



University College London  
Department of Genetics, Evolution and Environment

# **Traits influence responses to land-use and climate change in terrestrial vertebrates**

**Adrienne Etard**

Supervision: Dr. Tim Newbold, Dr. Alex Pigot

April 25, 2022

Submitted in part fulfilment of the requirements  
for the degree of Doctor of Philosophy in Ecology at University College London

Word count: ~XXXXXX



# Declaration

I, Adrienne Etard, confirm that the work presented in this thesis is my own. Some parts have been conducted in collaboration with other researchers, and the contributions of co-authors to specific Chapters are described in the *Thesis outline of contents* on page 6. All else is appropriately referenced, and where information has been derived from other sources, I confirm that this has been indicated in the Thesis.

A handwritten signature in black ink, appearing to read "Adrienne Etard".

# Acknowledgments

First and foremost, huge thanks to my primary supervisor, Tim Newbold, whose guidance has been pivotal to the completion of this journey. Tim, thank you for trusting and encouraging me throughout this PhD, and for the countless opportunities you have offered me.

I am equally grateful to Alex Pigot for hosting me in his lab, for the cutting-edge insights on my work, and for the weekly Zoom meetings that made remote-working more bearable throughout various phases of the Covid-19 pandemic.

I would like to extend my thanks to the wonderful people of CBER, both to those with whom I have directly collaborated (Jess Williams, Gonzalo Albaladejo-Robles, ), and to those who have facilitated my work in this journey with insightful conversations. Thank you for making the workplace a joy.

To my friends in London and elsewhere. In particular: Tom, Yumi. To my family for their spirits and their support, Anne, Jean-François, Vivien, Octave & Caroline.

# **Abstract**

# **Impact statement**

# Data and code access

## Code access

The main pieces of R code for Chapters 2, 3, 4 and 5 are available at: [https://github.com/AdrienneEtard/PhD\\_thesis\\_code](https://github.com/AdrienneEtard/PhD_thesis_code).

This document was compiled with L<sup>A</sup>T<sub>E</sub>X; the source code and files are available at: <https://github.com/AdrienneEtard/Thesis>

## Data

The data used in this Thesis are freely accessible and sources are referenced throughout. In particular:

- the PREDICTS database (Hudson et al., 2014, 2017) can be downloaded from: <https://data.nhm.ac.uk/dataset/the-2016-release-of-the-predicts-database>;
- the trait datasets I compiled in Chapter 2 are available at: [https://figshare.com/articles/dataset/Global\\_gaps\\_in\\_terrestrial\\_vertebrate\\_trait\\_data/10075421](https://figshare.com/articles/dataset/Global_gaps_in_terrestrial_vertebrate_trait_data/10075421).

# **Thesis outline of contents, authorship and collaborations**

# Contents

<b>List of Tables</b>	<b>10</b>
<b>List of Figures</b>	<b>12</b>
<b>1 General introduction</b>	<b>16</b>
<b>2 Global gaps and biases in trait data for terrestrial vertebrates</b>	<b>17</b>
2.1 Introduction . . . . .	18
2.2 Methods . . . . .	21
2.2.1 Trait data collection . . . . .	21
2.2.2 Investigating gaps and biases in trait data . . . . .	27
2.3 Results . . . . .	29
2.3.1 Taxonomic biases in trait information . . . . .	29
2.3.2 Phylogenetic biases in trait completeness . . . . .	30
2.3.3 Spatial biases in trait completeness . . . . .	33
2.4 Discussion . . . . .	35
<b>3 Intensive human land uses negatively affect vertebrate functional diversity</b>	<b>40</b>
3.1 Introduction . . . . .	40
3.2 Methods . . . . .	42
3.2.1 Vertebrate assemblages . . . . .	42
3.2.2 Functional traits and diversity indices . . . . .	44
3.2.3 Effects of land use and land-use intensity on FRic and FDis (Hypothesis 1) . . . . .	45
3.2.4 Investigating functional under-dispersion (Hypothesis 2) . . . . .	47
3.2.5 Functional loss and functional gain (Hypothesis 3) . . . . .	47
3.3 Results . . . . .	48
3.3.1 Effects of land use on FRic and FDis . . . . .	48

