

An empirical assessment of the extent to which the realized niche of bird species is truncated

Ivan Mattatia

rue de la Filature 19, 1227 Carouge, Switzerland

E-mail : ivan.mattatia@unil.ch

1st step project

directed by Prof. Antoine Guisan

supervised by Mathieu Chevalier

Master in Behaviour, Evolution and Conservation,
specialisation Geosciences, Ecology and Environment

University of Lausanne, 2019

Abstract

Species distribution models (SDM) are widely used in conservation biology to identify conservation priorities or predicting the effect of changes in landuse or climate. Based on the ecological niche, a key assumption of SDM is that relevant environmental gradients have been sampled in a way that the realized niche faithfully represents the set of conditions that the species can tolerate. However, there are several cases where this assumption can be violated e.g. when a geographically restricted dataset is used to calibrate the model or when the margins of the realized niche coincide with the margins of the available environmental space, i.e. the conditions available on Earth. Here we show an evidence for truncation in bird species realized niche. A majority of 49 bird species niche showed truncation, which suggest they can tolerate more conditions than the actual available conditions on Earth. Furthermore, this issue was proved to be more acute for generalist species with larger niche. We anticipate our exploratory study to be a starting point to develop methods to estimate truncation and experiment tools in order to take this truncation into account in species distribution models for spatial predictions or predictions using climate change scenarios.

Keywords

Niche truncation ; Environmental conditions ; Species Distribution Models ; Climate change scenarios ; projections

Introduction

Nowadays, climate change is widely recognized as a reality (Stocker et al., 2013). This phenomenon is a great concern for many fields, among them, conservation biology. It is largely accepted by the scientific community that global climate change is a threat to biodiversity and ecosystems which are not able to adapt to the actual stress that is beyond the natural climate change that occurred in the past (Omann, Stocker, & Jäger, 2009; Wilson & Peter, 1988). Several response to climate change effects have been observed such as adaptive evolution or tolerance, spatial or temporal shift, with a failure to respond by one of these means leading to population range reduction and local, regional, or even complete extinction (Holt, 1990; Nogués-Bravo et al., 2018). With the hope to avoid species extinctions and develop adapted conservation strategies for the different species, species distribution models (SDM) and environmental niche modelling (ENM) are useful tools. SDM are based on the notion of environmental niche conceptualized by Hutchinson (Hutchinson, 1953) and are built by statistically linking species occurrence (presence-absence or presence-only) and environmental data to identify an n-dimensional space representing a set of suitable environmental and biotic conditions for the species: the realized niche. SDMs have been developed in the mid-1980s (Pecchi et al., 2019) and are widely used since the early 20th century for a large panel of purposes including quantifying species ecological niches, informing conservation decision and identifying conservation priorities (Loiselle et al., 2003) or predicting the effect of a range of changes such as land use or climate (Hallstan, 2011). A key assumption of SDMs is that relevant environmental gradients have been sampled in a way that the realized niche faithfully represents the set of conditions that the species can tolerate.

However, there are several cases where this assumption can be violated e.g. when a geographically restricted dataset is used to calibrate the model or when the margins of the realized niche coincide with the margins of the available environmental space, i.e. the conditions available on Earth. It is also important to keep in mind that niche truncation can also be caused by restricted access to certain area due to species dispersal limitations (Soberón & Nakamura, 2009) or biotic interactions (Peterson, Cobos, & Jiménez-García, 2018). However these two latter situations will not be assessed in this study.

In the first case, when the scale of the study is a geographically restricted area, a truncation in the data will occur. This is the case for instance if, in the idea of predicting future distribution of one species in Switzerland, only the species occurrences in this country are taken into account. The study will thus not capture the whole realized niche that one species occupies and will then generate bias in the estimate of future species distribution (Titeux et al., 2017). This truncation induce other effect as for instance an overestimation of local extinction rate (Barbet-Massin, Thuiller, & Jiguet, 2010) or more conservative scenarios for projections of distribution changes (Thuiller, Brotons, Araújo, & Lavorel, 2004). But this situation can easily

be avoided by using distribution data beyond the boundaries of the study area (Titeux et al., 2017). Taking into account data from the neighbourhood allows to capture the response of the species for climatic conditions that are expected within the area in the future.

The second case where the SDM assumption is violated is linked to the available conditions on Earth. It happens when the margins of the niche of interest coincide with the margins of available environmental conditions on Earth. We hypothesize that it induces truncation in the niche since all the conditions that the species could tolerate are not observed. Since no solution is proposed in the literature to palliate to this issue, the predictions are only based on actual environmental conditions and the species behaviour is not known nor taken into account for future climate predictions. This is a big issue when extrapolation is used to model species distribution with non-analogous conditions to predict future changes in the distribution (Thuiller et al., 2004), species extinctions or potential of invasive species (Smolik et al., 2010).

To our knowledge, there is no estimation of the extent to which the limits of the environmental conditions that are available on our planet are causing niche truncation. The aim of the study is to analyse through an exploratory approach whether or not realized niche are truncated and to provide a quantitative assessment to which extend this realized niche is truncated for bird species due to environmental restrictions.

In order to answer this question, datasets for 49 birds species coming from two different sources, the Global Biodiversity Information Facility (GBIF) and the International Union for Conservation of Nature (IUCN), were analysed. 19 bioclimatic variables and species occurrences at worldwide scale were used to estimate (i) the world background envelope characterizing the range of environmental conditions available on Earth and (ii) the environmental envelope of each species characterizing the environmental conditions experienced by species across their geographical range. The realized niche envelope was compared to the environmental background envelope (i.e. environmental conditions available on Earth) to estimate truncation. This truncation was defined as the percentage of the realized niche that exceeds the environmental background margins.

Since a large niche has more chance to reach the margins of the environmental background, we hypothesize a higher truncation for species with a larger realized niche. Since generalist species tend to have larger niche than specialist species (Pagani-Núñez et al., 2019), this hypothesis suggest higher truncation for generalist species. A second hypothesis is that more or less truncation is to be expected depending on groups species and their characteristic traits, e.g. climate change sensibility, diet, thermic tolerance. Since some traits can influence species distribution (Aubin et al., 2018) we expect them to potentially influence niche truncation as well.

Our results suggest a niche truncation for the vast majority of the species tested. It appears to be a tendency for species with the largest niche to have the more truncation in the realized niche.

Methods

All analyses were performed using the R environment software (R Core Team, 2018).

