This is the final accepted manuscript published in *Limnologica*
https://doi.org/10.1016/j.limno.2023.126139

Downscaling future land cover scenarios for freshwater fish distribution models under climate change

Annika Brunner[1,2*], Jaime R. García Márquez[1], Sami Domisch[1*]

1. Leibniz Institute of Freshwater Ecology and Inland Fisheries, Department of Community and Ecosystem Ecology, Müggelseedamm 310, 12587 Berlin, Germany

2. Free University Berlin, Institute of Biology, Department of Biology, Chemistry and Pharmacy, Königin-Luise-Straße 28-30, 14195 Berlin, Germany

[*]**corresponding author:** annika.brunner@gmx.de, sami.domisch@igb-berlin.de

Abstract

The decreasing freshwater biodiversity trend can be attributed to anthropogenic impacts in terms of climate and land cover change. For targeted conservation efforts, mapping and understanding the distribution of freshwater organisms consists of an important knowledge gap. Spatial modelling approaches offer valuable insights into present-day biodiversity patterns and potential future trajectories, however methodological constraints still hamper the applicability of addressing future climate and land cover change concurrently in one modelling workflow. Compared to climate-only projections, spatially explicit and high-resolution land cover projections have seen less attention, and the lack of such data challenges modelling efforts to predict the possible future effects of land cover change especially on freshwater organisms. Here we demonstrate a workflow where we downscale future land cover projection data from the Shared Socioeconomic Pathway (SSP) scenarios for South America at 1 km^2 spatial resolution, to then predict the future habitat suitability patterns of the Colombian fish fauna. Specifically, we show how the land cover data can be converted from plain numbers into a spatially explicit representation for multiple SSP scenarios and at high spatial resolution, employing freshwater-specific downscaling aspects when spatially allocating the land cover category grid cells, and how it can be fitted into an ensemble species distribution modelling approach of 1209 fish species. Our toolbox consists of a suite of open-source tools, including Dinamica EGO, R, GRASS GIS and GDAL, and we provide the code and necessary steps to reproduce the workflow for other study areas. We highlight the feasibility of the downscaling, but also underline the potential challenges regarding the spatial scale and the size of the spatial units of analysis.

Keywords: Shared Socioeconomic Pathway (SSP), Dinamica EGO, land cover, freshwater fish, species distribution model, climate change

1 Introduction

Freshwater habitats can be considered biodiversity hotspots given that they cover less than 1% of the Earth's surface, yet they accommodate approximately 10% of all known animal species (Balian et al., 2007). At the same time, freshwater ecosystems can be considered the most anthropogenically transformed habitats (Cowx, Portocarrero Aya, 2011), and are exposed to pollution, habitat degradation, flow alterations and invasive species, which are mediated and amplified by climate and land cover change (Allan, 2004; Reid et al., 2019). Consequently, freshwater biodiversity has been impacted detrimentally (Dudgeon et al., 2006) resulting in a significant decline in e.g. freshwater vertebrate and megafauna populations (He et al., 2019; W.W.F., 2020).

In addition to the observed changes in freshwater biodiversity, model projections have proven their high potential in providing an advanced understanding towards the patterns, dynamics and consequences of climate and land cover change (Soares Filho et al., 2003). Several studies show, either through empirical investigations or modelling, that land cover and changes in land cover have an impact on e.g. water quality, river discharge or sedimentation processes due to pollution, hydrological changes, or riparian deforestation, and impact freshwater organisms (e.g. Allan, 2004; Kuemmerlen et al., 2015; Dala-Corte et al., 2016; Radinger et al., 2016; Leitão et al., 2018; Tóth et al., 2019; Lo et al., 2020).

Species distribution models (SDMs) are on the forefront for projecting possible changes in biodiversity (Elith, Leathwick, 2009) which in riverine ecosystems, require environmental information that is aggregated to stream segments or sub-catchments to match the spatial configuration of the stream network. While such present-day information regarding climate and land cover has only recently become available (Domisch et al., 2015a; Linke et al., 2019; Amatulli et al., 2022), future climate and land cover change projections tailored towards freshwater ecosystems remain scarce, where the downscaling of coarse-scale data to the network or sub-catchments still poses a challenge. Given the lack of such freshwater-specific future projections, care must be taken that studies that analyse possible future projections, do not run the risk of being overly simplified and hence of limited value to practical implementations (Schuwirth et al., 2019).

Following best practices by e.g. Eraso et al. (2013), Galford et al. (2015) or Gollnow et al. (2018), future trend analyses regarding global change effects should combine both quantitative and qualitative information. In this regard, the narratives of the Shared Socioeconomic Pathways (SSPs) have been used to examine plausible future changes by including such quantitative and qualitative information regarding economic growth, political stability, urbanization, environmental awareness or land cover, describing five different development paths until the year 2100 (Ebi et al., 2014; O'Neill et al., 2017; Riahi et al., 2017). Much effort has concentrated on high-resolution Earth Observation data, such as Digital Elevation Models (Yamazaki et al., 2017), present-day and possible future climate projections (Karger et al., 2017; Karger et al., 2018), as well as present-day land cover projections (Hoskins et al., 2016; ESA, 2017). Land cover projections for the future, however, are only available at a coarse spatial resolution (e.g. at ca. 50 km^2, Chen et al., 2020) which poses a challenge for incorporating these data in freshwater species distribution models due to the mismatch between the coarse gridded data and the structure of the stream network (Domisch et al., 2015b), where the information of coarse-grained grid cells can not be associated to specific stream segments or sub-catchments. High-resolution, downscaled land cover data remains therefore often constrained to small geographic extents and specific locations based on local case studies (da Silva Cruz et al., 2022).

Addressing the spatial scale and configuration is critical for modelling species distributions in freshwater ecosystems. For example, continuous climate data changes gradually within a given study area, whereas categorical land cover data changes between one grid cell to another, requiring special attention given the topographical and topological characteristics of the stream network. The data needs to be therefore tailored towards this spatial configuration that differs fundamentally from the terrestrial realm. For instance, two streams can have an euclidean distance of only few kilometers, but can have a network distance of hundreds of kilometers (Rodriguez-Iturbe, Rinaldo, 2001), such that e.g. coarse-scale, gridded land cover data does not match the stream network. This could be overcome by using high-resolution land cover data that also addresses e.g. the distance of a given land cover category to freshwater features such as streams, rivers, lakes, ponds and wetlands.

Bearing in mind the issues of scale and the special spatial configuration in the freshwater realm, our objective is to develop a generic workflow that allows integrating coarse, continental-scale SSP land cover information in freshwater SDMs, where such land cover projections can be concurrently used with climate projections, along with other environmental data. Such a workflow should ideally allow (i) future, high-resolution land cover projections to be available across the entire study area at a standardized 1 km^2 spatial resolution, where (ii) the downscaling should take freshwater-specific aspects into account when spatially allocating the land cover category grid cells. Most importantly, (iii) the newly-developed land cover projections should match the climate projections temporally, as well as other environmental data spatially, such that (iv) all data can be seamlessly integrated into a freshwater SDM approach.

Here we showcase and test such a workflow in a freshwater biodiversity hotspot region and selected the northest part of South America as our study region with 1209 freshwater fish species. We demonstrate the applicability of the workflow by employing a toolbox of open-source geospatial software and provide the code for reproducing the workflow for other study areas, and discuss possible challenges. We highlight that the fish distribution modelling analyses in our study are for demonstration purpose as to emphasize the feasibility of the approach, while more detailed and in-depth analyses would be needed to assess the potential global-change impacts of the Colombian freshwater fish fauna.

2 Methods

The general workflow consisted of three steps (Fig. 1 A-C). In the absence of spatially-explicit land cover projections for our study area, we first translated the continental-scale numerical (and non-spatial) SSP land cover projections developed for Latin America and the Caribbean region into spatially explicit information and simultaneously downscaled it to 1 km^2 spatial resolution. We then employed spatially explicit land cover models (LCM) which we calibrated using the original, projected changes in the SSP data from 2005 to 2100. Second, we extracted a river network and defined the sub-catchments as the new spatial units of analysis to which all spatial data was aggregated (from previous 1 km^2 grids). Third, we narrowed the study area to include Colombia and adjacent catchments in Peru, Ecuador, Brazil and Venezuela and collated freshwater fish data for this region. Finally, we modelled fish distributions occurring in Colombia for the present time step as well as for the year 2070, incorporating climatic, environmental and the newly-generated land cover data projections.

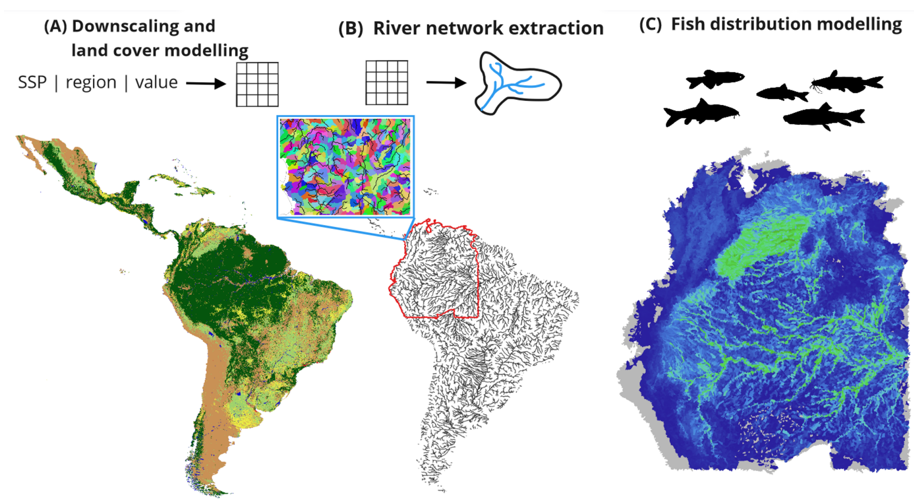

Figure 1: The workflow of the study is split into three parts: (A) the spatial downscaling and land cover modelling, (B) the extraction of the river network and defining the sub-catchments as the spatial units of analysis, and (C) the fish distribution modelling within a larger window of Colombia.

2.1 Study area

We focused on two spatial domains: for the LCM, as well as the river network and sub-catchment extraction, we used data across the entire Latin America and the Caribbean region (Fig. 1, A-B) since the SSP database lumps this region into a single spatial unit, which is an area of approximately 2200 million hectares (Mha). For the species distribution modelling and further analysis we narrowed the spatial domain to Colombia and adjacent catchments in Peru, Ecuador, Brazil and Venezuela, referred to as the study area, with a size of 513 Mha (Fig. 2, red polygon). The study area was larger than Colombia as to ensure that we account for the range-wide distribution of fish species in the SDMs, including a larger number of species occurrences within their ranges and to capture wider gradients of environmental predictors (Barbet-Massin et al., 2010).

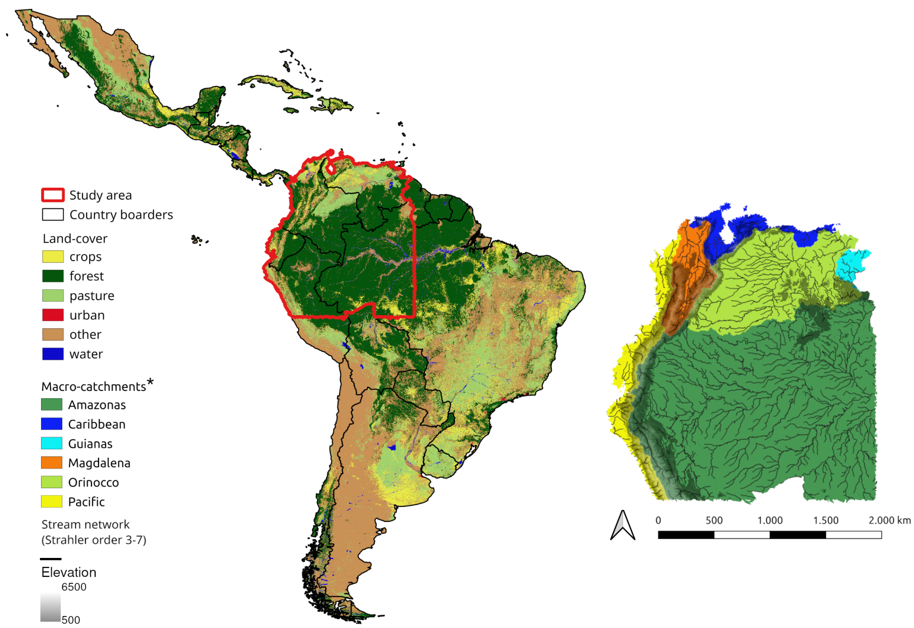

Figure 2: Maps of the two spatial working scales. **Left**: Land cover clustered after the SSP categories for Latin America and the Caribbean. The red line represents the study area. Country boarders are marked with black lines. **Right**: The spatial domain used for the species distribution modelling (i.e., the study area). The six colours represent the drainage basins, following the assignment by Reis et al. (2016). Black lines show the newly-delineated stream network, here showing the streams of Strahler order 3-7. Grey shading represents elevation between 500 and 6500m to highlight the Andes for orientation.

2.2 Data

2.2.1 Shared Socioeconomic Pathway (SSP) data

We obtained the land cover projection data that incorporates the SSP scenarios from the public SSP database at (`https://tntcat.iiasa.ac.at/SspDb`; Popp et al., 2017). This data corresponds to the so-called baseline storylines, where Representative Concentration Pathways (RCPs) are assigned to the respective scenario corresponding to a development where no mitigation or adaptation measures are applied. The SSP data includes the land cover types crops, forest, pasture, urban/built-up area and other natural land. This data is only available as tables, describing the broad coverage across broad regions or continents, and needs to be transformed into a spatially explicit format prior any spatial analysis.