3.3.2	Changes in the probability of occurrence of functional under-dispersion . . . . .	51
3.3.3	Functional loss and gain . . . . .	51
3.4	Discussion . . . . .	55
<b>4</b>	<b>Geographical range area, habitat breadth and specialisation on natural habitats predict land-use responses and climate-change sensitivity more consistently than life-history and dietary traits in terrestrial vertebrates</b>	<b>59</b>
4.1	Introduction . . . . .	59
4.2	Methods . . . . .	59
4.3	Results . . . . .	59
4.4	Discussion . . . . .	59
<b>5</b>	<b>Energetic constraints and trophic group explain species persistence in disturbed land uses</b>	<b>60</b>
5.1	Introduction . . . . .	61
5.2	Methods . . . . .	66
5.2.1	Vertebrate assemblage composition . . . . .	66
5.2.2	Energy availability by land-use type and land-use intensity . . . . .	66
5.2.3	Resting Metabolic Rates (RMR) & imputations of missing RMR values . . . . .	67
5.2.4	Trophic group and body mass information . . . . .	68
5.2.5	Effects of land use, land-use intensity and trophic group on assemblage-level total RMR (prediction 1; Figure 5.2a) . . . . .	68
5.2.6	Effects of land use, land-use intensity, trophic group and residual RMR on species occurrence probability (prediction 2; Figure 5.2b) . . . . .	69
5.3	Results . . . . .	70
5.3.1	Effects of land use, land-use intensity and trophic group on assemblage-level total RMR . . . . .	70
5.3.2	Effects of land use, land-use intensity, trophic group and residual RMR on species' occurrence probability . . . . .	71
5.4	Discussion . . . . .	74
<b>6</b>	<b>General discussion</b>	<b>77</b>
<b>7</b>	<b>Conclusion</b>	<b>78</b>

<b>Bibliography</b>	<b>79</b>
<b>Appendices</b>	<b>94</b>
<b>Appendix 0: Supporting information for Chapter 1</b>	<b>95</b>
<b>Appendix 1: Supporting information for Chapter 2</b>	<b>95</b>
S2.1 Taxonomic corrections . . . . .	96
S2.2 Additional information for trait compilation . . . . .	99
S2.3 Cutting distribution maps by altitudinal limits . . . . .	100
S2.4 Impact of taxonomic corrections on trait coverage . . . . .	101
S2.5 Assemblage-level median, mean and standard deviation of trait completeness (maps) . . . . .	102
S2.6 Phylogenetic patterns in trait completeness . . . . .	107
S2.7 Model coefficients for Range size against Number of sampled traits (Poisson model). . . . .	112
S2.8 Spatial models summaries . . . . .	114
S2.9 Trait coverage and taxonomic matching . . . . .	115
<b>Appendix 2: Supporting information for Chapter 3</b>	<b>115</b>
S3.1 Land-use categories in PREDICTS and sample sizes . . . . .	116
S3.2 Trait data & imputation of missing trait values . . . . .	119
S3.2.1 Choice of imputation technique . . . . .	121
S3.2.2 Phylogenetic signal in traits . . . . .	121
S3.2.3 Implementation of missForest imputations . . . . .	122
S3.3 Degree of multicollinearity among traits . . . . .	123
S3.4 Imputation performance . . . . .	124
S3.5 Functional loss and functional gain . . . . .	127
S3.6 Diagnostic plots . . . . .	130
S3.7 Model robustness . . . . .	139
S3.8 Model robustness – time since land-use conversion . . . . .	146
<b>Appendix 3: Supporting information for Chapter 4</b>	<b>147</b>
<b>Appendix 4: Supporting information for Chapter 5</b>	<b>149</b>
<b>Bibliography for the Appendices</b>	<b>158</b>