Data

Analyses were performed using data for 49 birds species which were extracted from two major databases that are commonly used for species distribution modelling: IUCN and GBIF databases. IUCN data are geographic range maps that are built using existing occurrence records together with expert knowledge of species ecology (Hawkins, Rueda, & Rodríguez, 2008). These data appear in the form of polygons that overlay a geographical map and within which the species is assumed present with a high probability (Alhajeri & Fourcade, 2019; Herkt, Skidmore, & Fahr, 2017). On the other hand, GBIF data are compiled records of georeferenced observation (Fig 1b) coming from different sources such as citizen science programs, museum collections and monitoring programs (Edwards, 2004). IUCN and GBIF data are commonly used in conservation biology e.g. to estimate extinctions probabilities under different climate change scenarios using niche modelling algorithms, or to quantify species ecological niches. We here assume that because these data are plentiful and cover a large geographical extent encompassing large environmental gradients, they would make it possible to characterize the full-realized niche of species. Differences were nevertheless expected between GBIF and IUCN datasets in their assessment of the realized niche owing to the strength and the weaknesses of the two datasets. For instance, since they are georeferenced observations, GBIF data are generally precise. However, as they come from different sources such as citizen science programs, a spatial bias is expected with lower sampling in areas with adverse conditions such as mountains, desert areas or far away from walking tracks. We thus expect the niche to be underestimated, especially in the extreme parts of the niche. In the other hand, IUCN data have a more coarse resolution which suggest that there are some points in the range where the species is not found. But in the contrary of GBIF data, IUCN data are not spatially biased. Since there are no evidence for a dataset to be better than the other and that both of them are widely used in SDM we decided to use the two datasets independently, as suggested by Alhajeri and Fourcade (Alhajeri & Fourcade, 2019).

Environmental layer (raster data) for the world were directly downloaded from WorldClim online database (“Bioclimatic variables | WorldClim - Global Climate Data,” n.d.). This data contains bioclimatic data for 19 variables representing annual trends, seasonality and extreme or limiting environmental factors. These values are the average for the years 1970 to 2000 and are derived from monthly temperature and rainfall values. Compiling annual trends but also taking variations throughout the year into account, they are commonly used in niche modelling analysis. The spatial resolution used is 2.5 minutes, which is about 4.5km at the equator for each cell. This somewhat coarse resolution is recommended to perform range-map

based analyses (Hurlbert & Jetz, 2007; Lam & Quattrochi, 1992) and was chosen to avoid putting too much emphasis on extreme climatic conditions that may distort species envelopes.

Defining the environmental space

In order to obtain a two-dimensional environment space for the 19 bioclimatic variables, a Principal Component Analysis (PCA) was performed on the bioclimatic variables. This statistical procedure convert the correlated bioclimatic variables in a set of uncorrelated values that are the principal components. The first principal component is defined in a way that it

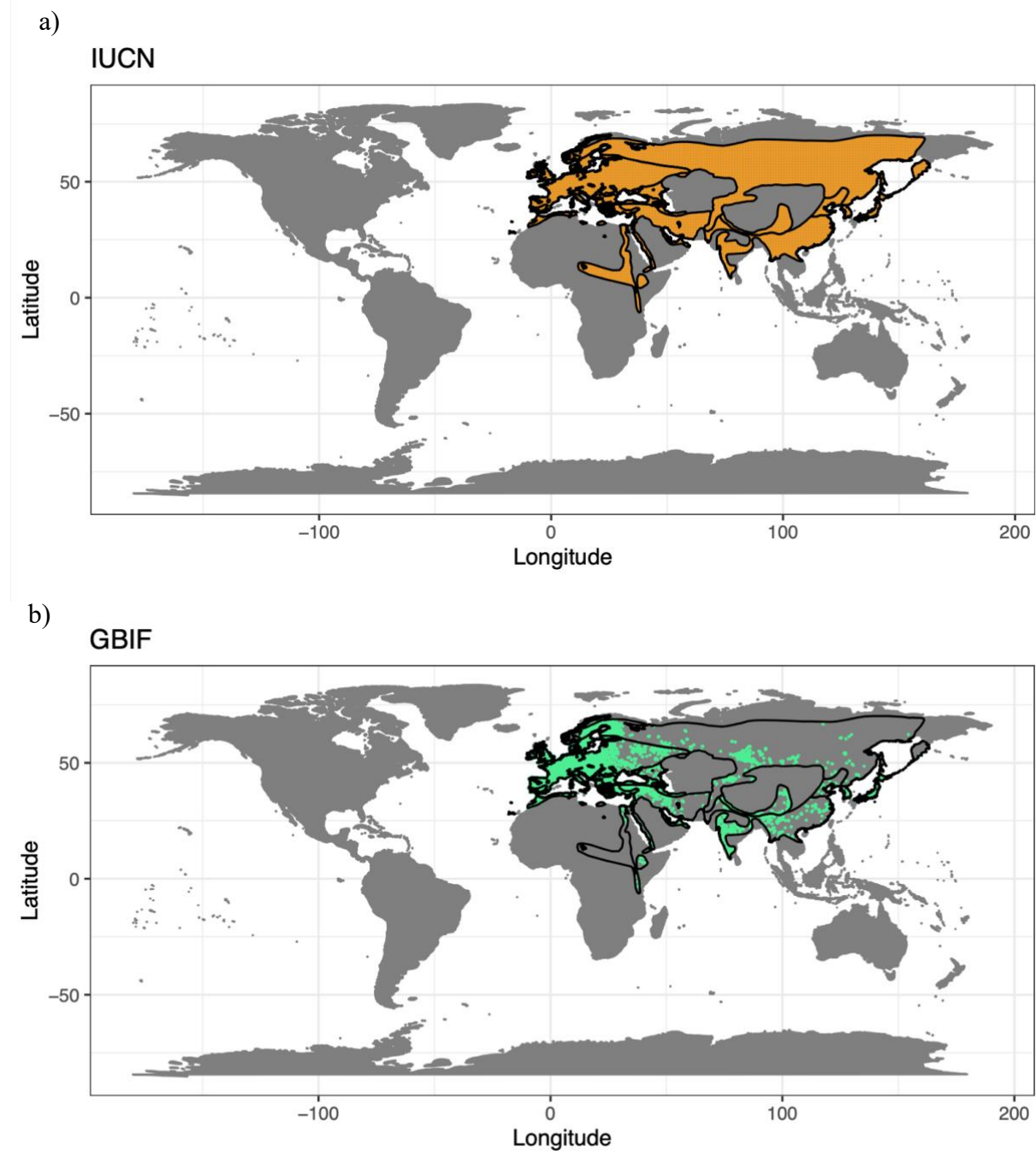


Figure 1: a) Geographic range map for Eurasian sparrowhawk (*Accipiter nisus*) obtained with IUCN data and b) georeferenced observation for Eurasian sparrowhawk obtained with GBIF data.

account for the largest variability as possible in the data and the next components have, in turn the highest variance possible as long as it is orthogonal to the preceding component. The environmental background envelope as well as species niche envelopes were then computed in this two-dimensional space, defined by the first two PCA axis, that were selected as representing the majority of the variability. 51% of the variance is explained by the first axis of the PCA and 26% by the axis. The two first axis thus explain 77% of the variance of the environmental data.

