060 and 2061-2080), and for each of the four RCP scenarios of climate change (see above). The distribution models were projected onto the future climate estimates at a spatial resolution of 5 arc-minutes, before being resampled (using bilinear interpolation) to 0.5° resolution to match the land-use projections. For each of the 12 scenario-time period combinations I made one of three assumptions about species ability to disperse in response to changing climatic suitability. For all dispersal scenarios, including the scenario of unlimited dispersal, I assumed that species could not move beyond the combination of biome and biogeographic realm in which they occurred in the reference time period, thus assuming that major habitats would not shift in response to climate changes within the time period simulated. Whether or not this assumption holds probably depends on the biome: on the one hand, trees still have not filled all potentially suitable areas following post-glacial climate changes [195]; on the other hand, desertification can occur very rapidly in response to climatic changes [196]. The least conservative dispersal scenario assumed unlimited dispersal ability, where a species was assumed to be able to occupy all climatically suitable areas (provided that they fell within an originally occupied realm-biome combination). The most conservative dispersal scenario assumed no ability of species to disperse in response to changing climatic suitability, with species only able to occupy areas that were suitable in the reference time period and that remained suitable in future time periods. Finally, I simulated an intermediate dispersal scenario (termed here, the ‘realistic’ scenario) assuming that species could move at a specified rate away from suitable areas in the reference time period, using clade-specific rates of dispersal: 0.5 km year⁻¹ for reptiles and amphibians, and 3 km year⁻¹ for mammals and birds. These rates correspond with the ‘optimistic’ clade-specific dispersal rates assumed in a previous study of vertebrate responses to climate change [7]. Specifically, a grid cell could be occupied in the future if its centroid fell within a certain distance of the nearest originally suitable cell, set according to the above
clade-specific dispersal rates. Dispersal rates vary substantially within these major clades of species [197–199], but species-specific estimates are not available for most vertebrate species.

Overall projections of species richness change were calculated for each grid cell by summing all of the species for which the cell was estimated to be climatically suitable, and then expressing this species richness as a percentage of the value for the same grid cell in the reference time period. Uncertainty in the climate projections was estimated as the full range of projected species richness values across all of the five model types in the ensemble.

**Projections of land-use-impact models**

The models of the response of vertebrate species richness to land use were projected onto the scenarios of land use, land-use intensity and human population density (see above) following previously published methods [19]. These projections assumed that the relative species richness of a given grid cell, as a percentage of the species richness estimated to occur naturally, is a simple multiplicative function of the proportional area and relative biodiversity values of the different land use and land-use intensity combinations within a grid cell. For example, if a cell were half covered by primary vegetation, which as the baseline land use has a species richness value of 100%, and half by intensively managed cropland, which has 65% of the species richness of primary vegetation, then the overall species richness estimate for the cell before accounting for the effect of human population density would be 82.5%. The effect of human population density was estimated assuming that humans were distributed uniformly throughout a grid cell. In reality, human population density probably co-varies with land use and land-use intensity, but the scenario data did not permit this level of detail in the projections. Uncertainty estimates (95% confidence intervals) for the projections of land-use impacts were derived from the uncertainty in the coefficients from the mixed-effects models of land-use responses (see above).

**Combining the land-use and climate projections**

The projections of land-use and climate impacts on local species richness were combined assuming that the pressures act on species independently of each other. For example, if a grid cell was predicted to retain 50% of its species as a result of climate impacts and 50% as a result of land-use impacts, then the resulting species richness value with the effect of both pressures would be 25%. This assumption is unlikely to hold for a number of reasons. First, climate and land use might disproportionately impact a similar set of species. For example, narrow-ranged species have been shown to be disproportionately sensitive to land use (Newbold et al., submitted), and are likely also to be the most sensitive to climate change. Second, climate has been shown – at least at small scales – to influence how species respond to land use [200,201]. Third, land use might influence the ability of species to disperse through landscapes in response to climate change [202]. At present there is insufficient information to account properly for these interactions, but they should be included in more refined models in future. I do however test whether climate and land use are likely to have the greatest effects on a similar set of species (see next section). Global averages for the combined projections of land-use and climate impacts (as well as for the projections of the effects of the pressures individually) were calculated as the average across all terrestrial grid cells, weighted by cell area and by total ‘natural’ vertebrate species richness, i.e. in the absence of any climate or land-use impacts. Total natural species richness was estimated by overlaying the extent-of-occurrence range maps – as described above – for all mammals, birds, reptiles and amphibians.
Supplementary Figures and Tables

Figure S1. Map of sites with data used to assess the impact of land use on vertebrate community diversity. From the PREDICTS database [20,21]. Overlaid on a map of the world’s terrestrial biomes [203]. Point diameter is proportional to the log-transformed number of species recorded at each site. All points have a fixed opacity, so areas that appear more opaque contain a greater number of sites.
Figure S2. Predicted species richness of ecological communities under future climate and land-use change, for individual distribution modelling algorithms. Algorithms used were: Maxent (a, b), GLM (c, d), Random Forests (e, f), BIOCLIM (g, h) and DOMAIN (i, j). All values are expressed relative to a pre-human baseline. Left-hand panels (a, c, e, g, i) show separate effects of climate (solid, opaque lines) and land use (dashed, translucent lines); right-hand panels (b, d, f, h, j) show combined effects of both pressures (assuming no interactions).