# List of Tables

<b>Chapter 2</b>	<b>17</b>
2.1    Data sources for each trait . . . . .	23
2.2    Number of species for each analysis . . . . .	27
<b>Appendix 1</b>	<b>95</b>
S2.1    Coefficients of the model investigating whether species range size explained the number of sampled traits . . . . .	112
S2.2    Coefficients of the model investigating whether species range size explained the number of sampled traits, using range maps not cut by altitudinal limits . . . . .	112
S2.3    Spatial model summary for amphibians . . . . .	114
S2.4    Spatial model summary for reptiles . . . . .	114
<b>Appendix 2</b>	<b>115</b>
S3.1    Land-use categories in the PREDICTS database . . . . .	116
S3.2    Phylogenetic signal in continuous and categorical traits . . . . .	122
S3.3    Variance inflation factors across considered (imputed) traits . . . . .	123
S3.4    Sample sizes (number of pairs of sites) for the calculation of functional loss and functional gain across all species . . . . .	128
S3.5    Sample sizes (number of pairs of sites) for the calculation of functional loss and functional gain within each class . . . . .	129
S3.6    Summary of the model explaining FRic by land use and time since land-use conversion, fitted on the subset of data for which there are information on time since land-use conversion	146
<b>Appendix 3</b>	<b>147</b>

S3.7	Summary of the model explaining FDis by land use and time since land-use conversion, fitted on the subset of data for which I have information on time since land-use conversion	148
------	---	-----

**Appendix 4** **149**

S5.1	Number of species for which RMR data was collected in each class & data sources, initial RMR coverage for PREDICTS species and phylogenetic signal in RMR . . . . .	150
------	--	-----

# List of Figures

<b>Chapter 2</b>	<b>17</b>
2.1    Procedure used to identify the accepted names of species . . . . .	22
2.2    Number of natural and artificial habitats used by a species against number of strictly natural habitats used by a species . . . . .	26
2.3    Trait coverage and completeness across species . . . . .	30
2.4    Within-family median trait completeness in herptiles . . . . .	32
2.5    Relationship between number of sampled traits and geographical range size . . . . .	33
2.6    Spatial distribution of assemblage-level median trait completeness in herptiles . . . . .	34
2.7    Spatial model trends for herptiles . . . . .	35
<b>Chapter 3</b>	<b>40</b>
3.1    Overview of the study design and functional metrics. . . . .	43
3.2    Effects of land use, land-use intensity and region on FRic (a) and on FDis (b) across all PREDICTS vertebrates. . . . .	49
3.3    Effects of land use, land-use intensity, region and taxonomic class on FRic (a) and FDis (b). .	50
3.4    Effects of land use, land-use intensity and region on the probability of occurrence of functional under-dispersion. . . . .	52
3.5    Effects of land use and land-use intensity on (a) functional loss and (b) functional gain across pairs of sites in temperate areas and tropical areas. . . . .	53
3.6    Effects of land use, land-use intensity, region and taxonomic class on functional loss and functional gain across pairs of sites. . . . .	54
<b>Chapter 4</b>	<b>59</b>

<b>Chapter 5</b>	<b>60</b>
5.1    Map of PREDICTS sites, sample sizes and NPP by land use and land-use intensity . . . . .	64
5.2    Framework for the predictions and models . . . . .	65
5.3    Effects of land use, land-use intensity and trophic group on assemblage-level total RMR. .	71
5.4    Slope estimates and predictions for the relationship between occurrence probability and residual RMR . . . . .	72
5.5    Predicted occurrence probabilities in primary vegetation and for the most disturbed land uses where residual RMR was found to have a significant effect on occurrence probability .	73
 <b>Appendix 1</b>	<b>95</b>
S2.1    Difference in species number before and after taxonomic correction (A) and distribution of number of synonyms across datasets (B) . . . . .	98
S2.2    (a) Body mass versus body length and (b) longevity versus age at sexual maturity in amphibians . . . . .	99
S2.3    Generation length versus longevity data in birds . . . . .	99
S2.4    Availability of altitudinal limits across species . . . . .	100
S2.5    Range sizes before versus after cutting by altitudinal limits . . . . .	100
S2.6    Trait coverage when no taxonomic correction is applied before matching sources by binomial names, versus when we applied the described procedure . . . . .	101
S2.7    Spatial distribution of assemblage-level median trait completeness in each class . . . . .	103
S2.8    Spatial distribution of assemblage-level mean trait completeness in each class . . . . .	104
S2.9    Spatial distribution of assemblage-level standard deviation of trait completeness in each class	105
S2.10    Assemblage-level species richness against standard deviation in completeness in each class	106
S2.11    Within-family median trait completeness in mammals . . . . .	108
S2.12    Within-family median trait completeness in birds . . . . .	109
S2.13    Within-family standard deviation in completeness . . . . .	110
S2.14    Within-family species richness against the within-family standard deviation of completeness	111
S2.15    Relationship between number of sampled traits and geographical range size using distribution maps not cut by altitudinal limits . . . . .	113
S2.16    Comparison of trait coverage among datasets corrected for taxonomy in different ways .	115
 <b>Appendix 2</b>	<b>115</b>