Box 1. The Shared Socioeconomic Pathways (SSPs)

The SSPs aim to examine plausible future changes by including different quantitative and qualitative drivers regarding economic growth, political stability, urbanization, environmental awareness or land-cover. In addition, they integrate climate change projection through the RCPs that were developed in the first phase of the scenario development process and describe the possible range of radiative forcing values in the year 2100 (Ebi et al., 2014). There are five different development paths until the year 2100 which are distinguished on the basis of challenges to adaptation and mitigation to climate change (O'Neill et al., 2017; Riahi et al., 2017).

SSP1 Sustainability – Taking the Green Road
(Low challenges to mitigation and adaptation)
SSP2 Middle of the Road
(Medium challenges to mitigation and adaptation)
SSP3 Regional Rivalry – A Rocky Road
(High challenges to mitigation and adaptation)
SSP4 Inequality – A Road Divided
(Low challenges to mitigation, high challenges to adaptation)
SSP5 Fossil-fueled Development – Taking the Highway
(High challenges to mitigation, low challenges to adaptation)

2.2.2 Land cover data

We obtained present-day land cover data from the European Space Agency (ESA) for the years 1992 to 2018 for 22 land cover categories and 14 regional sub-categories at 300 m spatial resolution (Bontemps et al., 2013; ESA, 2017). In the Latin America and the Caribbean region 20 of the land cover and seven of the regional sub-categories were present. We clustered all land cover categories in the ESA maps into six groups, namely crops, forest, pasture, urban/built-up area, water, and other natural land, to match the SSP categories, including water as an extra group as to emphasize the presence of inland waters (see Supplementary Material Table S1). We also resampled the maps to 1 km^2 spatial resolution as to standardize all data to a common spatial grid using the *gdalwarp* function from the Geospatial Data Abstraction Library (GDAL; GDAL/OGR contributors, 2021).

2.2.3 Environmental data for the land cover modelling (LCM)

We selected the environmental variables for the LCM (Table 1) following recommendations by Eraso et al. (2013), Galford et al. (2015), and Gollnow et al. (2018) to account for the influence of environmental features on land cover change and to integrate water-related information.

2.2.4 Environmental data for the SDM

We started with an initial set of 45 different bioclimatic, river topology and land cover variables that are frequently used in freshwater SDMs (Buisson et al., 2008; Bond et al., 2011; Martin et al., 2013; Porfirio et al., 2014; Kuemmerlen et al., 2015). Regarding the bioclimatic and topographical variables, we calculated the mean and range of values within each sub-catchment, whereas for the LCM-derived categorical land cover, we computed the percent coverage of each category within a given sub-catchment. The present-day bioclimatic data stems from the period 1979-2013, whereas the future projections for 2070 refers to the period 2061-2080 and are developed within the framework of the Coupled Model Intercomparison Project 5 (CMIP5, Taylor et al., 2012). We used future climate projections of the following global circulation models (GCMs) from the CHELSA database (Karger et al., 2017; Karger et al., 2018): CCSM (Community Climate System Model); IPSL-CM5A-MR (Institute Pierre Simon Laplace medium resolution climate model); and MIROC-ESM-CHEM (Model for Interdisciplinary Research on Climate Earth system Model with an atmospheric chemistry component).

Prior the SDMs we reduced the number of the variables by pair-wise correlation tests and excluded highly correlated variables. Here, we considered the Pearson correlation coefficient $r >= \mid 0.7 \mid$ as a threshold (Bahrenberg et al., 2017). When excluding the variables, care was taken to retain specific climate data such as annual mean temperature, as they provide information regarding long-term physiological constraints for the species. We finally used 16 environmental variables in the SDMs (Table 1).

Variable	LCM	SDM	Unit	Source
Annual mean temperature	X	X	°C *10	Karger et al. (2018)
Temperature range	X	X	°C *10	Karger et al. (2018)
Temperature seasonality		X	°C *10	Karger et al. (2018)
Annual precipitation (mean)	X	X	mm	Karger et al. (2018)
Annual precipitation (range)		X	mm	Karger et al. (2018)
Precipitation seasonality		X	mm	Karger et al. (2018)
Elevation	X		m	Amatulli et al. (2018)
Slope	X		°	Amatulli et al. (2018)
Inland water bodies	X		categorical	Tootchi et al. (2018)
Distance to highways primary and secondary roads	X		m	Meijer et al. (2018)
Distance to "all" roads	X		m	Meijer et al. (2018)
Distance to sea	X		m	ESA (2017)
Distance to inland water bodies	X		m	Barbarossa et al. (2018)
Distance to towns/settlements	X		m	reclassified LC maps, ESA (2017)
Protected Areas	X		categorical	UNEP-WCMC (2020)
Crops		X	%	LCM
Forest		X	%	LCM
Pasture		X	%	LCM
Urban		X	%	LCM
Water		X	%	LCM
Other natural land		X	%	LCM
Flow	X		m/s	Barbarossa et al. (2018)
Cumulative stream length		X	km	Amatulli et al. (2018) GRASS GIS / *r.stream.order*
Flow accumulation		X	km^2	Amatulli et al. (2018) GRASS GIS / *r.watershed*
Outlet distance		X	km	Amatulli et al. (2018) GRASS GIS / *r.stream.order*
Elevation drop		X	m	Amatulli et al. (2018) GRASS GIS / *r.stream.order*

Table 1: Environmental variables used in the Land cover modelling (LCM) and Species distribution modelling (SDM, marked with "X"), their units and sources. All distance-related maps were calculated using Dinamica EGO with the indicated source maps as input. Categorical inland water bodies refer to wetlands, lakes and streams. Categorical protected area definitions refer to strict nature reserves, wilderness areas, national parks, natural monuments or features, habitat/species management areas, protected landscapes/seascapes, protected areas with sustainable use of natural resources, as well protected areas as not reported, not applicable or not assigned.

2.2.5 Fish occurrence data

We retrieved fish occurrence data from five databases: the Global Biodiversity Information Facility (GBIF, 2020), Fishnet2 (fishnet2, 2020), the speciesLink Network (Canhos et al., 2022), Integrated Digitized Biocollections (iDigBio, 2020) and expert data from fish taxonomists across Colombia.

In total, this data consists of 188,362 occurrence records representing 2075 fish species collected between 1700 and 2020. We subsequently restricted the time frame to 1980-2018 to match the environmental data, resulting in 153,364 unique occurrence records for 2030 species. We further reduced the dataset towards species known to occur in Colombia using the "Checklist of the freshwater fishes of Colombia" (DoNascimiento et al., 2017) comprising 1494 species. These species' occurrence records often extend beyond Colombia, since we aimed to capture the wider range of species observed distributions as to not truncate the environmental gradients in the SDMs (Barbet-Massin et al., 2010). We aggregated the fish occurrence data to the sub-catchments which served as the spatial units in the SDMs. This procedure collapsed multiple observations of a given species per sub-catchment to one, also reducing the sampling bias (Porfirio et al., 2014; Melo-Merino et al., 2020). Finally, we considered all species with at least two unique occurrence records across sub-catchments as candidate species in the SDMs. Our final set hence comprised 1209 fish species occurring in Colombia, of which 219 are endemic to Colombia (DoNascimiento et al., 2017).

2.3 Data processing & analysis

2.3.1 River network extraction

We use the GMTED-Digital Elevation Model (DEM) (Danielson, Gesch, 2011) obtained from www.EarthEnv.org/topography at 1 $\mathbf{km^2}$ spatial resolution (Amatulli et al., 2018) to extract the stream network and the corresponding sub-catchments for Latin America and the Caribbean using GRASS GIS (Neteler et al., 2012). First, we 'burned' the water bodies from the Open Street Map (OpenStreetMap contributors, 2017) using an elevation depth of 20 m to ensure that water is flowing along the known flow paths. We then employed the *r.watershed* function to derive the amount of overland flow accumulation and the flow direction, and to extract the corresponding stream network (Neteler et al., 2012). Here, we used a threshold of 110 grid cells (approx 11000 ha) to initiate a river channel, i.e. defining the minimum upstream catchment size. Next we computed the corresponding sub-catchments using the *r.stream.basins* function along with the stream network and the flow direction map. We assigned each sub-catchment a unique ID that corresponds to the river segment. We also computed the macro-catchments using the "-l" flag of *r.stream.basins*, delineating basins for the stream outlet (Neteler et al., 2012). Finally, we computed the stream network topology and a stream vector map using the *r.stream.order* function.
Moreover, we computed for each species the network distance from each occurrence record to the drainage border, allowing us later to apply these species-specific layers to (i) mask those drainages that are not connected to the occurrences, assuming that these can not be reached by the given species, and to (ii) downweight the suitable habitats in sub-catchments that are distant from the species known range.

2.3.2 Land cover modelling

We developed the LCM with the open-source software Dinamica EGO (Environment for Geoprocessing Objects) that allows to build environmental simulation models through a combination of map algebra, cellular automata technique and tabular data manipulation (Soares Filho et al., 2003; Rodrigues, Soares-Filho, 2018).

For the LCM calibration we required the present-day land cover and the future SSP narratives and corresponding scenarios of how the land cover is likely to change. In general, the predictions of future land cover developments in the SSPs are made using various integrated assessment models (IAM) with land cover modules. For the implementation of each scenario a specific IAM is assigned as the so-called marker scenario (we refer to Popp et al., 2017 for further details). Given the variability in IAM input data for the base year (2005), the land cover outputs of the IAMs may differ (Popp et al., 2017; van Vuuren et al., 2017; Fricko et al., 2017). Here, we chose the IMAGE model as our IAM, as it covers all five SSP scenarios, features dynamic data for the built-up area category and focuses on environmental issues (opposed to e.g. an economic focus of other IAMs; Popp et al., 2017; van Vuuren et al., 2017). For the specific input parameters of Dinamica EGO we refer to the Supplementary Information.

2.3.3 Species distribution modelling

We used an ensemble model in the R-package biomod2 (R Core Team, 2020; Thuiller et al., 2020) that integrates different modelling algorithms, and combines the individual algorithm results into an overall trend while reducing the uncertainty (Thuiller et al., 2009). We chose three algorithms across different model families: Generalized Linear Model (GLM) as a regression algorithm (McCullagh, Nelder, 1989), Classification Tree Analysis (CTA) which is a classification algorithm (Breiman et al., 1984) and Maximum Entropy (MAXENT.Phillips) as a machine learning method (Phillips et al., 2006).

We extracted the environmental data (Table 1) for the occurrences as well as for 10,000 random background absences of each each species. We split the data into a 70% model training and 30% testing sets, and ran ten repetitions for every algorithm and species. Across all model runs, we calculated a weighted ensemble, where the weights were based on the model performance statistics for each algorithm (Hao et al., 2019). We evaluated the models using the Area Under the Curve (AUC) and True Skill Statistic (TSS). Species with TSS <0.4 were removed from further analyses. Moreover, we computed the relative variable importance for each species.

The probability of habitat suitability of each sub-catchment is given as a continuous value between 0 and 1 for each species. We then transformed the probabilities into a binary output, indicating whether a sub-catchment is deemed

suitable for a species (1) or not (0) employing the TSS cut-off value as a threshold. For the projections for 2070, three outputs per SSP, each generated with a different GCM (see section "Environmental data for the SDM"), were averaged and converted to a binary output. We ran model projections for each species and GCM separately, and averaged the future habitat suitability projections for each species to obtain one final projection. We ran future projections for the SSP2 (the business-as-usual scenario with assigned RCP 6.0) and SSP5 (the fossil-fueled development scenario with assigned RCP 8.5).
We addressed the connectivity and accounted for possible dispersal constraints by multiplying the probability and binary maps with the scaled distance map. This allowed us to weight the spatial distribution of the species by spatial limits and thus to simulate dispersal constraints: the habitat suitability was down-weighted linearly with increasing distance from the occurrence records and was set to zero in non-connected drainages.

2.3.4 Analysing land cover trends

We calculated the changes in the six land cover categories between 1992 and 2018 in hectares (Eq. 1) and percent (Eq. 2) as:

$$\text{rt} = A_{t2} - A_{t1} \text{ (Eq.1)}$$

$$\text{rt} = \left(\frac{A_{t2}}{A_{t1}} - 1\right) * 100 \text{ (Eq.2)}$$

where A=area and t=time.

We further calculated the projected coverage for cropland, forest, pasture, and urban areas from 2005 to 2100 resulting from the LCM using the built-in function *Area* in Dinamica EGO and the rates of change (see Eq. 2 above) and also located the main areas of change.

2.3.5 Analysing SDM outputs

We stacked all individual fish projections to estimate present-day and future species richness given the SSP2 and SSP5 model runs for 2070 (Distler et al., 2015). We further overlaid each species ensemble projection with forest, crop, pasture, urban land cover and average annual temperature and precipitation to evaluate the spatial habitat preferences. We used Pearson correlation tests to explore the relationships between forest cover, temperature, and precipitation against the future modelled changes in the occupied area, the average elevation and flow accumulation of those sub-catchments modelled as suitable habitat. We ran identical analyses for endemic vs. non-endemic fish species to analyse possible effects regarding their range-restrictedness using a Welch test (Lu, Yuan, 2010).