Background envelope

The environmental background was computed using a kernel density estimate. The kernel density estimate is a way to estimate the probability density function of random variables and thus estimate the hypervolume defined by the environmental variables. Since this estimate delimits the boundaries of the hypervolume (Blonder, Lamanna, Violle, & Enquist, 2014), it delineates the environmental background envelope.

Species envelopes

Two Species niche envelopes were computed for each species, one using IUCN data, the other using GBIF data. The IUCN polygons and GBIF data were first loaded. GBIF data were previously cleaned using the CoordinateCleaner R package (Zizka et al., 2019). Then, environmental data were extracted for 10,000 points sampled randomly in the range of IUCN data. The points of the two databases were then projected on the PCA axis. A kernel density estimate was again used to estimate the boundaries around the occurrence points and delineate the species niche envelope. Since this method is based on density occurrence, where only 99% of occurrences were used (limit with this method), the species niche envelope can thus go beyond the environmental background envelope. For instance, if a high density of occurrence is found at the margins of the background envelope, the kernel density will delimit the boundaries around our occurrence points further from the points and thus outside of the background envelope. These are these particular area where the niche envelope goes beyond the environmental background envelope that was considered as a niche truncation mark.

Niche truncation

Two approaches were used to estimate the extent to which the realized niche of species might be truncated by the available environmental space. These two methods are based on the part of the niche envelope that overlay the environmental background envelope, and both the area and the perimeter of these surface were used to estimate the truncation. The area and perimeter of the truncated part of the niche were standardized using a percentage of the total envelope (see formula in Fig 2).

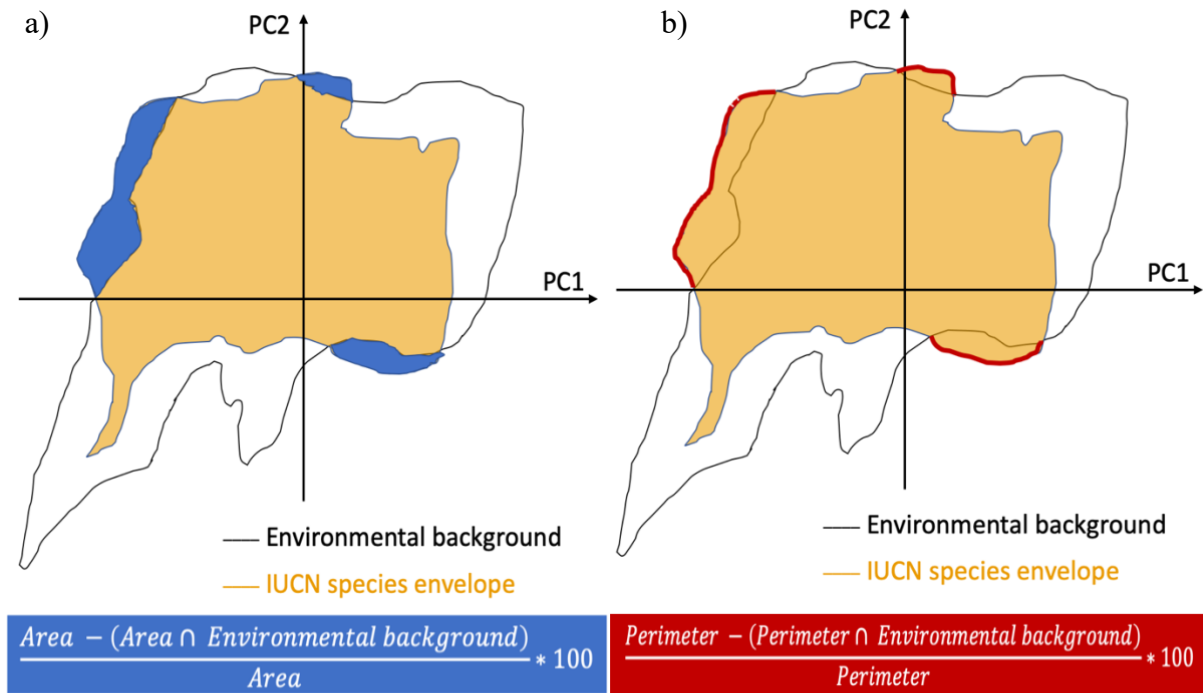


Figure 2: Estimation of niche truncation based on the a) area (in blue) and b) the perimeter (in red) of the surface of the niche species envelope that goes beyond the environmental background envelope.

Testing differences in truncation assessments

A Wilcoxon-Mann-Whitney test was performed between truncation obtained from IUCN and GBIF data for both area and perimeter approach in order to compare the two datasets.

In a second time, an analysis was performed taking into consideration the size of the realized niche to verify the first hypothesis that higher truncation is expected for species with a larger niche. A plot of the estimated truncation in function of niche size was drawn in order to establish whether or not a relation between truncation and niche size exists. The relation was analysed with a linear model.

In order to analyse the second hypothesis which stipulate that a difference in niche truncation is to be expected between groups of species and their characteristics traits, the truncation was analysed in functions of groups of species. In this case, groups of species were based on the taxonomic order of the species which was one of the traits that we could find with our actual data and that is a first hint if we expect other species traits to influence the extent of niche truncation.

Results

The superposition of the two realized niche envelopes obtained with GBIF and IUCN data and the environmental background envelope allows to notice that they differ considerably between GBIF and IUCN data for a majority of the 49 species analysed (Tab1, Fig3). The envelope computed with IUCN data are bigger than when computed with GBIF data for all species (Tab. 1).