S3.1	Number of sites in each land use and land-use intensity for which FRic and FDis were calculated, across all species . . . . .	117
S3.2	Number of sites in each land use, land-use intensity and for which FRic and FDis were calculated, within each class . . . . .	118
S3.3	Relationship between habitat breadth and geographical range size across species in each class	119
S3.4	Trait coverage for the vertebrate species sampled in the PREDICTS database . . . . .	120
S3.5	Distribution of trait completeness across the vertebrate species sampled in the PREDICTS database . . . . .	120
S3.6	Distribution of continuous traits considered in the calculation of the functional diversity metrics . . . . .	124
S3.7	Frequency distribution for categorical traits considered in the calculation of the functional diversity metrics . . . . .	125
S3.8	Out-of-bag imputation errors for the continuous traits (a) and categorical traits (b) . . . . .	126
S3.9	Diagnostic plots for Model 1a . . . . .	130
S3.10	Diagnostic plots for Model 1b . . . . .	131
S3.11	Diagnostic plots for Model 2a . . . . .	132
S3.12	Diagnostic plots for Model 2b . . . . .	133
S3.13	Diagnostic plots for Model 3 . . . . .	134
S3.14	Diagnostic plots for Model 4a . . . . .	135
S3.15	Diagnostic plots for Model 4b . . . . .	136
S3.16	Diagnostic plots for Model 5a . . . . .	137
S3.17	Diagnostic plots for Model 5b . . . . .	138
S3.18	Effects of land use, land-use intensity and region on FRic (a) and FDis (b) across vertebrates, for the subset of species with complete trait data . . . . .	139
S3.19	Effects of land use, land-use intensity and region on FRic (a) and FDis (b), for the subset of species with complete trait data, with geographical range size as an additional trait . . . . .	140
S3.20	Effects of land use and land-use intensity on FRic (a) and FDis (b), obtained when calculating FRic and FDis with each set of imputed traits . . . . .	141
S3.21	Effects of land use and land-use intensity on FRic (a) and FDis (b), obtained when resampling primary vegetation sites twenty independent times . . . . .	142
S3.22	Effects of land use, land-use intensity and taxonomic class on FRic (a) and FDis (c), for the subset of species with complete trait data . . . . .	143

S3.23	Effects of land use, region, land-use intensity and taxonomic class on FRic and FDis obtained with the imputed trait data (black points) or with the complete data subsets (red points) . . . . .	144
S3.24	Effects of land use, region, land-use intensity and taxonomic class on FRic and FDis, obtained when calculating FRic and FDis with each set of imputed traits . . . . .	145
S3.25	Effects of land use on FRic for the model that includes time since land-use conversion (blue points) versus the model that does not take time since land-use conversion into account (red points) . . . . .	147
<b>Appendix 3</b>		<b>147</b>
<b>Appendix 4</b>		<b>149</b>
S5.1	Formula for the abundance model, used to understand the role of shifts in the body mass of species on observed changes in tRMR . . . . .	150
S5.2	Effects of land use, land-use intensity, trophic group, body mass and their interactions on assemblage-level total abundance . . . . .	151
S5.3	Diagnostic plots for the linear mixed-effects model looking at the effects of land use, land-use intensity, trophic group and their interactions on assemblage-level total RMR . . . . .	152
S5.4	Estimates for the models looking at the effects of land use, land-use intensity, trophic group and their interactions on assemblage-level total RMR . . . . .	153
S5.5	Effects of land use, land-use intensity, trophic level and their interactions on assemblage-level total RMR: empirical versus imputed . . . . .	154
S5.6	Diagnostic plots for the generalised mixed-effects model looking at the effects of land use, land-use intensity, trophic group, residual RMR and their interactions on species' probability of occurrence . . . . .	155
S5.7	Model's coefficients from the occurrence model fitted using the lme4 package against coefficients from the model fitted using the MCMCglmm package . . . . .	156
S5.8	Slope estimates and predictions for the relationship between occurrence probability and residual RMR for the model fitted on empirical RMR values only . . . . .	157

# **1 | General introduction**

## **2 | Global gaps and biases in trait data for terrestrial vertebrates**

### **Keywords**

Terrestrial vertebrates; traits; coverage; completeness; taxonomic biases; spatial biases; phylogenetic biases.