3 Results

3.1 Past and future land cover trends

All five scenarios predict an increase in cropland (Fig. 3 A) from 2005 to 2070. The cropland coverage in SSP2 is projected to be 11.2 Mha (+ 39.9%) larger in 2070 than in 2005, and in SSP5 it is projected to increase by 18.8 Mha (+ 67.2%).

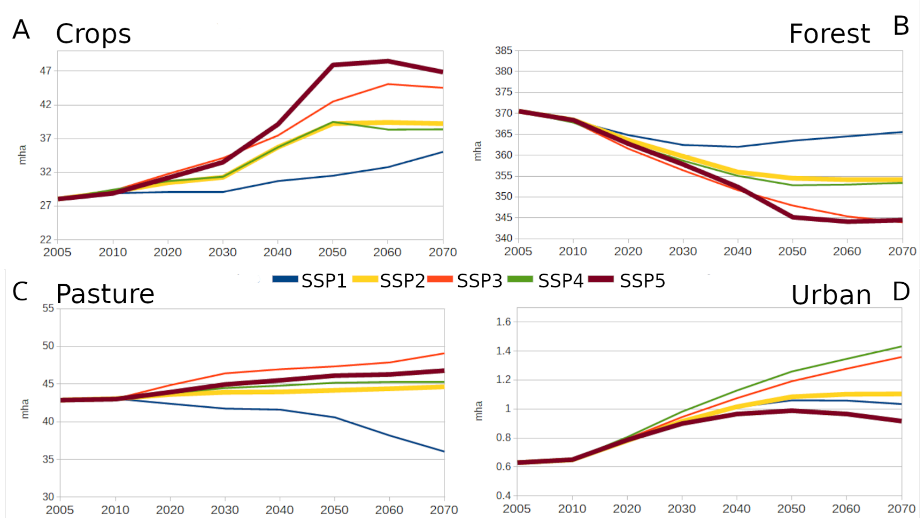

Figure 3: Projected change of the four land cover classes in the study area. The coloured graphs depict the changes of the five SSPs 2005 to 2070 with (A) crops, (B) forest, (C) pasture and (D) urban. SSPs 2 and 5 are shown in bold, because they are the SSPs that are described in more detail in the text.

Overall, forest coverage is projected to decrease from 2005 to 2070 for all scenarios (Fig. 3 B). Under the SSP2 scenario, a decrease in forest coverage is predicted from 370.53 Mha to 354 Mha (-4.4%) by 2070. For SSP5 a decrease of 26.4 Mha forest coverage is projected between 2005 and 2070 with an average change rate of -0.9% per 10 year interval.
In SSP2, the pasture coverage is predicted to increase 1.8 Mha (+ 4.1%) by 2070 (Fig. 3 C). Under the SSP5 scenario, an increase by 2070 of 3.4 Mha (+9.1%) is predicted.
All scenarios show an increase in urban area by 2050 (Fig. 3 D). Urban land growth is projected to be stagnant after 2050 for SSP2. The increase rates per decade are projected to remain almost the same in the years after 2050. In SSP 5, on the other hand, the built-up coverage declines after 2050, so that by 2070, 0.2 Mha less than in 2050 are covered by urban area.
Most of the land-cover changes projected in the LCM occur in the south-eastern part of the study area and along a broad section following the coastline, espe-

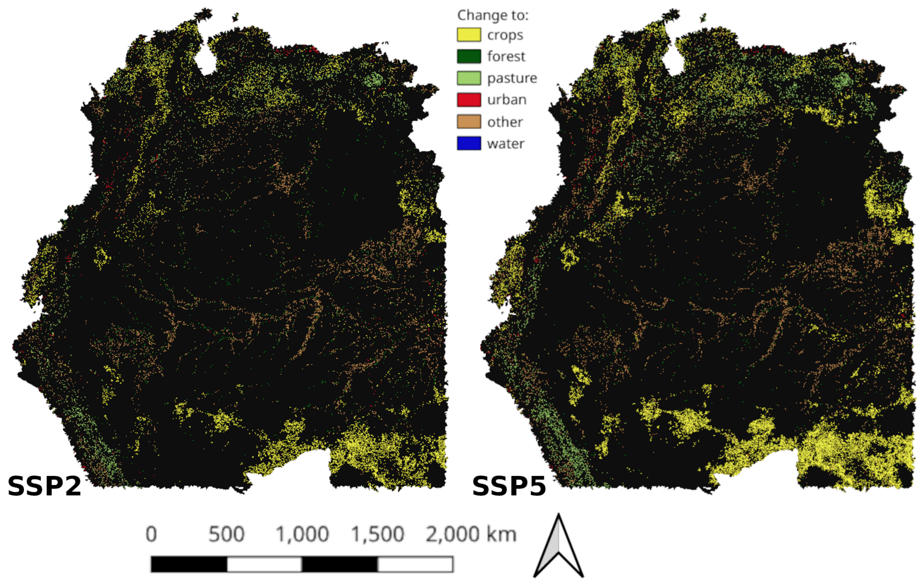

Figure 4: Changes in land-cover for 2010-2070. The colors represent the land-cover category to which the land has changed into. Black pixels indicate no change in the corresponding period. For interactive maps that allow to toggle and zoom individual layers, please see `https://glowabio.org/project/ssp_landcover/`.

cially in the north (Fig. 4). The main transformation corresponds to cropland, pasture and other natural land. Furthermore, the expansion or re-development of built up area, occurs mainly in the northern and north-western part of the study area. For the Amazon basin, changes are projected especially along the rivers.

3.2 Fish habitat distribution

Across all 1209 fish species, the models including all environmental variables achieved a True Skill Statistic (TSS) of 0.87 ± 0.11 (mean, $\pm$ standard deviation) and an AUC value of 0.97 ± 0.03 in the performance evaluation. We note that our intention was to be inclusive in that we built SDMs also for species with a low number of occurrence records as to highlight the feasibility of our approach. In this regard, the 20 species (2% of all) with only two unique occurrences had a mean TSS of 0.99 ± 0.01 and an AUC of 0.97 ± 0.006. No species was discarded from the final analyses.
The mean number of species for which a sub-catchment is considered suitable (here also referred to as species richness) is 98 in the present, with a maximum

of 605 species projected in a single sub-catchment. In 2070, SSP2 shows a mean richness of 166 and SSP5 of 203. The maxima are 661 and 666 for SSPs 2 and 5, respectively.
In the present, habitat suitability is projected for 24,675 sub-catchments which corresponds to an area of 4,678,022 km^2. Until 2070, 280 more suitable sub-catchments (+40,808 km^2) would arise for SSP2, and 346 (+49,365 km^2) more for SSP5 (Table S4).
Models projected that 27 (SSP2) and 32 (SSP5) species would loose their entire suitable habitat until 2070. This corresponds to a potential loss of 1% (SSP2) to 2% (SSP5) of species between the present and 2070 (Supplementary Material Table S4).

Of the 1209 species, 219 are designated as endemic to Colombia according to the checklist of Colombian freshwater fish (DoNascimiento et al., 2017). The projected mean species richness for endemic species in the study area is five in the present with 16,750 (=3,433,954 km^2) suitable sub-catchments. In 2070, the mean richness would remain at five for SSP2, while it would increase to six in the SSP5. Between the present and 2070, the number of sub-catchments with suitable habitats would increase by 1879 (359,650 km^2) for SSP2 and 4571 (732,222 km^2) for SSP5. The projected maximum species richness in the present is at 86 for SSP2 and 90 in SSP5. In 2070, the maximum richness would change to 84 (SSP2) and 87 (SSP5). The number of fish projected to loose their entire suitable habitat would be 16 and 20 for SSP2 and SSP5, respectively. This represents a potential loss between 5% and 9% of species, compared to the present.

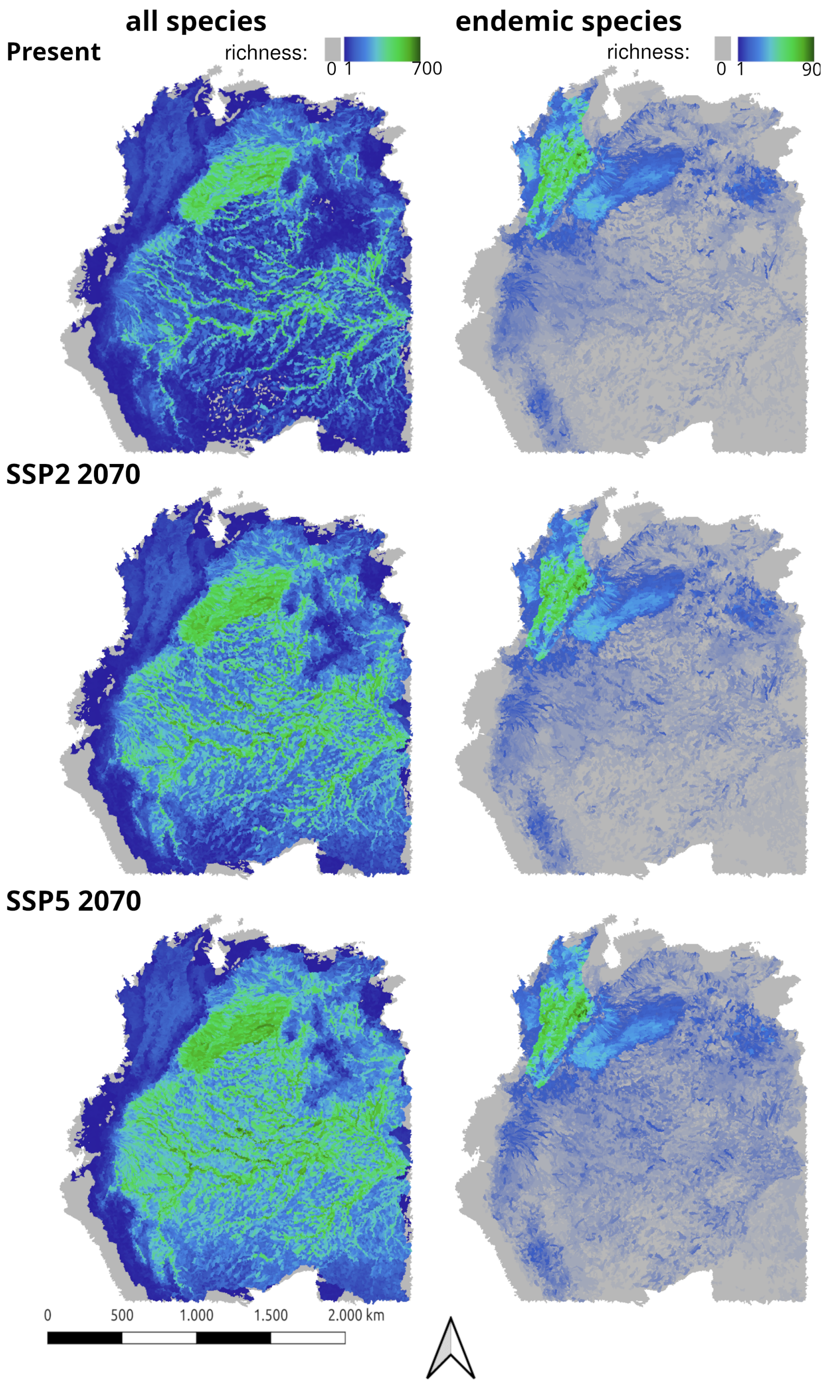

Figure 5: Species richness (i.e., stacked habitat suitability estimates) per sub-catchment. Number of species for which a sub-catchment is suitable ranging from one species (dark blue) to the highest suitability (dark green). Grey refers to no habitat suitability. For interactive maps that allow to toggle and zoom, please see `https://glowabio.org/project/ssp_landcover/`.

The spatial distribution of fish richness is concentrated in the central part of the study area along the Amazon River and in the north in the Llanos Orientales (Fig. 5). The spatial extent of projected suitable sub-catchments and their richness increases until 2070. Areas without any projected suitable habitats are located around the boundary of the study area, especially in the south-west (Fig. 5). This pattern remains throughout all present and future scenario model projections.

Present suitable habitats of the 219 species endemic to Colombia are mainly located in the north-west of the study area. This covers areas in the Magdalena basin and the western part of the Orinoco basin and has a total extent of 3,433,954 km^2 (Fig. 5). Medium richness is found east of it in the Llanos Orientales in the Orinoco basin and to the west and north towards the coast. For 2070 and especially for SSP5, inland areas are increasingly designated with a low habitat suitability, compared to the present.

In general, new potential suitable habitats emerge in the South, opposed to the present where suitable habitats seem to be fragmented and occur in larger streams. The strongest increase in future fish habitat suitability is found in the tributaries of the Amazon basin (Fig. 6A/B, left). A decrease of habitat suitability is projected at the coastal region in the north-west, and in the south-east of the study area.
A future increase in suitable habitats for endemic species is projected further south and east, while the decline in endemic species richness is particularly strong in the west to north-west of the Magdalena basin (see Fig. 6 A/B, right).
Subtracting the richness map of SSP2 from SSP5 shows the higher richness of SSP5 in most of the study area for both endemic and non-endemic species (Fig. 6 C). Those sub-catchments that show a higher richness under SSP2 than under SSP5 are mainly located in the east/southeast of the study area.