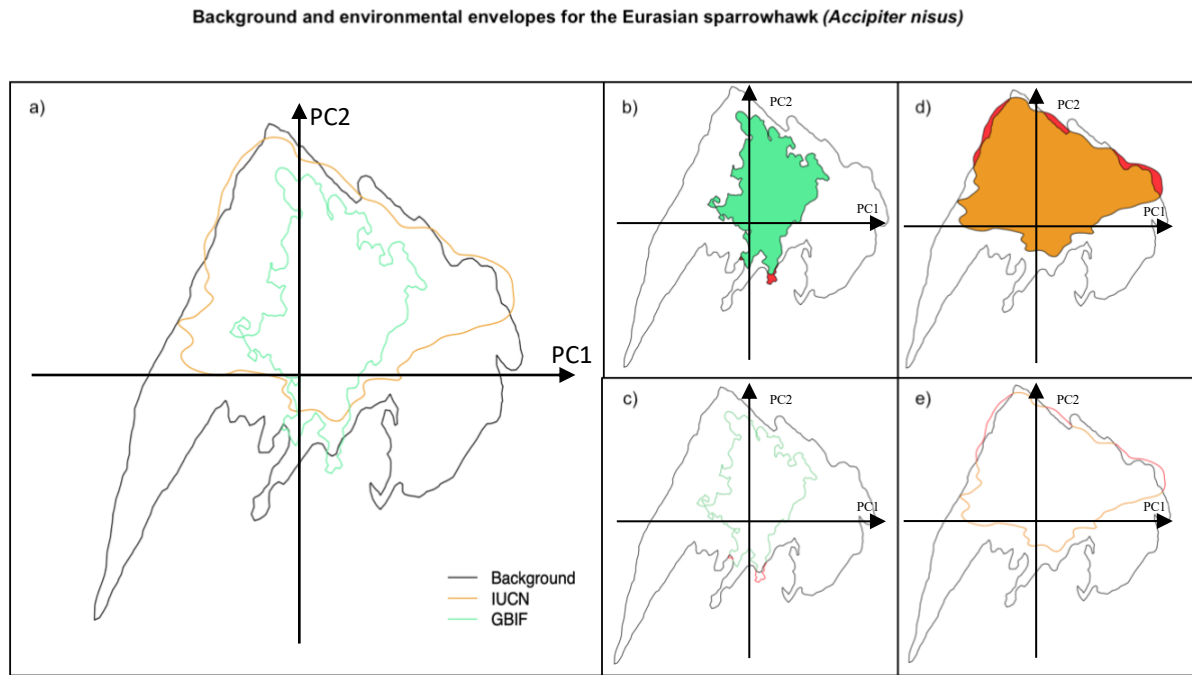


Figure 3: a) Environmental background envelope (black line) and species niche envelope based on IUCN (orange) and GBIF (green) datasets of the Eurasian sparrowhawk (*Accipiter nisus*) projected on the 2 dimensional scale provided by the Principal component axis 1 and 2 with b) truncated area (red area) with GBIF dataset, c) truncated perimeter (red line) with GBIF dataset, d) truncated area (red area) with IUCN dataset and e) truncated perimeter (red line) with IUCN dataset.

The truncation, considered as the percentage of the realized niche that goes beyond the environmental background envelope varied between 0 and 10.50% for the area and between 0 and 32.66% for the perimeter of niches obtained with GBIF data. The range of truncation value obtained with IUCN data was from 0 to 9.15% for the area and from 0 to 54.93% for the perimeter (Tab1, Fig4). Out of 49 species, 29 species showed niche truncation when analysed with GBIF data whereas all except four showed niche truncation based on the IUCN dataset (Tab1).

The mean area truncation for the GBIF dataset was 1.05% whereas this values was 2.45% for the IUCN dataset. The mean perimeter truncation with GBIF data was estimated at 5.62% and at 17.92% for the IUCN dataset.

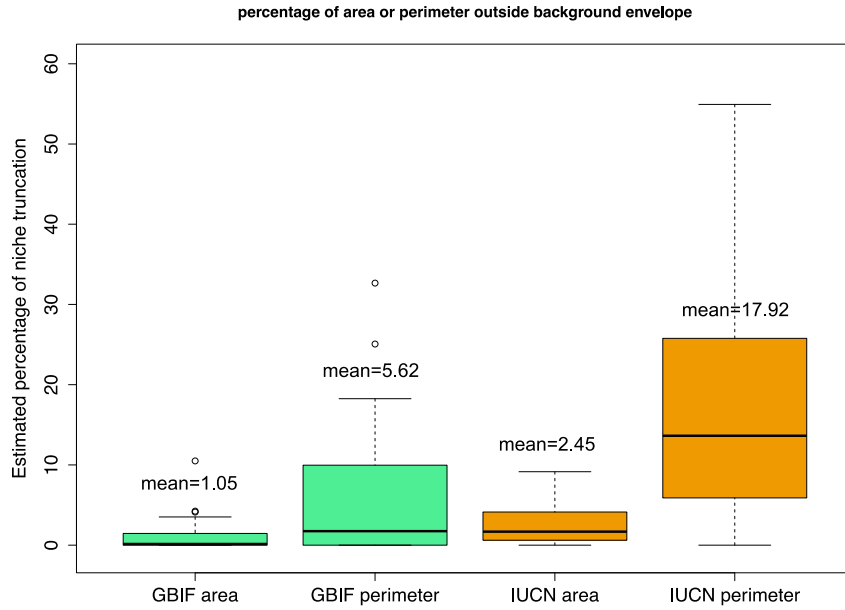


Figure 4: Estimated percentage of niche truncation in function of the database (GBIF in green, IUCN in orange) and measure method (area, perimeter).

The Wilcoxon-Mann-Whitney test indicates significant differences between GBIF and IUCN data regarding the area-based estimate of truncation level ($p\text{-value} = 3.741e-05$). The same conclusion was reached regarding the truncated perimeter of realized niches ($p\text{-value} = 3.057e-06$).

For both datasets, the perimeter outside the environmental background is higher than the area.

The figure 5 shows the percentage of truncation in function of the niche size in order to analyse whether there is a relation between the size and the degree of truncation of the niche. A strong relation is revealed indicating that we expect more truncation for species with a larger niche (Fig. 5). It implies that generalist species tend to have more niche truncation since their niche are bigger than specialist species.

The last analysis allowed to reveal differences between groups of species based on the taxonomic order of the 49 species. The mean estimated truncation is higher for some of the orders. For instance, Coraciiformes and Falconiformes order show a high estimated truncation based on GBIF and IUCN data, whereas Galliiformes shows a high estimated truncation only with IUCN data.