Figure S3. Predicted species richness of ecological communities under future climate and land-use change, for different dispersal assumptions. I assumed either clade-specific dispersal rates of 0.5 km year\(^{-1}\) for amphibians and reptiles and 3 km year\(^{-1}\) for birds and mammals (a, d); no ability to disperse beyond areas suitable in the baseline period (b, e); or unlimited dispersal ability (c, f). All values are expressed relative to a pre-human baseline. Upper panels (a, b, c) show separate effects of climate (solid, opaque lines) and land use (dashed, translucent lines); lower panels (d, e, f) show combined effects of both pressures (assuming no interactions). Error bars show estimated uncertainty in the projections for the year 2070: 95% confidence intervals for land-use impact models, and range of estimates across the distribution model ensemble for the climate impact models.
Figure S4. Predicted proportion of the Earth’s terrestrial surface exceeding 20% species loss under future climate and land-use change, for individual distribution modelling algorithms. Algorithms used were: Maxent (a, b), GLM (c, d), Random Forests (e, f), BIOCLIM (g, h) and DOMAIN (i, j). All values are expressed relative to a pre-human baseline. Lefthand panels (a, c, e, g, i) show separate effects of climate (solid, opaque lines) and land use (dashed, translucent lines); right-hand panels (b, d, f, h, j) show combined effects of both pressures (assuming no interactions).

Figure S5. Predicted proportion of the Earth’s terrestrial surface exceeding 20% species loss under future climate and land-use change, for different dispersal assumptions. I assumed either clade-specific dispersal rates of 0.5 km year\(^{-1}\) for amphibians and reptiles and 3 km year\(^{-1}\) for birds and mammals (a, d); no ability to disperse beyond areas suitable in the baseline period (b, e); or unlimited dispersal ability (c, f). All values are expressed relative to a pre-human baseline. Upper panels (a, b, c) show separate effects of climate (solid, opaque lines) and land use (dashed, translucent lines); lower panels (d, e, f) show combined effects of both pressures (assuming no interactions). Error bars show estimated uncertainty in the projections for the year 2070: 95% confidence intervals for land-use impact models, and range of estimates across the distribution model ensemble for the climate impact models.
Figure S6. Predicted percentage of species losing more than half of their distribution area under future climate change, for different dispersal assumptions. I assumed either clade-specific dispersal rates of 0.5 km year$^{-1}$ for amphibians and reptiles and 3 km year$^{-1}$ for birds and mammals (a); no ability to disperse beyond areas suitable in the baseline period (b); or unlimited dispersal ability (c). All values are expressed relative to the baseline period (1960-1990). Error bars show estimated uncertainty in the projections for the year 2070, as the range of estimates across the distribution model ensemble. A comparison of the results in a with Figure 2 in ref. [7] shows broad consistency, although note that slightly different climate scenarios were used.
Figure S7. Response of local species richness to land use and land-use intensity (a) and human population density (b). Species richness in a is shown as a percentage change relative to the baseline condition (primary vegetation), with error bars showing 95% confidence intervals. Species richness in b is shown as absolute expected average species richness at each level of human population density and for each land use, with shading showing ±0.5 × SEM, for clarity. Colours used for each land use are the same in b as in a.
Table S1. Summary of predictions about climate, land use and human population made by the Representative Concentration Pathways Scenarios. Adapted from Table 1 in ref. [19]. Land use and human population estimates are from ref. [188], and climate from ref. [204]. The scenarios are labelled in terms of the radiative forcing (the amount by which the energy balance of the Earth is altered, in W/m$^2$) in 2100: 2.6, 4.5, 6.0 and 8.5.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Climate in 2070</th>
<th>Land use</th>
<th>Human population</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCP 2.6</td>
<td>Approximately 1°C.</td>
<td>Shift in agriculture from developed to developing countries. Large expansion of biofuel plantations. Urban area assumed to be constant.</td>
<td>10.1 billion by 2100.</td>
</tr>
<tr>
<td>RCP 4.5</td>
<td>Approximately 1.6°C.</td>
<td>Carbon markets lead to slower loss of primary forest and rapid expansion of secondary forest, especially in the tropics. Improved crop yields and agricultural efficiency, as well as dietary changes, cause decreases in agricultural area. Urban area assumed to be constant.</td>
<td>Reaches peak at 9 billion around 2065.</td>
</tr>
<tr>
<td>RCP 6.0</td>
<td>Approximately 1.6°C.</td>
<td>Human population growth drives an increase in cropland and urban areas. Decline in livestock grazing areas (pastures).</td>
<td>9.1 billion by 2100.</td>
</tr>
<tr>
<td>RCP 8.5</td>
<td>Approximately 2.7°C.</td>
<td>Agricultural areas expand to feed a growing human population, despite yield increases. Urban areas expand to house the human population.</td>
<td>12 billion by 2100.</td>
</tr>
</tbody>
</table>
Supplementary References


44. Centro Agronómico Tropical de Investigación y Enseñanza. 2010 Unpublished data of reptilian and amphibian diversity in six countries in Central America.


103. Malonza PK, Veith M. 2012 Amphibian community along elevational and habitat


120. Neuschulz EL, Botzat A, Farwig N. 2011 Effects of forest modification on bird community composition and seed removal in a heterogeneous landscape in South


201. Rogelj J, Meinshausen M, Knutti R. 2012 Global warming under old and new