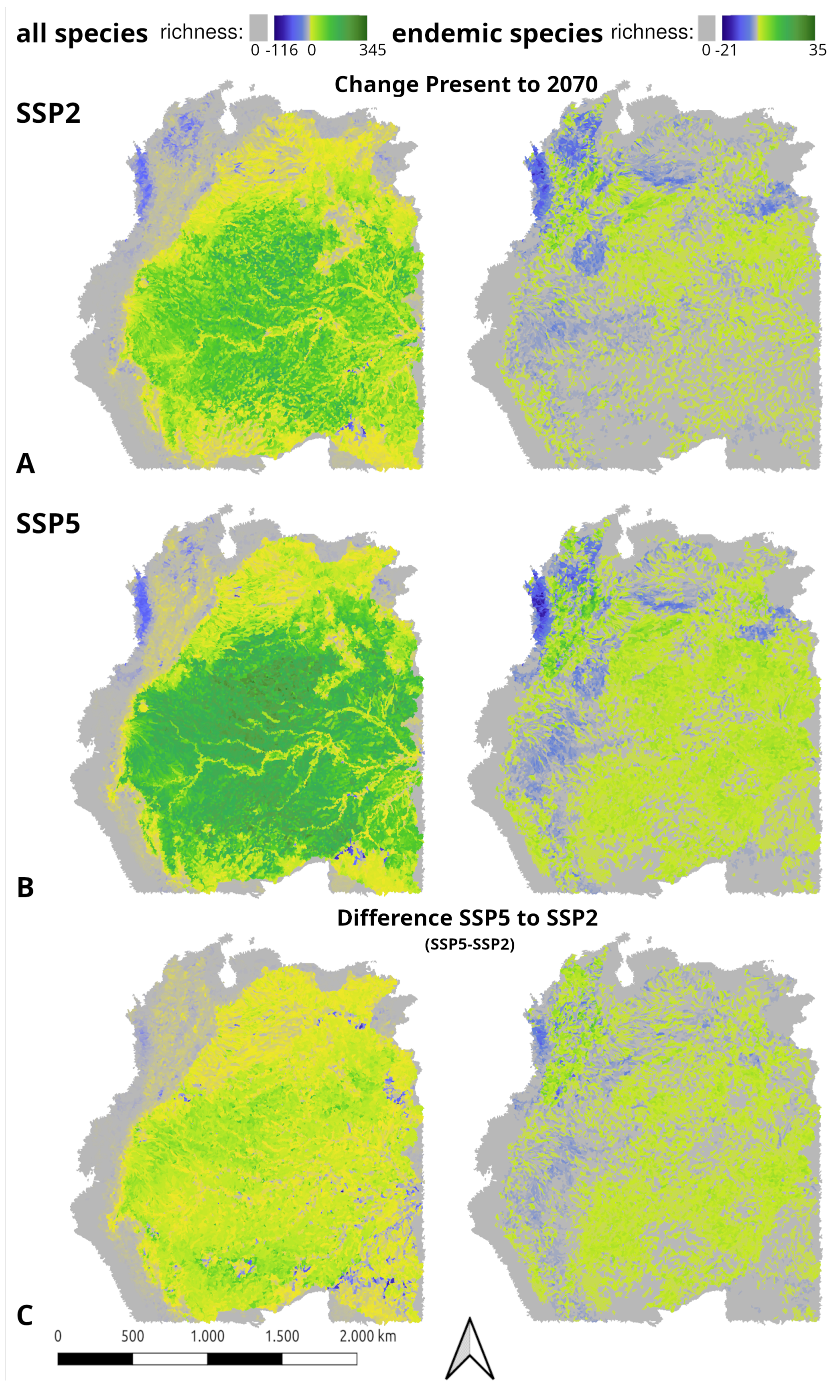

Figure 6: Richness changes. **A-B**: Change in sub-catchment richness (=number of species for which a sub-catchment is suitable) from the present to 2050 and 2070. Blue: decrease; Yellow/Green: increase; Grey refers to no change. **C**: Differences in richness between SSPs 2 and 5. Blue: SSP2 higher species richness; Yellow/Green: SSP5 higher species richness. Grey refers to no change. For interactive maps that allow to toggle and zoom, please see `https://glowabio.org/project/ssp_landcover/`.

3.3 Environmental variables influencing freshwater fish distributions

SDMs project that until 2070, 75% (SSP2) and 77% (SSP5) of the fish species show on average a gain in suitable habitat area of 61% (+296,184 km^2; SSP2) and 91% (+445,561 km^2; SSP5) (see Supplementary Material Table S3).
Future suitable habitats in 2070 are projected to be at higher altitudes for 76% (SSP2) and 78% (SSP5) of the species, with a mean increase in elevation by +22 m (SSP2) and +37 m (SSP5).
The mean sub-catchment flow accumulation (referring to the sub-catchment size) across all species would be 34% (19,096 km^2; SSP2) and 31% (25,195km^2; SSP5) lower in 2070 than in the present. Overall, 27% (SSP2) and 26% (SSP5) of species in 2070 have a lower projected mean flow accumulation than in the present.
SSPs 2 and 5 display similar patterns in suitable habitat area, elevation and flow accumulation change in relation to the mean present forest cover, temperature or precipitation (Figure 7). No significant correlation is observed between species mean present habitat forest cover with future change in suitable habitat area ($r=0.02$, $p=0.5$; mean of SSP2 & 5), mean habitat elevation ($r=-0.003$, $p=0.7$) and mean future flow accumulation ($r=-0.03$, $p=0.2$, Fig. 7).
The majority of species suitable habitats have a mean annual air temperature between 24 and 27 °C. The higher the mean present air temperature within suitable habitats, (i) the greater the variability of change in total future suitable habitat area (Pearson: $r=0.2$; $p<0.05$; mean of SSP2 & 5) and the likelihood of gaining suitable habitat area (Fig. 7 A), (ii) the lower the variability in the change in their future habitat elevation ($r=-0.3$; $p<0.05$; Fig. 7 B) and (iii) the higher the variability in future flow accumulation ($r=-0.16$; $p<0.05$.) with a greater tendency for their future habitats to have lower mean flow accumulation (Fig. 7 C).

The majority of the present suitable habitats of the 1209 Colombian fish species have on average an annual precipitation between 2,000-3,000 mm. The increase in suitable habitat area in the future (Fig. 7 A) is negatively correlated with increasing mean precipitation in the present ($r=-0.09$ $p<0.05$; mean of SSP2 & 5). Habitat elevation change (Fig. 7 B) and increasing mean present habitat precipitation are negatively correlated as well ($r=-0.16$; $p<0.05$). Change in flow accumulation (Fig. 7 C) is not significantly correlated with mean present precipitation ($r=0.04$; $p=0.2$).
The present suitable habitats of endemic species show, compared to non-endemics, significantly lower present temperatures (22.9 °C mean, -2.3°C difference to non-endemic; Welch test: $t=11$, $p<0.05$), higher present precipitation (2,792 mm mean, +178 mm difference; $t=-2$, $p<0.05$), higher elevation (750 m mean, +559 m difference; $t= -14$, $p<0.05$) and lower flow accumulation (9,904 km^2 mean, -86,951 km^2 difference; $t=19$, $p<0.05$), but no significant difference in forest cover ($p<0.05$) (Figure 7, marked grey). Moreover, in the future projections, endemic species loose significantly more suitable habitat area (-9% until 2050

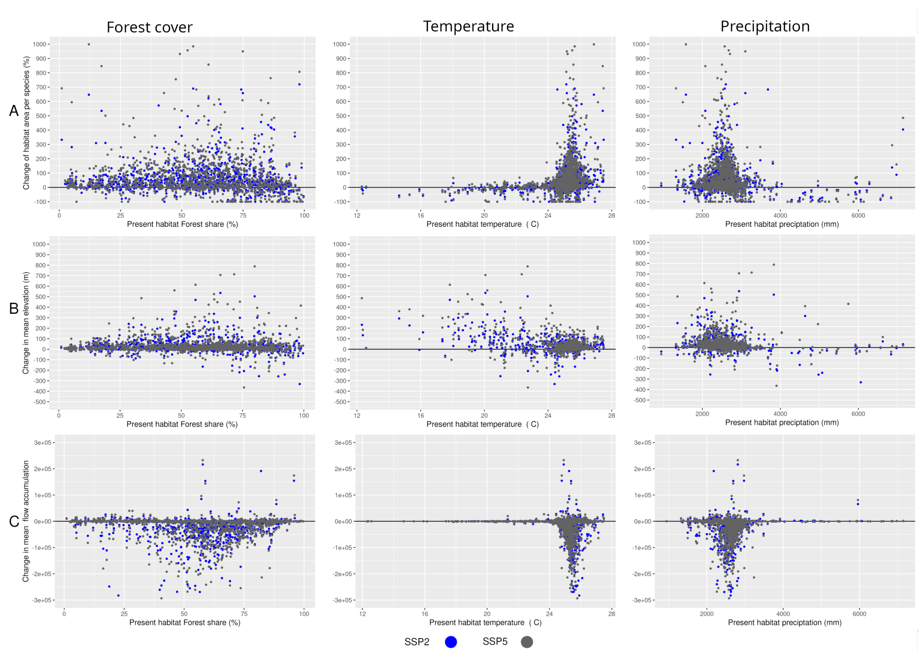

Figure 7: Relationship between mean present forest cover, temperature and precipitation (x-axis) and **A** the change of area projected to be suitable for a species (y-axis). **B** Change in mean elevation of a species (y-axis). **C** Change in mean flow accumulation in a given species' habitat (y-axis). Changes are calculated between present and 2070, for SSP2 (blue) and SSP5 (grey).

and -7% until 2070 (mean SSP2 & SSP5); t=12, p<0.05), move to higher altitudes (+49 m (2050), +80m (2070), t=-4, p<0.05) and the flow accumulation changes less in their suitable habitats (-369 km^2 (2050), -268km^2 (2070); t=-14, p<0.05), compared to non-endemic species.

4 Discussion

We demonstrate the workflow and the potential issues of downscaling tabular SSP land cover projection information to be used in spatial freshwater biodiversity analyses. Our study shows that (i) broad-scale future SSP information can be used to derive spatially explicit land cover estimates, and that (ii) such data can be integrated into a freshwater SDM workflow to provide additional insights into species potential susceptibility to global change. Nevertheless, we point out the challenges regarding (iii) how the scale and spatial resolution of analysis impacts the downscaling, and (iv) the cascading effect of the SSP-derived estimates on the mapped land cover patterns and the modelled fish distribution patterns. Finally, (v) we make recommendations towards future research avenues in freshwater biodiversity modelling approaches that integrate such land cover downscaling exercises.

4.1 Feasibility of the downscaling approach

The conversion of the numerical SSP data into a spatially explicit representation showed how the downscaled land cover data maintained the trends as provided by the initial tabular SSP data, for each land cover category across the entire LAC region. The variation within each land cover category (i.e., crops, forest, pasture and urban, Fig. 3A-D) fluctuated over the period from 2005-2070, corresponding to the specific SSP narrative (Box 1). As expected, the environmental variables and freshwater-specific settings in the LCM contributed towards generating consistent results. We found land cover changes to be prominent along roads and rivers, most probably because of a better accessibility to e.g. deforestation, while changes were less prominent in protected areas. Various direct and indirect drivers in the SSPs, such as population, consumption patterns, dietary preferences, education, technological development or environmental protection efforts, shape the projections of the SSP scenarios and have further influenced the land cover modelling regarding the spatial allocation of land cover grid cells. (see also the detailed maps at `https://glowabio.org/project/ssp_landcover/`). For instance, accessibility is, on the one hand, provided by roads, which are well developed in the north and east of the study area, as well as on the border with Bolivia, compared to the Amazon basin (see Global Roads Inventory Project (GRIP) data set, Meijer et al., 2018). On the other hand, the Amazon river serves as a transportation route for people and goods, which may explain why the majority of land cover changes are projected to occur along waterways in the Amazon basin (Delson, 2008; Barber et al., 2014). The newly-developed land cover data matched the identical 1 km^2 grid as the topography, stream-topology and climate data, thus allowing for a seamless integration of all data types. Following (Hermoso et al., 2011) we aggregated all data to sub-catchments and further harmonized the spatial data acquired from different sources.

4.2 Fish distribution models

The SDMs for the 1209 fish species across Colombia suggest that the amount of future suitable habitat would, on average, increase across all modelled fish species in the study area. While the LCM showed that crop, pasture and urban area coverage would increase, and forest cover would decrease, the SDMs showed the land cover variables had only a minor contribution in the SDMs, whereas climate variables contributed stronger to the modelled future development of suitable habitats. We attribute this to two effects: first, at the given spatial resolution of the sub-catchments, climate tends to be the primary driver of freshwater fish habitat suitability patterns (Friedrichs-Manthey et al., 2020; Maloney et al., 2013; Domisch et al., 2015b; Schmidt et al., 2020), and local effects such as land cover or hydrology contribute to fish habitat suitability at higher resolutions. For instance, Kuemmerlen et al. (2015) showed that land cover had the strongest impact on macroinvertebrate communities at the catchment scale. At even finer spatial scales, hydrological parameters tend to contribute stronger to the models (Friedrichs-Manthey et al., 2020). Second, the aggregation of single land cover pixels to sub-catchments, opposed to the continuous climate surface, yielded the effect that the contribution of local land cover effects flattened in the models, as identified by the minor contribution of land cover variables in the models. As a result, the SDMs remained mostly climate-driven where warm-adapted species would increase their future suitable habitat in terms of area, towards higher elevations and lower flow accumulation than species adapted to cooler temperatures. For the 219 endemic fish species of Colombia, SDMs projected a reduction in suitable habitat, and a shift to higher elevations. These projections are in line with previous observations and modelling studies regarding climate-change impacts e.g. in Europe (Domisch et al., 2013; Comte et al., 2013; Jarić et al., 2019).

4.3 Scale-dependency in the land cover projections

Though the general workflow proved its applicability and shows a high potential, we identify several aspects which are critical for further development. This relates especially to the scale-dependency in the downscaling and the subsequent species distribution modelling, and possible uncertainties deriving from the LCM.