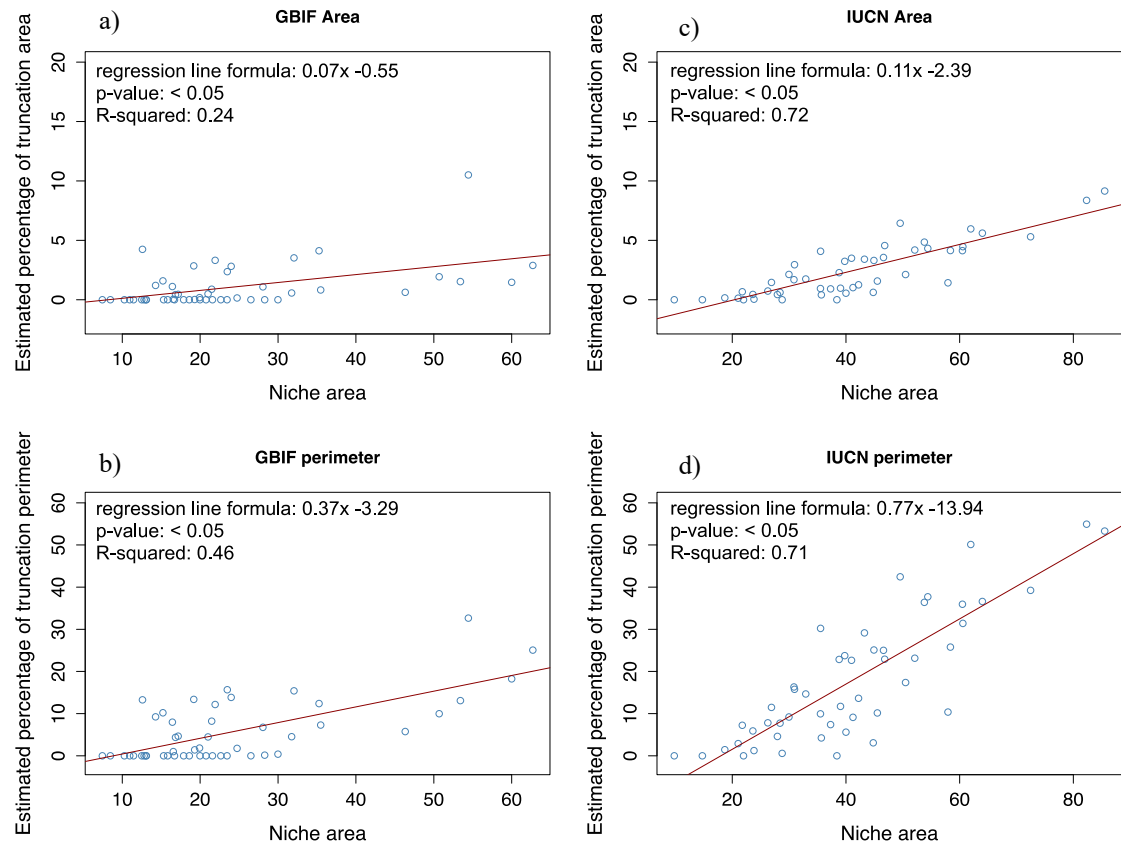


Figure 5: Estimated percentage of truncation in function of the niche size. Niche truncation percentage was estimated with a) truncated area with GBIF dataset, b) truncated perimeter with GBIF dataset, c) truncated area with IUCN dataset, d) truncated perimeter with IUCN dataset.

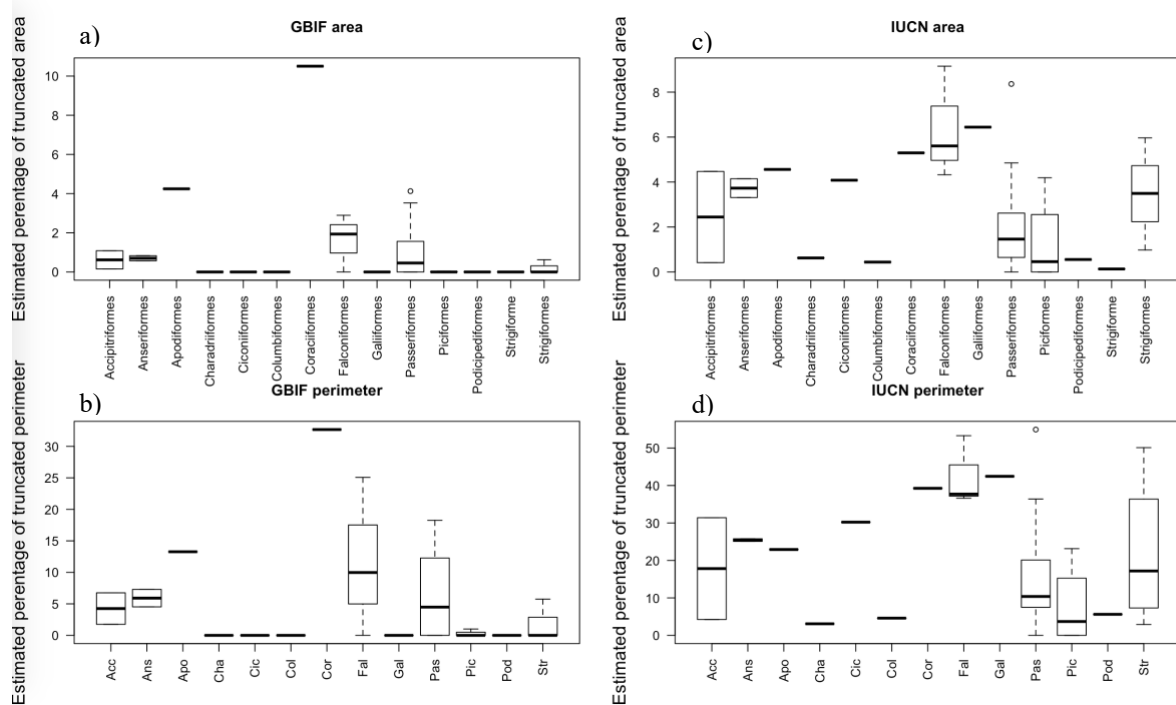


Figure 6: Estimated percentage of truncation in function of groups of species based on taxonomic order. Truncation estimated with a) truncated area with GBIF dataset, b) truncated perimeter with GBIF dataset, c) truncated area with IUCN dataset, d) truncated perimeter with IUCN dataset.