### **Abstract**

Trait data are increasingly used in studies investigating the impacts of global changes on the structure and functioning of ecological communities. Despite a growing number of trait data collations for terrestrial vertebrates, there is to date no global assessment of the gaps and biases the data present. Here, I assess whether terrestrial vertebrate trait data are taxonomically, spatially and phylogenetically biased. I compile seven ecological traits and quantify coverage as the proportion of species for which an estimate is available. For a species, I define completeness as the proportion of non-missing values across traits. I assess whether coverage and completeness differ across classes and examine phylogenetic biases in trait data. To investigate spatial biases, I test whether wider-ranging species have more complete trait data than narrow-ranging species. Additionally, I test whether species-rich regions, which are of most concern for conservation, are less well-sampled than species-poor regions. My results show that mammals and birds are well-sampled even in species-rich regions. For reptiles and amphibians (herptiles), only body size presents a high coverage (>80%), as well as habitat related variables (for amphibians). Herptiles are poorly sampled for other traits. The shortfalls are particularly acute in some species-rich regions and for certain clades. Across all classes, geographically rarer species have less complete trait information. Hence, trait information is less available on average in some of the most diverse areas and in geographically rarer species, both critical for biodiversity conservation. Gaps in trait data may impede our ability to conduct large scale analyses, while

23 biases can impact the validity of extrapolations. A short-term solution to the problem is to estimate missing  
24 trait data using imputation techniques, while a longer-term and more robust filling of existing gaps requires  
25 continued data collection efforts.

## 26 **2.1 Introduction**

27 Species traits are fundamental to ecological and evolutionary research. Comparative studies regularly use  
28 trait data across organisms to understand evolutionary processes and species coexistence (Escudero and  
29 Valladares, 2016; Zamudio et al., 2016), to investigate global patterns of life forms and functions (Díaz et  
30 al., 2016), or to assess species' vulnerability to environmental changes (Böhm et al., 2016; Pacifici et al.,  
31 2015; Pearson et al., 2014). Because traits influence species' ability to cope with environmental changes  
32 (Newbold et al., 2013) and underpin species' contributions to ecosystem processes (Lavorel and Garnier,  
33 2002; Violle et al., 2007; Wong et al., 2018), they play an increasingly important role in functional and  
34 conservation ecology.

35 Past and recent efforts to collate and release trait data in the public domain have facilitated the develop-  
36 ment of trait-based research. For instance, a global trait database has been published for plants (Kattge et al.,  
37 2011). As of May 2020, data from this database had been used in 297 publications since its release (Ac-  
38 tivity report, 18/06/2020, <https://www.try-db.org/TryWeb/Home.php>). Such databases hence  
39 constitute invaluable research tools and have the potential to greatly advance the field.

40 Vertebrates are one of the most studied taxa (Titley et al., 2017). There are now diverse sources of eco-  
41 logical traits for vertebrate groups (primates: Galán-Acedo et al., 2019; mammals: 'PanTHERIA', Jones et  
42 al., 2009; amniotes: Myhrvold et al., 2015; amphibians: 'AmphiBIO', Oliveira et al., 2017). These datasets  
43 stem from important efforts to collate published estimates of trait data and make them readily available. Trait  
44 data have also been made available on online platforms (for instance, the Global Assessment of Reptile Dis-  
45 tribution initiative: <http://www.gardinitiative.org/>; the IUCN Red List of Threatened Species:  
46 <https://www.iucnredlist.org/>; BirdLife data zone: <http://datazone.birdlife.org/>  
47 home).