First, by collapsing the 28 initial ESA land cover categories to the four plus two additional SSP categories, we diminished the level of detail in the present-day land cover data and projections, hence simplifying the species-environment relationships. Employing a higher number of land cover categories could therefore provide a more detailed present-day species-environment relationship, however at the expense of additional uncertainties in the future projections (Schuwirth et al., 2019) since the future location of such specific land cover types remains less certain than the broad types.

Second, given the spatial extent of the study area, with 513 Mha (5,130,000 km^2), we opted to use large sub-catchments as spatial units (187 km^2 on average) as to balance the size of the study area and the number of species in this workflow demonstration in terms of computation efficiency, and to compute spatial predictions across the entire study area. The proportions of each land cover category were aggregated across sub-catchments, similar to fish occurrences, such that the detailed local information at a given fish occurrence location is not used in the model. This aggregation to the sub-catchments flattens specific effects emanating from land cover at this spatial resolution (Domisch et al., 2015b; Friedrichs-Manthey et al., 2020).

Third, and again related to scale, is the spatial resolution of the environmental information of 1 km^2 which can be still considered relatively coarse within the freshwater realm, such that small-scale effects are not being detected (Kuemmerlen et al., 2014). For instance, land cover changes in the riparian zone have shown to impact on freshwater species habitats (Allan, 2004; Leitão et al., 2018). In this regard, the new Hydrography90m dataset (Amatulli et al., 2022) allows to increase the level of detail by several orders of magnitude, thus being able to reflect the influence of local land cover effects on freshwater species given the high resolution of the stream network and the corresponding sub-catchments. This would mean that future approaches could explore possible land cover changes on a higher resolution, producing several portfolios under specific (technical downscaling) assumptions that are then being used as input for e.g. SDMs. Diversifying the land cover change portfolio by means of several possible trajectories would thus have the potential to uncover possible, emerging patterns given local land cover changes.

Fourth, while the sub-catchments and the stream-topological data is routed across the network, the SDMs did not account for spatial effects which can improve the predictions considerably (Domisch et al., 2019). In terms of land cover, the routing of upstream land cover information provides the possibility to mimic the downstream transportation effects as river systems are considered cumulative collectors of upstream processes and effects (Kuemmerlen et al., 2014; Radinger et al., 2016; Schmidt et al., 2020). Spatially-explicit SDMs are however computationally intense and hence were not considered in this workflow demonstration. Similarly, instead of a random pseudo-absence extraction for the SDMs, another possibility could be to extract the pseudo-absences starting at certain network distances from the point occurrences, analogous to circular buffers in terrestrial or marine SDMs (Iturbide et al., 2015). That said, we acknowledge that the SDM analyses are for demonstration purpose only, and for a detailed analysis of how the Colombian fish fauna might respond to global change, a more detailed and in-depth modelling analysis would be required.

4.4 Future recommendations

In line with the aforementioned challenges, we aim to give recommendations when downscaling future land cover information. We aimed to showcase the downscaling, however management-oriented analyses would most likely apply a higher spatial resolution which would then allow to allocate local land cover changes more precisely than in our example. This would also include information regarding the hydrographic network and hydropower dams, and incorporate the hydrological dimension in the land cover estimates (which however requires also future hydrological estimates to be included in the downscaling). The LCM require land cover change-rates and so-called transition matrices from one time step to another, which broadly follow the given SSP storylines (O'Neill et al., 2017) where a given land cover category changes to another category. The calibration of such change-rates is non-trivial and requires more than a "one-size-fits-all" approach as showcased in our example, especially regarding pasture which is known to lead to a high variability depending on different definitions of pasture (Oliveira et al., 2020). In combination with the SSP storylines which are by definition broad, the resulting future land cover estimates do not necessarily match the projected development of specific regions. Future modifications in the downscaling would therefore need to take e.g. country-specific information regarding the economic growth and demographic development into account to pinpoint the possible changes in SSP-suggested changes more accurately (Frame et al., 2018). Moreover, we follow Otero et al. (2020) and Otero et al. (2022) and recommend extending the SSP narratives by e.g. a "no growth" scenario "SSP0" as to capture the full range of possible future trajectories.

Despite some limitations, our example workflow demonstrate the feasibility of integrating future land cover data in freshwater biodiversity models. We conclude that the spatial scale and resolution is key towards further improvements, which would allow highlighting local-to-regional changes in possible land cover changes. We encourage potential users to apply the workflow in custom study areas and provide the code and instructions at `https://github.com/Anni-R-T/Code_Brunner_Downscaling-future-land-cover-scenarios`.

5 Acknowledgements

We wish to thank Thomas Tomiczek for his help in preparing the fish data, and Rodrigo Rivero Castro, Leticia Lima and Leticia Hissa for the great help in setting up Dinamica EGO. We thank the Yale Center for Research Computing for guidance and use of the research computing infrastructure, and we acknowledge funding by the German Federal Ministry of Education and Research (BMBF grant agreement number no. 033W034A). SD acknowledges funding by the Leibniz Competition (J45/2018). We thank an anonymous reviewer for the constructive comments on an earlier version of the manuscript.

References

Allan, J. D., 2004. "Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems". In: *Annual Review of Ecology, Evolution, and Systematics* 35.1, pp. 257–284. DOI: 10.1146/annurev.ecolsys.35.120202.110122.

Amatulli, G., Garcia Marquez, J., Sethi, T., Kiesel, J., Grigoropoulou, A., Üblacker, M. M., Shen, L. Q., Domisch, S., 2022. "Hydrography90m: a new high-resolution global hydrographic dataset". In: *Earth System Science Data* 14.10, pp. 4525–4550. DOI: 10.5194/essd-14-4525-2022.

Amatulli, G., Domisch, S., Tuanmu, M.-N., Parmentier, B., Ranipeta, A., Malczyk, J., Jetz, W., 2018. "A suite of global, cross-scale topographic variables for environmental and biodiversity modeling". In: *Scientific Data* 5.1, p. 180040. DOI: 10.1038/sdata.2018.40.

Bahrenberg, G., Giese, E., Nipper, J., 2017. *Statistische Methoden in der Geographie. Band 1, Univariate und bivariate Statistik*. Print.

Balian, E. V., Segers, H, Martens, K, Lévêque, C, 2007. "The freshwater animal diversity assessment: an overview of the results". In: *Freshwater animal diversity assessment*, pp. 627–637.

Barbarossa, V., Huijbregts, M. A., Beusen, A. H., Beck, H. E., King, H., Schipper, A. M., 2018. "FLO1K, global maps of mean, maximum and minimum annual streamflow at 1 km resolution from 1960 through 2015". In: *Scientific Data* 5.1, p. 180052. DOI: 10.1038/sdata.2018.52.

Barber, C. P., Cochrane, M. A., Souza, C. M., Laurance, W. F., 2014. "Roads, deforestation, and the mitigating effect of protected areas in the Amazon". In: *Biological Conservation* 177, pp. 203–209. DOI: 10.1016/j.biocon.2014.07.004.

Barbet-Massin, M., Thuiller, W., Jiguet, F., 2010. "How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models?" In: *Ecography* 33.5, pp. 878–886.

Bond, N., Thomson, J., Reich, P., Stein, J., 2011. "Using species distribution models to infer potential climate change-induced range shifts of freshwater fish in south-eastern Australia". In: *Marine and Freshwater Research* 62.9, pp. 1043–1061. DOI: 10.1071/MF10286.

Bonham-Carter, G. F., 1994. "CHAPTER 9 - Tools for Map Analysis: Multiple Maps". In: *Geographic Information Systems for Geoscientists*. Ed. by G. F. Bonham-Carter. Vol. 13. Computer Methods in the Geosciences. Pergamon, pp. 267–337. DOI: 10.1016/B978-0-08-042420-0.50014-6.

Bontemps, S., Defourny, P., Radoux, J., Van Bogaert, E., Lamarche, C., Achard, F., Mayaux, P., Boettcher, M., Brockmann, C., Kirches, G., Zulkhe, M., Kalogirou, V., Seifert, F. M., Arino, O., 2013. "Consistent Global Land Cover Maps For Climate Modelling Communities: Current Achievements Of The ESA' Land Cover CCI". In: *ESA Living Planet Symposium*. Vol. 722. ESA Special Publication, p. 62.

Breiman, L., Friedman, J., Stone, C. J., Olshen, R., 1984. *Classification and Regression Trees*. Chapman and Hall/CRC.

Buisson, L., Blanc, L., Grenouillet, G., 2008. "Modelling stream fish species distribution in a river network: the relative effects of temperature versus physical factors". In: *Ecology of Freshwater Fish* 17.2, pp. 244–257. DOI: 10.1111/j.1600-0633.2007.00276.x.

Canhos, D. A. L., Almeida, E. A., Assad, A. L., Cunha Bustamante, M. M. d., Canhos, V. P., Chapman, A. D., Giovanni, R. D., Imperatriz-Fonseca, V. L., Lohmann, L. G., Maia, L. C., https://specieslink.net/, others. Data downloaded on 12.05.2020 at, 2022. "speciesLink: rich data and novel tools for digital assessments of biodiversity". In: *Biota Neotropica* 22.

Chen, M., Vernon, C. R., Graham, N. T., Hejazi, M., Huang, M., Cheng, Y., Calvin, K., 2020. "Global land use for 2015–2100 at 0.05° resolution under diverse socioeconomic and climate scenarios". In: *Scientific Data* 7.1, p. 320. DOI: 10.1038/s41597-020-00669-x.

Comte, L., Buisson, L., Daufresne, M., Grenouillet, G., 2013. "Climate-induced changes in the distribution of freshwater fish: observed and predicted trends". In: *Freshwater Biology* 58.4, pp. 625–639. DOI: 10.1111/fwb.12081.

Cowx, I. G., Portocarrero Aya, M., 2011. "Paradigm shifts in fish conservation: moving to the ecosystem services concept". In: *Journal of Fish Biology* 79.6, pp. 1663–1680. DOI: 10.1111/j.1095-8649.2011.03144.x.

da Silva Cruz, J., Blanco, C. J. C., de Oliveira Júnior, J. F., 2022. "Modeling of land use and land cover change dynamics for future projection of the Amazon number curve". In: *Science of The Total Environment* 811, p. 152348. DOI: https://doi.org/10.1016/j.scitotenv.2021.152348.

Dala-Corte, R. B., Giam, X., Olden, J. D., Becker, F. G., Guimarães, T. d. F., Melo, A. S., 2016. "Revealing the pathways by which agricultural land-use affects stream fish communities in South Brazilian grasslands". In: *Freshwater Biology* 61.11, pp. 1921–1934. DOI: 10.1111/fwb.12825.

Danielson, J. J., Gesch, D. B., 2011. *Global multi-resolution terrain elevation data 2010 (GMTED2010)*. US Department of the Interior, US Geological Survey Washington, DC, USA.

Delson, R. M., 2008. "Navigation: Inland Navigation in Amazonia". In: *Encyclopaedia of the History of Science, Technology, and Medicine in Non-Western Cultures*. Ed. by H. Selin. Dordrecht: Springer Netherlands, pp. 1–6. DOI: 10.1007/978-94-007-3934-5_9920-1.

Distler, T., Schuetz, J. G., Velásquez-Tibatá, J., Langham, G. M., 2015. "Stacked species distribution models and macroecological models provide congruent projections of avian species richness under climate change". In: *Journal of Biogeography* 42.5, pp. 976–988.

Domisch, S., Amatulli, G., Jetz, W., 2015a. "Near-global freshwater-specific environmental variables for biodiversity analyses in 1 km resolution". In: *Scientific data* 2.1, pp. 1–13.

Domisch, S., Araújo, M. B., Bonada, N., Pauls, S. U., Jähnig, S. C., Haase, P., 2013. "Modelling distribution in E uropean stream macroinvertebrates under future climates". In: *Global Change Biology* 19.3, pp. 752–762.

Domisch, S., Friedrichs, M., Hein, T., Borgwardt, F., Wetzig, A., Jähnig, S. C., Langhans, S. D., 2019. "Spatially explicit species distribution models: A missed opportunity in conservation planning?" In: *Diversity and Distributions* 25.5, pp. 758–769.

Domisch, S., Jaehnig, S. C., Simaika, J. P., Kuemmerlen, M., Stoll, S., 2015b. "Application of species distribution models in stream ecosystems: the challenges of spatial and temporal scale, environmental predictors and species occurrence data". In: *Fundamental and Applied Limnology*, pp. 45–61.

DoNascimiento, C., Herrera-Collazos, E. E., Herrera-R., G. A., Ortega-Lara, A., Villa-Navarro, F. A., Oviedo, J. S. U., Maldonado-Ocampo, J. A., 2017. "Checklist of the freshwater fishes of Colombia: a Darwin Core alternative to the updating problem". In: *ZooKeys* 708, pp. 25–138. DOI: 10.3897/zookeys.708.13897.

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., Naiman, R. J., Prieur-Richard, A.-H., Soto, D., Stiassny, M. L. J., Sullivan, C. A., 2006. "Freshwater biodiversity: importance, threats, status and conservation challenges". In: *Biological Reviews* 81.2, pp. 163–182. DOI: 10.1017/S1464793105006950.

Ebi, K. L., Hallegatte, S., Kram, T., Arnell, N. W., Carter, T. R., Edmonds, J., Kriegler, E., Mathur, R., O'Neill, B. C., Riahi, K., Winkler, H., Van Vuuren, D. P., Zwickel, T., 2014. "A new scenario framework for climate change research: background, process, and future directions". In: *Climatic Change* 122.3, pp. 363–372. DOI: 10.1007/s10584-013-0912-3.