SP.NAMES	SP.NAMES	ORDRE	GBIF TRUNCATION AREA	GBIF TRUNCATION PERIMETER	GBIF TOTAL AREA	IUCN TRUNCATION AREA	IUCN TRUNCATION PERIMETER	IUCN TOTAL AREA
ACCIPITER NISUS	Epervier d'Europe	Accipitriformes	1.08470795	6.75158087	28.0635434	4.4723486	31.4035631	60.5620091
ALCEDO ATTHIS	Martin pêcheur	Coraciiformes	10.5043716	32.6620949	54.4336769	5.29769588	39.248252	72.4730176
ANAS PLATYRHYNCHOS	Canard colvert	Anseriformes	0.82973846	7.29514677	35.4793091	4.14698322	25.775725	58.3645711
ASIO OTUS	Hibou moyen-duc	Strigiformes	0	0	17.8873705	0.97745081	11.717624	39.0402346
BOMBYCILLA GARRULUS	Jaseur boréal	Passeriformes	0	0	13.0499068	1.69031399	16.3195143	30.8474918
BUTEO BUTEO	Buse variable	Accipitriformes	0.15422489	1.75324774	24.7392955	0.4165062	4.23413199	35.6733773
CARDUELIS CARDUELIS	Chardonneret élégant	Passeriformes	0.49331937	4.48263804	21.0193218	0.45193431	5.8984872	23.6282288
CHLORIS CHLORIS	Verdier d'Europe	Passeriformes	0.02323289	1.4073951	19.2926148	0	0	21.9867647
CICONIA CICONIA	Cigogne blanche	Ciconiiformes	0	0	21.5906424	4.08176604	30.2180575	35.5285499
COLUMBA PALUMBUS	Pigeon ramier	Columbiformes	0	0	15.3182567	0.43995904	4.5944933	27.9425845
CORVUS CORONE	Corneille noire	Passeriformes	0.88737797	8.21713767	21.4913802	2.28388133	22.8456101	38.8044183
COTURNIX COTURNIX	Caille des blés	Galliiformes	0	0	22.6590278	6.44222245	42.4470721	49.5431368
CYANISTES CAERULEUS	Mésange bleue	Passeriformes	0	0	13.0949811	0.16197606	1.45131567	18.6857512
DELICHON URBICUM	Hirondelle de fenêtre	Passeriformes	0.0015611	0.1569264	28.2905178	2.12311491	17.3793149	50.484781
DENDROCOPOS MAJOR	Pic épeiche	Piciformes	0.0390543	1.00628037	16.5601075	0.91766497	7.38697182	37.2961034
DENDROCOPOS MEDIUS	Pic mar	Piciformes	0	0	7.45168448	0	0	9.79767889
EMBERIZA CITRINELLA	Bruant jaune	Passeriformes	1.59381487	10.2013918	15.2207965	0.67479778	7.24353838	21.7695919
FALCO PEREGRINUS	Faucon pèlerin	Falconiformes	2.89301164	25.0726785	62.7111125	9.15376523	53.2988742	85.5057288
FALCO SUBBUTEO	Faucon hobereau	Falconiformes	0	0	26.5167946	4.32392688	37.704494	54.3897284
FALCO TINNUNCULUS	Faucon crécerelle	Falconiformes	1.93478488	9.97330983	50.6868402	5.60606764	36.6105511	63.9838262
FICEDULA HYPOLEUCA	Gobemouche noir	Passeriformes	0	0	10.2834732	2.9518979	15.7738531	30.9534199
FRINGILLA COELEBS	Pinson des arbres	Passeriformes	2.37071472	15.6780354	23.4820368	0.04608601	1.23816936	23.8218871
GARRULUS GLANDARIUS	Geai des chênes	Passeriformes	0.45781248	4.65654618	17.18323	0.7357089	7.82773304	26.2586297
HIRUNDO RUSTICA	Hirondelle rustique	Passeriformes	1.46646513	18.2569583	59.998295	8.36272301	54.929244	82.319702
JYNX TORQUILLA	Torcol fourmilier	Piciformes	0	0	16.7152818	4.19090965	23.1450281	52.1059626
LOXIA CURVIROSTRA	Bec-croisé des sapins	Passeriformes	3.52370405	15.4070452	32.04886	1.02578379	9.12358903	41.2382897
MERGUS MERGANSER	Harle bièvre	Anseriformes	0.57138332	4.52211202	31.7456774	3.31129368	25.0841548	44.9107201
ORIOLOS ORIOLOS	Loriot d'Europe	Passeriformes	0	0	18.6095008	3.41474523	29.1479196	43.2411191
OTUS SCOPS	Petit-duc scops	Strigiformes	0	0	11.4441606	3.49425574	22.6458929	40.9768354
PARUS MAJOR	Mésange charbonnière	Passeriformes	0	0	19.9840297	1.42853395	10.3788001	57.9418705
PASSER DOMESTICUS	Moineau domestique	Passeriformes	1.53032366	13.0875233	53.4143896	4.13690687	35.9488867	60.4792858
PERIPARUS ATER	Mésange noire	Passeriformes	3.31751444	12.1644503	21.9287296	0.94293352	9.94444183	35.5105381
PHOENICURUS OCHRURUS	Rougequeue noir	Passeriformes	4.12580852	12.3895075	35.2657784	4.85058324	36.4003247	53.7826278
PICA PICA	Pie bavarde	Passeriformes	0	0	20.7537816	1.25896631	13.6338679	42.2160399
PICUS VIRIDIS	Pic vert	Piciformes	0	0	8.47684087	0	0	14.7407036
PODICEPS NIGRICOLLIS	Grèbe à cou noir	Podicipediformes	0	0	23.459069	0.55514801	5.61410281	39.9906987
REGULUS REGULUS	Roitelet huppé	Passeriformes	1.20693194	9.24150156	14.2856858	0.00551509	0.57276845	28.7765805
SCOLOPAX RUSTICOLA	Bécasse des bois	Charadriiformes	0	0	12.8146544	0.62401031	3.08300766	44.7972615
SITTA EUROPAEA	Sitelle torchepot	Passeriformes	0.42962955	4.36122856	16.8437624	1.74666347	14.6712086	32.9295468
STRIX ALUCO	Chouette hulotte	Strigiformes	0	0	10.9669312	0.13366124	2.90285378	21.0377516
STURNUS_VULGARIS	Etourneau sansonnet	Passeriformes	0.00302687	0.4032178	29.9878891	1.576731	10.1560127	45.5191938
SYLVIA ATRICAPILLA	Fauvette à tête noire	Passeriformes	1.11554913	7.9750782	16.4342719	3.23921733	23.7421554	39.7811753
SYLVIA BORIN	Fauvette des jardins	Passeriformes	0.17957714	1.83961772	19.920339	3.55869096	25.0155254	46.5926271
TACHYMARPTIS MELBA	Martinet à ventre blanc	Apodiformes	4.2436525	13.2683013	12.6004372	4.56142889	22.912417	46.8165293
TROGLODYTES TROGLODYTES	Troglodyte mignon	Passeriformes	2.85396538	13.3754511	19.1845639	0	0	38.4033903
TURDUS MERULA	Merle noir	Passeriformes	2.81423595	13.8446541	23.9912039	2.1333194	9.2000377	29.9636915
TURDUS PHILOMELOS	Grive musicienne	Passeriformes	0	0	15.8304872	1.46157166	11.4872294	26.8957123
TURDUS PILARIS	Grive litorne	Passeriformes	0	0	12.5135581	0.61905747	7.72903535	28.4088905
TYTO ALBA	Effraie des clochers	Strigiformes	0.62530905	5.75132544	46.3289109	5.9644834	50.1070941	61.9474136

Table 1: Estimated niche truncation for the two databases (IUCN and GBIF) and the two methods (Area and Perimeter). Total area represents the total niche size.

Discussion

From the two cases where the species model distribution assumption is violated, the first one, i.e. the truncation induced by geographically restricted datasets was already assessed and a solution to avoid this problem proposed (Titeux et al., 2017). The second case consisting of a truncation induced by a limit of the environmental background in which we suppose not to observe all the conditions that a species can tolerate had never been assessed. Based on bird species niche, this study assess the extent to which the realized niche can be truncated.