48 Nevertheless, despite the importance of vertebrate species in global research outputs, there is no single  
49 source for vertebrate ecological traits. Consequently, researchers wishing to conduct comparative studies  
50 across vertebrate groups may have to collate trait data from a range of sources (such as in Cooke et al.  
51 (2019a,b) or in González-Suárez et al. (2018)), a time-consuming prerequisite which may be a limiting  
52 step of the research process. Indeed, collating data from heterogeneously-formatted sources presents many

53 challenges (Schneider et al., 2019), particularly when working across a large number of species. For instance,  
54 traits may be measured differently across datasets; units may be inconsistent; and taxonomic resolution and  
55 nomenclature may vary.

56 The lack of a curated, readily available global database for vertebrate ecological traits impedes our ability  
57 to conduct cross-taxon comparative studies at global scales. However, efforts to collate data into a single  
58 database are limited by the availability of underlying data. Because there exist important gaps in biodiversity  
59 knowledge (Hortal et al., 2015), trait datasets are often incomplete, with many species lacking estimates for  
60 many traits. The incompleteness of ecological trait data at the species level has been termed the ‘Raunkiærán  
61 shortfall’ by Hortal et al. (2015). Furthermore, incomplete trait data are likely to be biased. Biases in trait data  
62 may be the consequence of uneven taxonomic and spatial collection effort, with a set of charismatic or easily  
63 detectable species being more completely sampled. For instance, González-Suárez et al. (2012) investigated  
64 biases in global trait information in mammals. They notably found that the availability of mammalian trait  
65 data was geographically and phylogenetically biased, with larger and more widely distributed species being  
66 better sampled. In addition, data availability also differed across IUCN Red List extinction risk categories,  
67 with threatened species (Critically Endangered, Endangered or Vulnerable) being less well sampled for traits  
68 than non-threatened species (Least Concern or Near Threatened).

69 A major issue with incomplete, biased data is the introduction of bias in subsequent analyses. Assessing  
70 the amount of missing data as well as the so-called ‘missingness mechanism’ – whether missing data are  
71 missing at random, as opposed to there being systematic biases in the way missing values are distributed, see  
72 Baraldi and Enders (2010) – is an important prerequisite. Indeed, there exist diverse techniques to deal with  
73 data missingness. The simplest one consists of retaining complete cases only by filtering out missing values  
74 (case deletion, see Nakagawa and Freckleton (2008)). Nevertheless, case deletion may lead to biased param-  
75 eter estimates and erroneous conclusions when values are not missing at random (González-Suárez et al.,  
76 2012). Therefore, it is critical to determine the most appropriate way to deal with data incompleteness. For  
77 instance, previous studies using terrestrial vertebrate trait data have employed multiple imputation techniques  
78 to fill in the gaps (Cooke et al., 2019a; González-Suárez et al., 2012). Yet, imputation techniques could be  
79 sensitive to non-randomness in trait data. Phylogenetic biases (where some clades are under-sampled com-  
80 pared to other clades) could notably impact the performance of several imputation approaches. It is thus  
81 vital to characterise the gaps in trait data prior to any analysis. However, there has been no study to date  
82 investigating global patterns in the availability of trait data across terrestrial vertebrates.

83 Here, I aim to assess the global state of trait data in terrestrial vertebrates. I focus on a set of traits that are  
84 available across the four classes and that are commonly used by ecologists: body size; litter or clutch size;

85 longevity; trophic level; activity time; habitat breath; and a measure of habitat specialisation. I quantify and  
86 compare the gaps in trait data across classes by calculating the coverage of each trait across species, and the  
87 completeness of trait estimates for each species (Box 1). I investigate taxonomic, spatial and phylogenetic  
88 biases in trait coverage and completeness.

89 Given that biodiversity research is globally biased towards birds and mammals (Titley et al., 2017),  
90 I hypothesise that herptiles are less well sampled for traits than mammals and birds, having both lower  
91 coverage and completeness.

92 Furthermore, building upon previous studies conducted on mammals (González-Suárez et al., 2012), I  
93 hypothesize that species rarity influences completeness, focusing on species geographical range size as one  
94 aspect of rarity. Widely distributed species could be better sampled than narrowly distributed species because  
95 their ranges overlap with more study sites, regardless of their abundance. As such, I test whether species  
96 geographical range size explains trait completeness.