Elith, J., Leathwick, J. R., 2009. "Species Distribution Models: Ecological Explanation and Prediction Across Space and Time". In: *Annual Review of Ecology, Evolution, and Systematics* 40.1, pp. 677–697. DOI: 10.1146/annurev.ecolsys.110308.120159.

Eraso, N. R., Armenteras-Pascual, D., Alumbreros, J. R., 2013. "Land use and land cover change in the Colombian Andes: dynamics and future scenarios". In: *Journal of Land Use Science* 8.2, pp. 154–174. DOI: 10.1080/1747423X.2011.650228.

ESA, 2017. *Land Cover CCI Product User Guide Version 2. Tech. Rep.* Available at: URL: https://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf.

fishnet2, 2020. "Fish data used in this study obtained from the Field Museum of Natural History, Museum of Vertebrate Zoology, University of Kansas Museum of Natural History, Tulane University Museum of Natural History (Accessed through the Fishnet2 Portal, www.fishnet2.org, 2020-05-12)". In.

Frame, B., Lawrence, J., Ausseil, A.-G., Reisinger, A., Daigneault, A., 2018. "Adapting global shared socio-economic pathways for national and local scenarios". In: *Climate Risk Management* 21, pp. 39–51. DOI: 10.1016/j.crm.2018.05.001.

Fricko, O., Havlik, P., Rogelj, J., Klimont, Z., Gusti, M., Johnson, N., Kolp, P., Strubegger, M., Valin, H., Amann, M., Ermolieva, T., Forsell, N., Herrero, M., Heyes, C., Kindermann, G., Krey, V., McCollum, D. L., Obersteiner, M., Pachauri, S., Rao, S., Schmid, E., Schoepp, W., Riahi, K., 2017. "The marker quantification of the Shared Socioeconomic Pathway 2: A middle-of-the-road scenario for the 21st century". In: *Global Environmental Change* 42, pp. 251–267. DOI: 10.1016/j.gloenvcha.2016.06.004.

Friedrichs-Manthey, M., Langhans, S. D., Hein, T., Borgwardt, F., Kling, H., Jähnig, S. C., Domisch, S., 2020. "From topography to hydrology—The modifiable area unit problem impacts freshwater species distribution models". In: *Ecology and evolution* 10.6, pp. 2956–2968.

Galford, G., Filho, B., Sonter, L., Laporte, N., 2015. "Will Passive Protection Save Congo Forests?" In: *PloS one* 10, e0128473. DOI: 10.1371/journal.pone.0128473.

GBIF, 2020. "GBIF.org (05 March 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.pizntb". In.

GDAL/OGR contributors, 2021. *GDAL/OGR Geospatial Data Abstraction software Library*. Open Source Geospatial Foundation.

Gollnow, F., Göpel, J., Viana Hissa, L. deBarros, Schaldach, R., Lakes, T., 2018. "Scenarios of land-use change in a deforestation corridor in the Brazilian Amazon: combining two scales of analysis". In: *Regional Environmental Change* 18.1, pp. 143–159. DOI: 10.1007/s10113-017-1129-1.

Hao, T., Elith, J., Guillera-Arroita, G., Lahoz-Monfort, J. J., 2019. "A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD". In: *Diversity and Distributions* 25.5, pp. 839–852. DOI: 10.1111/ddi.12892.

He, F., Zarfl, C., Bremerich, V., David, J. N., Hogan, Z., Kalinkat, G., Tockner, K., Jähnig, S. C., 2019. "The global decline of freshwater megafauna". In: *Global Change Biology* 25.11, pp. 3883–3892.

Hermoso, V, Linke, S, Prenda, J, Possingham, H., 2011. "Addressing longitudinal connectivity in the systematic conservation planning of fresh waters". In: *Freshwater Biology* 56.1, pp. 57–70.

Hoskins, A. J., Bush, A., Gilmore, J., Harwood, T., Hudson, L. N., Ware, C., Williams, K. J., Ferrier, S., 2016. "Downscaling land-use data to provide global 30 estimates of five land-use classes". In: *Ecology and Evolution* 6.9, pp. 3040–3055. DOI: https://doi.org/10.1002/ece3.2104.

iDigBio, 2020. "iDigBio.org (13 March 2020) Species Occurrence Download, data available at https://www.iDigBio.org". In.

Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., Gutiérrez, J. M., 2015. "A framework for species distribution modelling with improved pseudo-absence generation". In: *Ecological Modelling* 312, pp. 166–174. DOI: https://doi.org/10.1016/j.ecolmodel.2015.05.018.

Jarić, I., Lennox, R. J., Kalinkat, G., Cvijanović, G., Radinger, J., 2019. "Susceptibility of European freshwater fish to climate change: Species profiling based on life-history and environmental characteristics". In: *Global Change Biology* 25.2, pp. 448–458. DOI: https://doi.org/10.1111/gcb.14518.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., Kessler, M., 2017. "Climatologies at high resolution for the earth's land surface areas". In: *Scientific Data* 4.1, p. 170122. DOI: 10.1038/sdata.2017.122.

Karger, D., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R., Zimmermann, N., Linder, H., Kessler, M., 2018. *Data from: Climatologies at high resolution for the earth's land surface areas.* Ed. by Dryad. DOI: 10.5061/dryad.kd1d4.

Kuemmerlen, M., Schmalz, B., Cai, Q., Haase, P., Fohrer, N., Jähnig, S. C., 2015. "An attack on two fronts: predicting how changes in land use and climate affect the distribution of stream macroinvertebrates". In: *Freshwater Biology* 60.7, pp. 1443–1458. DOI: 10.1111/fwb.12580.

Kuemmerlen, M., Schmalz, B., Guse, B., Cai, Q., Fohrer, N., Jähnig, S. C., 2014. "Integrating catchment properties in small scale species distribution models of stream macroinvertebrates". In: *Ecological Modelling* 277, pp. 77–86. DOI: 10.1016/j.ecolmodel.2014.01.020.

Leitão, R. P., Zuanon, J., Mouillot, D., Leal, C. G., Hughes, R. M., Kaufmann, P. R., Villéger, S., Pompeu, P. S., Kasper, D., Paula, F. R. de, Ferraz, S. F. B., Gardner, T. A., 2018. "Disentangling the pathways of land use impacts on the functional structure of fish assemblages in Amazon streams". In: *Ecography* 41.1, pp. 219–232. DOI: 10.1111/ecog.02845.

Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., et al., 2019. "Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution". In: *Scientific data* 6.1, pp. 1–15.

Lo, M., Reed, J., Castello, L., Steel, E. A., Frimpong, E. A., Ickowitz, A., 2020. "The Influence of Forests on Freshwater Fish in the Tropics: A Systematic Review". In: *BioScience* 70.5, pp. 404–414. DOI: 10.1093/biosci/biaa021.

Lu, Z., Yuan, K.-H., 2010. "Welch's t test". In: pp. 1620–1623. DOI: 10.13140/RG.2.1.3057.9607.

Maloney, K., Weller, D., Michaelson, D., Ciccotto, P., 2013. "Species Distribution Models of Freshwater Stream Fishes in Maryland and Their Implications for Management". In: *Environmental Modeling and Assessment* 18, pp. 1–12. DOI: 10.1007/s10666-012-9325-3.

Martin, Y., Van Dyck, H., Dendoncker, N., Titeux, N., 2013. "Testing instead of assuming the importance of land use change scenarios to model species distributions under climate change". In: *Global Ecology and Biogeography* 22.11, pp. 1204–1216. DOI: 10.1111/geb.12087.

McCullagh, P., Nelder, J., 1989. *Generalized Linear Models.* 2nd. London: Chapman and Hall.

Meijer, J. R., Huijbregts, M. A. J., Schotten, K. C. G. J., Schipper, A. M., 2018. "Global patterns of current and future road infrastructure". In: *Environmental Research Letters* 13.6, p. 064006. DOI: 10.1088/1748-9326/aabd42.

Melo-Merino, S. M., Reyes-Bonilla, H., Lira-Noriega, A., 2020. "Ecological niche models and species distribution models in marine environments: A literature review and spatial analysis of evidence". In: *Ecological Modelling* 415, p. 108837.

Neteler, M., Bowman, M. H., Landa, M., Metz, M., 2012. "GRASS GIS: A multi-purpose open source GIS". In: *Environmental Modelling & Software* 31, pp. 124–130.

Oliveira, J., Campbell, E. E., Lamparelli, R. A., Figueiredo, G. K., Soares, J. R., Jaiswal, D., Monteiro, L. A., Vianna, M. S., Lynd, L. R., Sheehan, J. J., 2020. "Choosing pasture maps: An assessment of pasture land classification definitions and a case study of Brazil". In: *International Journal of Applied Earth Observation and Geoinformation* 93, p. 102205. DOI: 10.1016/j.jag.2020.102205.

O'Neill, B. C., Kriegler, E., Ebi, K. L., Kemp-Benedict, E., Riahi, K., Rothman, D. S., van Ruijven, B. J., van Vuuren, D. P., Birkmann, J., Kok, K., Levy, M., Solecki, W., 2017. "The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century". In: *Global Environmental Change* 42, pp. 169–180. DOI: 10.1016/j.gloenvcha.2015.01.004.

OpenStreetMap contributors, 2017. *Planet dump retrieved from https://planet.osm.org.*

Otero, I., Farrell, K. N., Pueyo, S., Kallis, G., Kehoe, L., Haberl, H., Plutzar, C., Hobson, P., García-Márquez, J., Rodríguez-Labajos, B., Martin, J.-L., Erb, K.-H., Schindler, S., Nielsen, J., Skorin, T., Settele, J., Essl, F., Gómez-Baggethun, E., Brotons, L., Rabitsch, W., Schneider, F., Pe'er, G., 2020. "Biodiversity policy beyond economic growth". In: *Conservation Letters* 13.4, e12713. DOI: 10.1111/conl.12713.

Otero, I., Rigal, S., Pereira, L., Kim, H., Grêt-Regamey, A., 2022. "A degrowth scenario for biodiversity? Some methodological avenues and a call for collaboration". In.

Phillips, S. J., Anderson, R. P., Schapire, R. E., 2006. "Maximum entropy modeling of species geographic distributions". In: *Ecological Modelling* 190.3, pp. 231–259. DOI: 10.1016/j.ecolmodel.2005.03.026.

Popp, A., Calvin, K., Fujimori, S., Havlik, P., Humpenöder, F., Stehfest, E., Bodirsky, B. L., Dietrich, J. P., Doelmann, J. C., Gusti, M., Hasegawa, T., Kyle, P., Obersteiner, M., Tabeau, A., Takahashi, K., Valin, H., Waldhoff, S., Weindl, I., Wise, M., Kriegler, E., Lotze-Campen, H., Fricko, O., Riahi, K., Vuuren, D. P. van, 2017. "Land-use futures in the shared socio-economic pathways". In: *Global Environmental Change* 42, pp. 331–345. DOI: 10.1016/j.gloenvcha.2016.10.002.

Porfirio, L. L., Harris, R. M. B., Lefroy, E. C., Hugh, S., Gould, S. F., Lee, G., Bindoff, N. L., Mackey, B., 2014. "Improving the use of species distribution models in conservation planning and management under climate change". In: *PloS one* 9.11, e113749–e113749. DOI: 10.1371/journal.pone.0113749.

R Core Team, 2020. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing. Vienna, Austria.

Radinger, J., Hölker, F., Horký, P., Slavík, O., Dendoncker, N., Wolter, C., 2016. "Synergistic and antagonistic interactions of future land use and climate change on river fish assemblages". In: *Global Change Biology* 22.4, pp. 1505–1522. DOI: 10.1111/gcb.13183.

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J., Kidd, K. A., MacCormack, T. J., Olden, J. D., Ormerod, S. J., Smol, J. P., Taylor, W. W., Tockner, K., Vermaire, J. C., Dudgeon, D., Cooke, S. J., 2019. "Emerging threats and persistent conservation challenges for freshwater biodiversity”. In: *Biological Reviews* 94.3, pp. 849–873. DOI: 10.1111/brv.12480.

Reis, R., Albert, J., Di Dario, F., Mincarone, M., Petry, P., Rocha, L., 2016. “Fish biodiversity and conservation in South America”. In: *Journal of Fish Biology* 89. DOI: 10.1111/jfb.13016.

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O’Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Da Silva, L. A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J. C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., Tavoni, M., 2017. “The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview”. In: *Global Environmental Change* 42, pp. 153–168. DOI: 10.1016/j.gloenvcha.2016.05.009.

Rivero Castro, R., 2019. “Land Cover Change under Future Scenarios of Zero-deforestation Commitment Implementation in South America”. Available at: http://hdl.handle.net/10125/66239. MA thesis.

Rodrigues, H., Soares-Filho, B., 2018. “A Short Presentation of Dinamica EGO.” In: *Geomatic Approaches for Modeling Land Change Scenarios. Lecture Notes in Geoinformation and Cartography.* Ed. by M. C. Olmedo, M. Paegelow, J. Mas, F. Escobar. Springer, Cham, pp. 493–498.

Rodriguez-Iturbe, I., Rinaldo, A., 2001. *Fractal river basins: chance and self-organization.* Cambridge University Press.

Schmidt, H., Radinger, J., Teschlade, D., Stoll, S., 2020. “The role of spatial units in modelling freshwater fish distributions: Comparing a subcatchment and river network approach using MaxEnt”. In: *Ecological Modelling* 418, p. 108937. DOI: 10.1016/j.ecolmodel.2020.108937.