The main result is that niche truncation could be observed in the majority of species with both datasets. With realized niche truncation ranging from 0 to more than 50%, this empirical result give rise to an important questioning about the accuracy of predictions for future species distribution models (Peterson et al., 2018) but also for spatial prediction as it is done with SDMs for invasive species (Jiufeng, Zhao, Zhao, & Zhang, 2018). As it was hypothesized, a niche truncation exists in the margins of the environmental conditions available on Earth. The predictions based on actual realized niche thus does not take into account the environmental conditions that are not available nowadays but that may become available in the future. For species with a part of the margins of the realized niche corresponding to the margins of the environmental conditions, we expect the realized niche to be larger in the future due to the change in the environmental conditions. If new environmental conditions that the species can tolerate become available in the future, the realized niche will increase. And this opportunity is for now not taken into account in predictions using climate change scenarios.

The Wilcoxon Mann-Whitney test, performed between IUCN and GBIF perimeter or area truncation also confirm that the two datasets have significant difference in their perimeter and area truncation distribution ($p\text{-values} = 3.057e-06$ and $3.741e-05$ respectively). It would be interesting to analyse which of the two datasets represents the best realized niche models, but nowadays, they remain the most used in the Niche modelling and Species Distribution Modelling (Alhajeri & Fourcade, 2019; Hawkins et al., 2008).

The analyse performed with the two datasets obtained from GBIF and IUCN values allowed to notice that the two realized niche computed with these values differed considerably. However, this is a different result than what showed Alhajeri & Fourcade who found no considerable difference in environmental data estimates between IUCN and GBIF data (Alhajeri & Fourcade, 2019). Both study yet used large spatial and taxonomic scale. Since this study showed no better database than the other, we thus need to consider both results and take the strength and weakness of the two datasets into account to explain the truncation that is expected. The maps displaying the IUCN polygons and GBIF dots representing the occurrences of the species shows that GBIF points are relatively close to the IUCN data in area such as the western part of Europe, South-Africa and Northern America but displays a noteworthy

difference for other regions. It can be explained by the fact that GBIF data comes from species presence observations by scientists and average citizen whereas the IUCN data are not only evaluated by presence-observations but also by experts predictions. We expect a sampling bias for the GBIF data since the species observation occurrences are less likely to be realized in regions with adverse conditions, far away from populated area and out of the tracks. This results in an underestimation of the species occurrences and thus of the realized niche. A shift bias can also result from these data if the extreme values are not taken into account due to a lack of occurrence observations in adverse environmental conditions. On the other hand, the expert prediction extrapolations in the IUCN data that also have a more coarse resolution are expected to overestimate presence occurrence of species in area with sudden spatial environmental conditions changes such as mountain, different land use, etc. For instance, with this kind of coarse resolution, a sudden altitude elevation on a mountain can not to be taken into account and a species living in low elevation habitat would be considered as present in our occurrence sampling since with the extrapolation this area is part of the IUCN polygons although the species is not found in reality. We thus expect the percentage of truncation to be in between the values obtained with these two datasets.

Although truncation was noticed with both methods, i.e. when analysing the area or the perimeter of the realized niche that was beyond the environmental background envelope, the perimeter truncation proved to be higher. We can then question which is the best mean to analyse truncation and define more precisely the realized niche truncation, now that these results revealed it, with these two experimental methods. The perimeter is thought to overestimate the truncation since the n-environmental space is probably rugged (Blonder et al., 2014) which will largely increase the perimeter whereas the area will remain approximately the same. Since holes can also be found in the niche (Blonder et al., 2014), this will also increase the perimeter while reducing the area. For these reasons we expect the perimeter to overestimate the percentage of truncation but further analysis are required. Other methods are also conceivable in order to estimate truncation with the best accuracy.

The analyse of the truncation extent in function of the realized niche size returned a strong relation for all data test and, all the more, a positive relation. It thus confirm the hypothesis that a higher truncation is linked with a larger realized niche. And since generalist species tend to have larger niche than specialist species (Pagani-Núñez et al., 2019), the niche truncation issue is then more acute for generalist species. Other hypothesis that should be tested are if specialist species can also present truncation if their environmental suitable conditions are near the environmental conditions margins and if species with extreme conditions tolerance would be more exposed to niche truncation, since they are more susceptible to have realized niche margins that coincide with environmental conditions margins.

The last hypothesis expecting differences in the niche truncation depending on species groups raise questions about other species characteristic traits that could influence the extent to which the realized niche is truncated. Since the results revealed different niche truncation in function of the taxonomic order of the species, we expect the niche truncation to be related to other species traits than niche size and taxonomic order. Knowing which species traits influence the extent of niche truncation is important since they species traits are a key to species vulnerability assessment (Aubin et al., 2018). Other life-history traits either affecting populations dynamics such as dispersal distance, generation time, affecting species distribution such as thermic tolerance or affecting biotic relations such as trophic level, home range size (Santini et al., 2016) would be interesting to analyse the effects of these traits on niche truncation.

This exploratory study revealed several limits. Besides the IUCN and GBIF databases used to estimate species, other limits were encountered. The dataset that was used to compute the environmental background space consisted of temperature and precipitation variables. Other environmental variables should also be taken into account in order to estimate the environmental conditions available with the best accuracy. For instance, soil occupation, tree cover, light exposure are variables that could be interesting to analyse as well. The kernel density estimate tool is also a limiting parameter of this study. Other estimates are available such as convexhull (Blonder et al., 2014) and alphahull (Arias-Castro & Casal, 2015) to define the envelopes. As Blonder et al. shown, the convexhull estimate is sensitive to outlier occurrences points and the kernel density estimate allows to model holes in the envelope (Blonder et al., 2014). The kernel density estimate was thus chosen for this first study, but other estimates should be tested in future studies.

In conclusion, this study revealed the extent to which the realized niche is truncated for bird species with IUCN and GBIF data. We expect this study to be the trigger of further analyses on other species, and with different methods in order to determine which is the best mean to estimate truncation. This exploratory study revealed niche truncation and pave the way to numerous other studies. The next step will be to test other methods to define the more acute way to estimate niche truncation and experiment tools in order to take this truncation into account when modelling niche and species distribution models for spatial predictions for instance for invasive species analysis or for predictions using climate change scenarios.

Acknowledgement

I want to thank my project supervisor Mathieu Chevalier for his precious help throughout the project and particularly for his help for the analysis with R and his comments on the first draft of this manuscript. I also want to thank my project director, Prof. Antoine Guisan for his help as well as all people from his group for their useful comments during the project presentation.

References

- Alhajeri, B. H., & Fourcade, Y. (2019). High correlation between species-level environmental data estimates extracted from IUCN expert range maps and from GBIF occurrence data. *Journal of Biogeography*, 46(7), 1329–1341. doi: 10.1111/jbi.13619
- Arias-Castro, E., & Casal, A. (2015). On Estimating the Perimeter Using the Alpha-Shape. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 53. doi: 10.1214/16-AIHP747
- Aubin, I., Boisvert-Marsh, L., Kebli, H., McKenney, D., Pedlar, J., Lawrence, K., ... Ste-Marie, C. (2018). Tree vulnerability to climate change: improving exposure-based assessments using traits as indicators of sensitivity. *Ecosphere*, 9(2), e02108. doi: 10.1002/ecs2.2108
- 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? *Ecography*, 33(5), 878–886. doi: 10.1111/j.1600-0587.2010.06181.x
- Bioclimatic variables | WorldClim - Global Climate Data. (n.d.). Retrieved December 5, 2019, from <http://worldclim.org/bioclim>
- Blonder, B., Lamanna, C., Violle, C., & Enquist, B. J. (2014). The n-dimensional hypervolume. *Global Ecology and Biogeography*, 23(5), 595–609. doi: 10.1111/geb.12146
- Edwards, J. L. (2004). Research and Societal Benefits of the Global Biodiversity Information Facility. *BioScience*, 54(6), 485–486. doi: 10.1641/0006-3568(2004)054[0486:RASBOT]2.0.CO;2
- Hallstan, S. (2011). *Species Distribution Models - Ecological Applications for Management of Biodiversity*.
- Hawkins, B. A., Rueda, M., & Rodríguez, M. Á. (2008). What Do Range Maps and Surveys Tell Us About Diversity Patterns? *Folia Geobotanica*, 43(3), 345. doi: 10.1007/s12224-008-9007-8
- Herkt, K. M. B., Skidmore, A. K., & Fahr, J. (2017). Macroecological conclusions based on IUCN expert maps: A call for caution. *Global Ecology and Biogeography*, 26(8), 930–941. doi: 10.1111/geb.12601
- Holt, R. D. (1990). The microevolutionary consequences of climate change. *Trends in Ecology*

- & *Evolution*, 5(9), 311–315. doi: 10.1016/0169-5347(90)90088-U
- Hurlbert, A. H., & Jetz, W. (2007). Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proceedings of the National Academy of Sciences*, 104(33), 13384–13389. doi: 10.1073/pnas.0704469104
- Hutchinson, G. E. (1953). The Concept of Pattern in Ecology. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 105, 1–12. Retrieved from JSTOR.
- Jiufeng, W., Zhao, Q., Zhao, W., & Zhang, H. (2018). Predicting the potential distributions of the invasive cycad scale *Aulacaspis yasumatsui* (Hemiptera: Diaspididae) under different climate change scenarios and the implications for management. *PeerJ*, 6, e4832. doi: 10.7717/peerj.4832
- Lam, N. S.-N., & Quattrochi, D. A. (1992). On the Issues of Scale, Resolution, and Fractal Analysis in the Mapping Sciences. *The Professional Geographer*, 44(1), 88–98. doi: 10.1111/j.0033-0124.1992.00088.x
- Loiselle, B. A., Howell, C. A., Graham, C. H., Goerck, J. M., Brooks, T., Smith, K. G., & Williams, P. H. (2003). Avoiding Pitfalls of Using Species Distribution Models in Conservation Planning. *Conservation Biology*, 17(6), 1591–1600. doi: 10.1111/j.1523-1739.2003.00233.x
- Nogués-Bravo, D., Rodríguez-Sánchez, F., Orsini, L., de Boer, E., Jansson, R., Morlon, H., ... Jackson, S. T. (2018). Cracking the Code of Biodiversity Responses to Past Climate Change. *Trends in Ecology & Evolution*, 33(10), 765–776. doi: 10.1016/j.tree.2018.07.005
- Omann, I., Stocker, A., & Jäger, J. (2009). Climate change as a threat to biodiversity: An application of the DPSIR approach. *Ecological Economics*, 69(1), 24–31. doi: 10.1016/j.ecolecon.2009.01.003
- Pagani-Núñez, E., Liang, D., He, C., Zhou, X., Luo, X., & Goodale, E. (2019). Niches in the Anthropocene: Passerine assemblages show niche expansion from natural to urban habitats. *Ecography*. doi: 10.1111/ecog.04203
- Pecchi, M., Marchi, M., Burton, V., Giannetti, F., Moriondo, M., Bernetti, I., ... Chirici, G. (2019). Species distribution modelling to support forest management. A literature review. *Ecological Modelling*, 411, 108817. doi: 10.1016/j.ecolmodel.2019.108817
- Peterson, A. T., Cobos, M. E., & Jiménez-García, D. (2018). Major challenges for correlational ecological niche model projections to future climate conditions. *Annals of the New York Academy of Sciences*, 1429(1), 66–77. doi: 10.1111/nyas.13873
- R Core Team (2018). (Version R version 3.5.1). (2018). Retrieved from <https://www.r-project.org/index.html>
- Santini, L., Cornulier, T., Bullock, J. M., Palmer, S. C. F., White, S. M., Hodgson, J. A., ...

- Travis, J. M. J. (2016). A trait-based approach for predicting species responses to environmental change from sparse data: how well might terrestrial mammals track climate change? *Global Change Biology*, 22(7), 2415–2424. doi: 10.1111/gcb.13271
- Smolik, M. G., Dullinger, S., Essl, F., Kleinbauer, I., Leitner, M., Peterseil, J., ... Vogl, G. (2010). Integrating species distribution models and interacting particle systems to predict the spread of an invasive alien plant. *Journal of Biogeography*, 37(3), 411–422. doi: 10.1111/j.1365-2699.2009.02227.x
- Soberón, J., & Nakamura, M. (2009). Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences*, 106(Supplement 2), 19644–19650. doi: 10.1073/pnas.0901637106
- Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M. M. B., Allen, S. K., Boschung, J., ... Midgley, P. M. (2013). *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. 14.
- Thuiller, W., Brotons, L., Araújo, M. B., & Lavorel, S. (2004). Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, 27(2), 165–172. doi: 10.1111/j.0906-7590.2004.03673.x
- Titeux, N., Maes, D., Daele, T. V., Onkelinx, T., Heikkinen, R. K., Romo, H., ... Luoto, M. (2017). The need for large-scale distribution data to estimate regional changes in species richness under future climate change. *Diversity and Distributions*, 23(12), 1393–1407. doi: 10.1111/ddi.12634
- Wilson, E. O., & Peter, F. M. (1988). *The Effect of Global Climatic Change on Natural Communities*. Retrieved from <https://www.ncbi.nlm.nih.gov/books/NBK219279/>
- Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Ritter, C. D., Edler, D., ... see <<https://github.com/ropensci/onboarding/issues/210>>). (2019). CoordinateCleaner: Automated Cleaning of Occurrence Records from Biological Collections (Version 2.0-11). Retrieved from <https://CRAN.R-project.org/package=CoordinateCleaner>