Schuwirth, N., Borgwardt, F., Domisch, S., Friedrichs, M., Kattwinkel, M., Kneis, D., Kuemmerlen, M., Langhans, S. D., Martínez-López, J., Vermeiren, P., 2019. “How to make ecological models useful for environmental management”. In: *Ecological Modelling* 411, p. 108784.

Soares Filho, B., Filho, L., Cerqueira, G., Araújo, W., 2003. “Simulating the spatial patterns of change through the use of the DINAMICA model”. In: *Anais XI SBSR.*

Soares Filho, B. S., Rodrigues, H. O., Costa, W, 2009. “Modeling environmental dynamics with Dinamica EGO”. In: *Centro de Sensoriamento Remoto. Universidade Federal de Minas Gerais. Belo Horizonte, Minas Gerais* 115.

Taylor, K. E., Stouffer, R. J., Meehl, G. A., 2012. “An overview of CMIP5 and the experiment design”. In: *Bulletin of the American meteorological Society* 93.4, pp. 485–498.

Thuiller, W., Georges, D., Engler, R., Breiner, F., 2020. *biomod2: Ensemble Platform for Species Distribution Modeling.* R package version 3.4.6.

Thuiller, W., Lafourcade, B., Engler, R., Araújo, M. B., 2009. "BIOMOD – a platform for ensemble forecasting of species distributions". In: *Ecography* 32.3, pp. 369–373. DOI: 10.1111/j.1600-0587.2008.05742.x.

Tootchi, A., Jost, A., Ducharne, A., 2018. *Multi-source global wetland maps combining surface water imagery and groundwater constraints.* data set. Sorbonne Université, Paris, France. DOI: 10.1594/PANGAEA.892657.

Tóth, R., Czeglédi, I., Kern, B., Erős, T., 2019. "Land use effects in riverscapes: Diversity and environmental drivers of stream fish communities in protected, agricultural and urban landscapes". In: *Ecological Indicators* 101, pp. 742–748. DOI: 10.1016/j.ecolind.2019.01.063.

UNEP-WCMC, I., 2020. "UNEP-WCMC and IUCN (2020), , Protected Planet: The World Database on Protected Areas (WDPA). Online, accessed 11/2020, Cambridge, UK: UNEP-WCMC and IUCN. Available at: www.protectedplanet.net". In.

van Vuuren, D. P., Stehfest, E., Gernaat, D. E., Doelman, J. C., van den Berg, M., Harmsen, M., de Boer, H. S., Bouwman, L. F., Daioglou, V., Edelenbosch, O. Y., Girod, B., Kram, T., Lassaletta, L., Lucas, P. L., van Meijl, H., Müller, C., van Ruijven, B. J., van der Sluis, S., Tabeau, A., 2017. "Energy, land-use and greenhouse gas emissions trajectories under a green growth paradigm". In: *Global Environmental Change* 42, pp. 237–250. DOI: 10.1016/j.gloenvcha.2016.05.008.

W.W.F., 2020. *Living Planet Report 2020. Bending the curve of biodiversity loss: a deep dive into freshwater.* Ed. by R. Almond, G. M., T. Petersen. WWF, Gland, Switzerland.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., Bates, P. D., 2017. "A high-accuracy map of global terrain elevations". In: *Geophysical Research Letters* 44.11, pp. 5844–5853.

A Supplementary Material

Table S1: The original ESA legend numbers and categories, the single categories clustered into larger categories, and the area (in million hectare) of the clustered/reclassified categories and for each of the five SSP marker scenarios. The clustered land-cover categories are crops (yellow shading); forest (dark green); pasture (light green); built up area / urban (red); other natural land (brown); water (blue).

No Group; ESA legend	group name; ESA legend name	reclassified	SSP1	SSP2	SSP3	SSP4	SSP5
1	**crops**	**167.421**	**173.72**	**164.63**	**160.25**	**161.29**	**167.32**
10	Cropland, rainfed						
12	Cropland tree or shrub cover						
20	Cropland, irrigated or post-flooding						
30	Mosaic cropland(>50%) /natural vegetation (tree, shrub, herbaceous cover) (<50%)						
2	**forest**	**891.96**	**741.07**	**902.16**	**922.76**	**840.06**	**1095.78**
50	Tree cover, broadleaved, evergreen, closed to open (>15%)						
60	Tree cover, broadleaved, deciduous, closed to open (>15%)						
61	Tree cover, broadleaved, deciduous, closed (>40%)						
62	Tree cover, broadleaved, deciduous, open (15-40%)						
70	Tree cover, needleleaved, evergreen, closed to open (>15%)						
80	Tree cover, needleleaved, deciduous, closed to open(>15%)						
90	Tree cover, mixed leaf type (broadleaved and needleleaved)						
3	**pasture**	**307.12**	**539.64**	**557.03**	**549.05**	**548.7**	**518.40**
11	Cropland herbaceous cover						
110	Mosaic herbaceous cover (>50%) / tree and shrub (<50%)						
153	Sparse herbaceous cover (<15%)						
130	Grassland						
4	**built up area (urban)**	**5.09**	**6.89**	**-**	**8.64**	**5.54**	**6.19**
190	Urban areas						
5	**other natural land**	**638.28**	**-**	**-**	**-**	**-**	**-**
40	Mosaic natural vegetation (tree, shrub, herbaceous cover) (>50%) / cropland(<50%)						
100	Mosaic tree and shrub (>50%) / herbaceous cover (<50%)						
120	Shrubland						
121	Evergreen shrubland						
122	Deciduous shrubland						
150	Sparse vegetation (tree, shrub, herbaceous cover) (<15%)						
160	Tree cover, flooded, fresh or brakish water						
170	Tree cover, flooded, saline water						
180	Shrub or herbaceous cover, flooded, fresh/saline/brakishwater						
200	Bare areas						
220	Permanent snow and ice						
6	**water**	**34.54**	**-**	**-**	**-**	**-**	**-**
210	Water bodies						

Input for Dinamica Ego

Initial Year Map Dinamica EGO requires a map that serves as the spatial basis for the first cycle of the LCM (=initial year map). The reclassified LC map of 2005 served as the initial year map in Dinamica EGO.

Transition Matrix
The transition matrix depicts the percent change of area per time step from one land-cover category to another.

In a first step we calculated the rates of change per ten-year time steps per land-cover category and for each SSP (Figure S2 A) as:

$$\text{rt} = \frac{A_{t2}}{A_{t1}} - 1$$

A= area, t= time

We then applied these change rates to the areas derived from the reclassified 2005 LC map (Figure S2 B). Then distinct changes from one category to another (e.g. which specific percentage of forest becomes crop land) need to be specified (Figure S2 C). For determining the next time step transformations (i.e. a new transformation matrix), we calculated the next set of SSP change rates (e.g. 2010-2020) and applied these to the areas that resulted from the first transition calculation (Figure S2 D). For the water category, we set the rate of change between 2005 and 2100 to zero and hence kept it constant, because no information or projections were available regarding changes in water bodies.

As a result, we created separate sets of transition matrices for each SSP scenario. Hence, the specific land-cover change rates for each of the five SSP scenarios is hereby reflected in the LCM.

		Forest area SSP1 2005	Forest area SSP1 2010	change rate	
A	Calculation of SSP change rate:	750	700	-6.67%	
		Forest area 2005	change rate	Forest area 2010	
B	Application of first change rate:	900	-6.67%	840	
C	Splitting of change rate:	Forest to Crop Forest to Pasture Forest to Urban	2.5% 1.5% 0.5%		
		Forest area 2010	SSP change rate (2010-2020)		Forest area 2020
D	Application of second change rate:	840	-3.65%		810

Figure S1: Exemplary illustration of the transition matrix creation process - *note that the numbers shown here are only examples and do not correspond to actual values.* The numbers in gray shading represents the change rate that is deducted from the SSP data (row A) and is applied to the area derived from the reclassified LC maps (green) (B). It is prepared for input into Dinamica EGO by splitting it into more specific change rates (C). The area which results from the application of the change rates (marked blue) is then the base area for the transition matrix of the next time step (D).

Weights of evidence

The weights of evidence (WOE) are calculated in a Bayesian framework and give the probability of transition from one category to another, depending on the influence of the environmental variables (i.e., spatial determinants) on the location of changes (Bonham-Carter, 1994; Soares Filho et al., 2009). Two maps (initial and final year) are needed to track the LCCs occurring between these years and to link them to the environmental conditions at the location of the change, thus calculating/training the WOE accordingly. We used the reclassified 2000 and 2010 LC maps for the WOE training to match the 10-year intervals of the SSP data and to account for changes before and after the 2005 base year.
When observing the historical trend no decline is registered in urban areas, but in three of the SSPs negative development for urban areas is described. For instance, for the specific change of urban area being transformed into another category, we calculated the WOE "backwards" using 2010 as the initial and 2000 as the final year, using a model in Dinamica EGO provided by Rodrigo Rivero (Rivero Castro, 2019).
In addition to environmental factors, the actual spatial distance between land cover types is another element that played a role in the calculation of WOE. For example, when calculating the probability of forest changing into cropland, we took the distance to already existing cropland into account.
The same environmental maps used for the calculation of the WOE are used as input for the LCM, such that the transformation allocations can be made. Since an adjustment of the WOE with environmental projections for the future is complex, the environmental data remains static for the entire modelled period, implying an assumption of constant environmental conditions over the modelling period. The distance is reassessed with every modelling round from the latest map created.

Dinamica also provides measures of correlation between spatial variables and shows whether a variable is significant for the LCM or not. We took the Crammer coefficient into account in order to minimize the impact of correlated variables, potentially distorting the results. The coefficient describes the correlation between spatial variables from 0 (independent) to 1 (high correlation). We removed variables with a Crammer coefficient greater than 0.45 from the "weight list", following Galford et al., 2015. In case of pair.wise highly correlated variables, we only kept one, i.e. the one marked more significant over the others. We removed variables not deemed significant. This evaluation is done per transition. Accordingly, the environmental variables of e.g. the transformation "forest to crops" show different correlation and significance values than the transformation forest to pasture.

Patcher Expander

The "patcher" and "expander" functions were designed to reflect the spatial change pattern: The expander function determines the expansion of existing patches of a certain land-cover category. The patcher function forms new patches of a category where it did not occur before, using a seedling mechanism. In addition, the relationship between expansion and formation of new patches is indicated by determining the percentage of transition by expansion (Soares Filho et al., 2003; Soares Filho et al., 2009). The patcher and expander values were also derived with a Dinamica EGO model by Rodrigo Rivero (Rivero Castro, 2019).
We applied one change prior to integrating the patcher and expander values into the LCM: we changed the percentage of transition by expansion for changes from "forest to urban area" from 0.4 to 0.6 to have urban areas slightly rather expand than form new patches, as first test runs showed (too) many new single pixels of urban area. Overall, the expansion of areas outweighs the formation of new patches.

Validation of the LCM

To test how well the LCM method used projected LCCs, we used a function in Dinamica EGO to compare projected maps to observed maps. We created a second, so-called validation model that followed the principle of the created LCM for calculating WOE and patcher and expander values. However, the rates of the transformation matrix were derived from observed maps so that the modelled results could later be compared with the observed ones. The initial year of the validation model was 2000 and projections for 2005 and 2015 were generated to match the length of the first two periods of "my" LCM. However, the initial year is five years earlier because there was no observed map for 2020 at the time of processing. The changes projected between 2000 and 2005 and 2015 were then compared to the observed changes via a model in Dinamica EGO, and the similarity of the changes was reported (to check if the LCM projected the same or similar changes between the two years as seen in the observed maps). The output is then the proportion (in percent) of agreement between projected and observed.

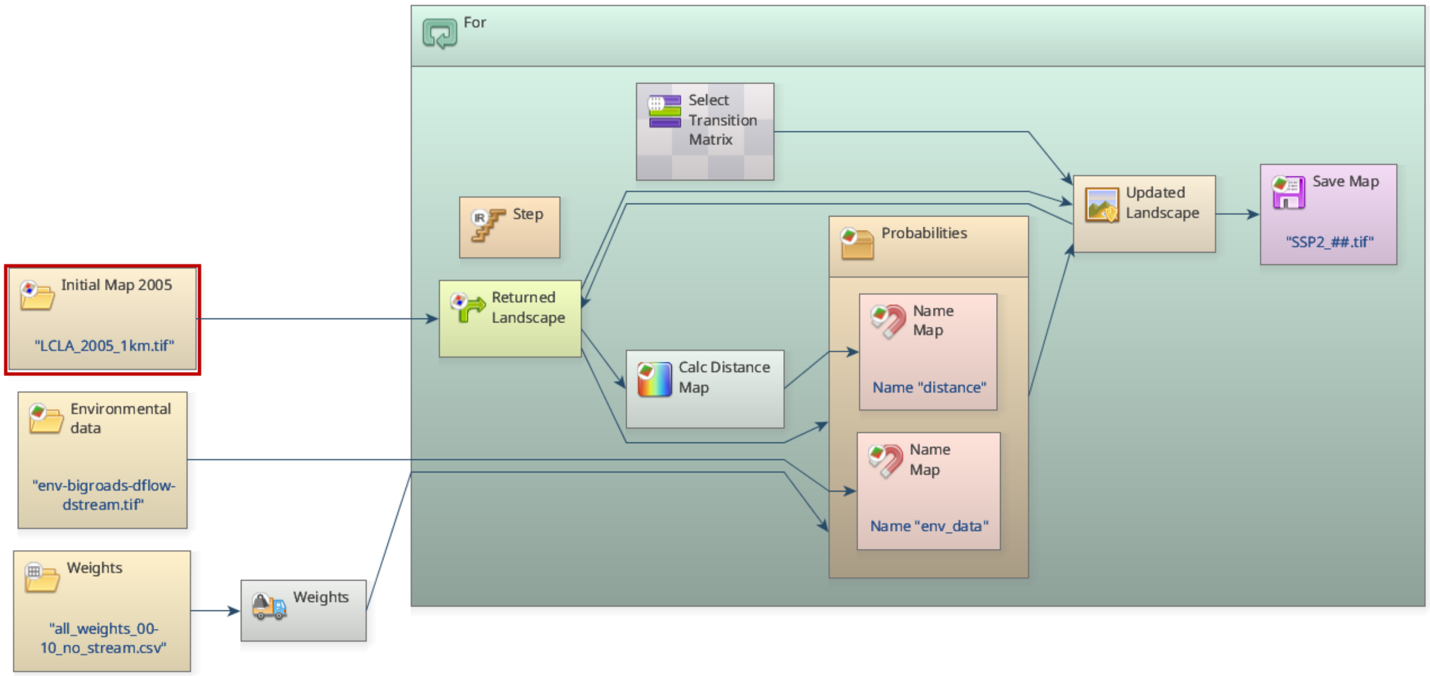

Figure S2: Exemplary illustration of the LCM Structure in Dinamica for SSP2. Left: Model Inputs - initial map, stacked environmental data and pre-calculated weights of evidence. The large blue contains the elements of the Model-Loop. For each round a new transition matrix is selected and the distances between LC-categories is reiterated (Calc distance map)

Environmental statistics for the projected habitats

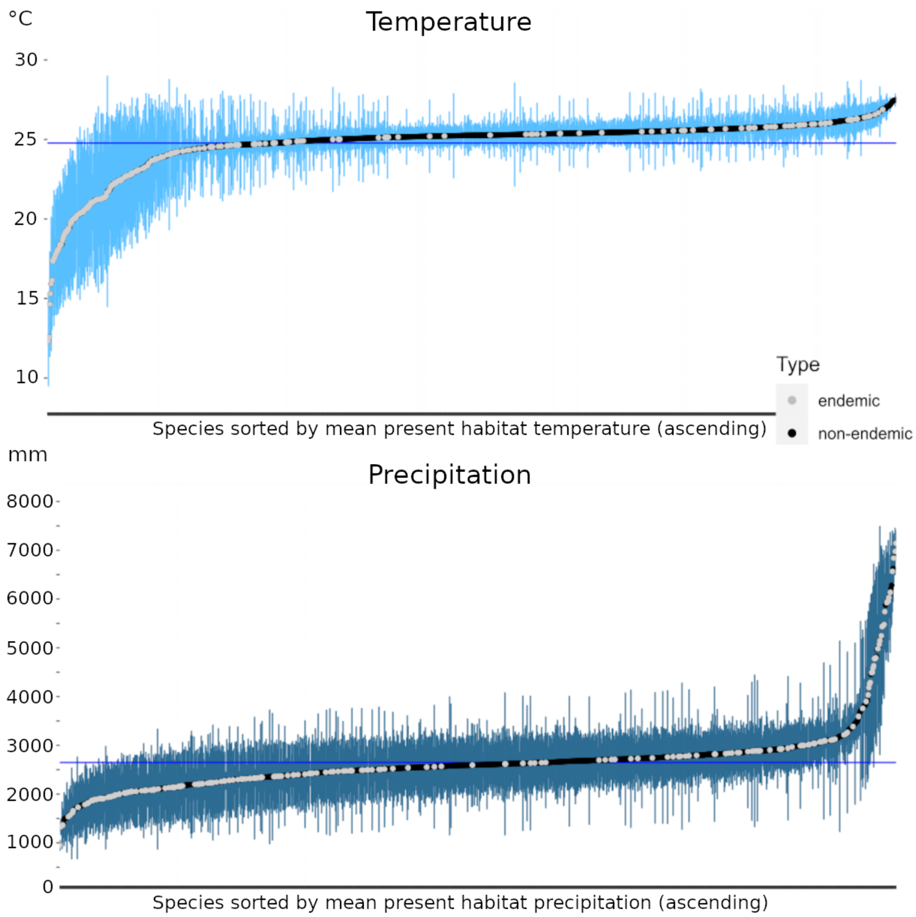

Figure S3: Mean present temperature and precipitation (y-axis) across the modelled suitable habitats for all 1,209 species (x-axis). The species are sorted by mean temperature in their present projected habitats. Endemic species are marked in grey. The standard deviation is shown in blue vertical lines. The dark blue horizontal line represents the overall mean.

The mean temperature across all present projected fish habitats is 24.8°C (Fig. S3), increasing to 26.7°C and 28°C for SSPs 2 and 5, respectively (Supplementary Material, Table S2). Species whose suitable habitats are projected in cooler temperatures in the present, show a greater standard deviation (Fig. S3) than species whose mean projected present habitat temperature is above the overall mean.

The mean precipitation projected across all present-day modelled fish habitats is 2,646 mm (Fig. S3), increasing to 2,791 mm (SSP2) and 2,815 mm (SSP5) for 2070 (Supplementary Material, Table S2).

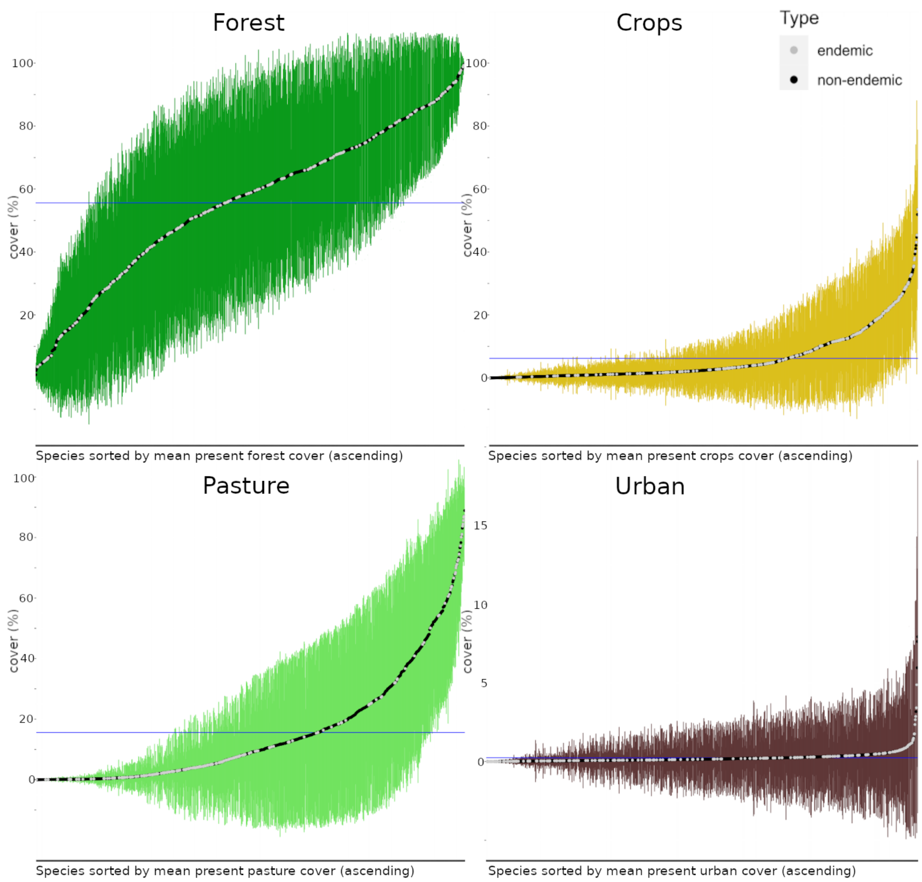

Figure S4: Mean present land-cover proportions for all species. Mean share of of the corresponding land-cover category across all habitats (y-axis) across the modelled suitable habitats for all 1,209 species (x-axis). The species are sorted by their mean land cover share in their present projected habitats. *Note that for an improved visualisation the x-axis has no numeric scaling.* Endemic species are marked in grey. Colors represent the standard deviation of the land-cover category per species: Dark green - forest; Yellow - crops; Light green - pasture; Red/brown - urban. The blue horizontal line represents the overall mean.

The average proportion of forest within all sub-catchments projected to host suitable fish habitat is 55.6% (Figure S4, blue horizontal line). The proportion of forest in the sub-catchment occupied by a species varies given the standard deviation. In 2070, the mean share of forest across all 1,209 species is 51.4% and

58% for SSP2 and SSP5, respectively (see Table S2 in Supplementary Material for the future coverages).

Cropland occupies a mean of 6.2% of a sub-catchment projected as suitable habitat in the present (Fig. S4). In 2070, the projected mean habitat cropland proportion is 7.3% for SSP2 and 7.8% for SSP5. The maximum crop share in the suitable habitats increases by 16% and 8 % until 2070, for SSPs 2 and 5, respectively (see Supplementary Material Table S2 for maximum values).

Under present conditions, a mean of 15.6% within suitable fish habitats is pasture (Fig. S4), and is projected to decrease to 14.2% and 14% in 2070 under SSPs 2 and 5, respectively. The maximum value of sub-catchment pasture cover is 9% higher than in the present year for both SSPs in 2070. The higher the mean proportion of either crop or pasture coverage in a given species suitable habitat, the higher the standard deviation of the respective land-cover proportions.

Urban cover in suitable fish habitats is on average 0.2 % in the present (Fig. S4), and is projected to be between 0.3% (SSP5) and 0.4% (SSP2) in the future (Table S2). In the present, the averaged maxima of urban areas across sub-catchments is 7.93%. SSP2 projects urban cover up to 28 % of a subcatchment area for 2070. For SSP5, the maximum urban share is projected to be 10.2% in 2070. The higher the mean present urban cover, the greater the standard deviation.

Table S2: Mean and maximum values of land-cover shares, temperature and precipitation across the habitats/sub-catchments of all 1,209 species. Projected for the present as well as for 2050 and 2070 for SSP2 and SSP5. Further changes from the present to 2050 and the present to 2070 for each variable are listed.

			2070		present to 2070	
		present	**SSP2**	**SSP5**	**SSP2**	**SSP5**
forest	mean	55.6	57.4	58.0	1.1	1.3
(%)	max	99.7	99.7	100.0	0.2	0.0
crops	mean	6.2	7.3	7.8	1.1	1.6
(%)	max	53.5	69.4	61.5	15.9	8.0
pasture	mean	15.6	14.2	14.0	-1.4	-1.6
(%)	max	88.9	98.1	98.2	9.2	9.2
urban	mean	0.2	0.4	0.3	0.1	0.0
(%)	max	7.9	28.0	10.2	20.1	2.2
mean annual temperature	mean	24.8	26.7	28.1	1.9	3.3
°C	max	27.5	29.4	31.2	2.0	3.7
annual precipitation	mean	2646	2791	2816	145	170
mm	max	7145	7723	8168	578	1023

Table S3: Changes in suitable habitat area, elevation and flow accumulation from the present to 2070 by mean change and number of species loosing or gaining in the respective area. Species showing no elevation or flow accumulation values in the future are the fish no longer finding suitable habitats.

		Present to 2070	
		SSP2	**SSP5**
area	Mean area change per species	+ 296,184 km^2 (61%)	+ 445,561 km^2 (91%)
	Species gaining area	906 (75%)	934 (77%)
	Species loosing area	302 (25%)	275 (23%)
	Area remaining the same	1	-
	NA	-	-
elevation	Mean elevation change per species	+22 m (7%)	+37m (13%)
	Species moving to higher elevation	919 (76%)	945 (78%)
	Species moving to lower elevation	262 (22%)	232 (19%)
	Elevation remaining the same	1	-
	NA	27 (2%)	32 (3%)
flow accumulation	Mean flow change per species	-19,096 km^2 (34%)	-25,195 km^2 (31%)
	Species with higher flow accumulation	331 (27%)	318 (26%)
	Species with lower flow accumulation	850 (70%)	859 (71%)
	Flow remaining the same	1	-
	NA	27 (2%)	32 (3%)

Table S4: Statistics of present and 2070 projections for the 1,209 fish species: The mean, standard deviation and maximum richness averaged across all suitable habitats. The total number of habitats projected to be suitable or not suitable and the number of species for which no suitable habitats are projected in the future (extinct species).

		2070				
	present	**SSP1**	**SSP2**	**SSP3**	**SSP4**	**SSP5**
mean richness	98	165	166	167	167	203
stdev richness	120	145	145	146	146	155
max richness	605	660	661	660	660	666
suitable sub-catchments	24675	24936	24955	24949	24954	25021
unsuitable sub-catchments	2331	2070	2051	2057	2052	1985
extinct species		28	27	30	28	32

Table S5: Statistics of present and 2070 projections for the endemic fish species: The mean, standard deviation and maximum richness averaged across all suitable habitats. The total number of habitats projected to be suitable or not suitable and the number of species for which no suitable habitats are projected in the future (extinct species).

		2070				
	present	**SSP1**	**SSP2**	**SSP3**	**SSP4**	**SSP5**
mean richness	5	5	5	5	5	6
stdev richness	10	10	10	10	10	10
max richness	86	86	84	85	85	87
suitable sub-catchments	16750	18551	18629	18704	18696	21321
unsuitable sub-catchments	10256	8455	8377	8302	8310	5685
extinct species		16	16	16	16	20