97 It is well established that global research effort is distributed unequally (United Nations Educational  
98 Scientific and Cultural Organization, 2015), with patterns underpinned by various geographical and so-  
99 cieconomic factors. For instance, countries with higher gross domestic product tend to host a larger number  
100 of research institutions (Martin et al., 2012). The proximity of research infrastructures and the accessibility  
101 of survey sites play an important part in explaining the global distribution of knowledge (Hortal et al., 2015).  
102 As a result of these factors, biodiversity data gaps tend to be greater in tropical areas (Collen et al., 2008).  
103 Tropical areas have the greatest species richness, and so these data biases are of great concern for biodiversity  
104 conservation. Whether species-rich regions are systematically under-sampled for traits compared to species-  
105 poor regions is thus important to assess, given the significance of species-rich areas for global conservation.  
106 Here, I investigate spatial biases in trait completeness, hypothesizing that species-rich areas are on average  
107 less well sampled than species-poor areas.

108 Finally, I investigate phylogenetic biases in the trait data. I assess whether particular clades have received  
109 more attention than others by looking for patterns in the distribution of trait completeness across the terminal  
110 branches of phylogenetic trees in each class.

### **Box 1. Definitions**

*Trait:* Sensu stricto, a characteristic measurable at the level of an individual and that influences organismal fitness or performance (Violle et al., 2007). In this thesis, I broaden this definition to include ‘ecological’ traits (e.g., the number of habitats used by a species), where the relationship of a species to the surrounding environment needs to be considered. Ecological traits may be estimated by aggregating data across multiple individuals.

*Trait completeness:* For a given species, the proportion of traits for which an estimate is available.

*Trait coverage:* For a given trait, the proportion of species for which an estimate is available.

## **111 2.2 Methods**

112 I produced class-specific trait datasets that were made available on figshare (DOI: 10.6084/m9.figshare.  
113 10075421). Data compilation and all analyses were conducted with R version 3.5.1 (R Core Team, 2018).  
114 Distribution maps were processed using both R and the ArcPy package available in ArcGIS v.10.6 (ESRI,  
115 2011) (implemented in Python 2.7; van Rossum, G. (1995)).

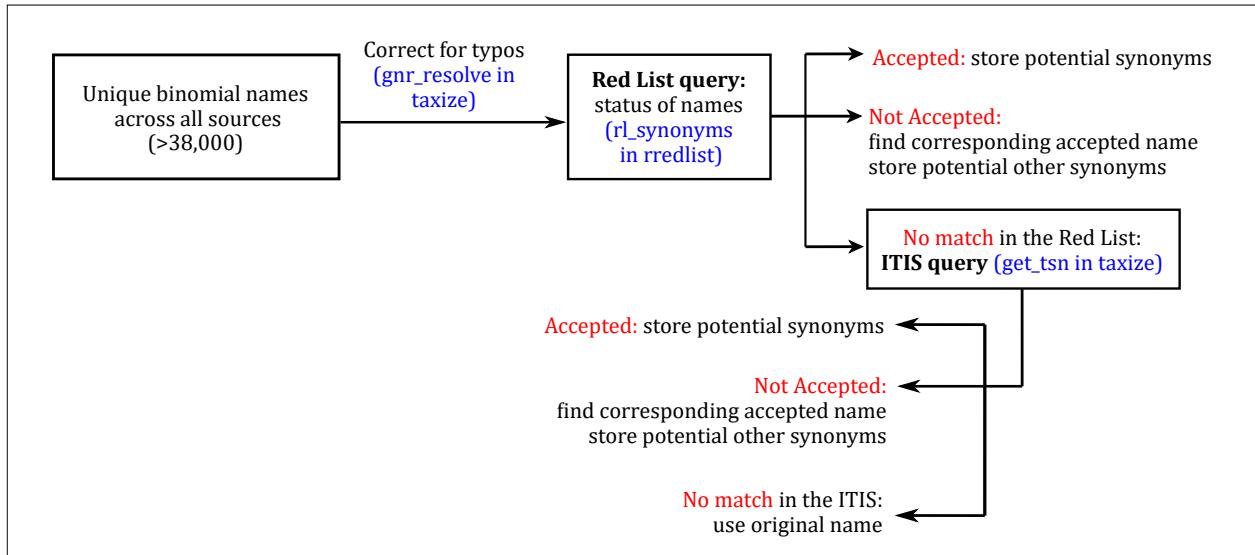
### **116 2.2.1 Trait data collection**

#### **117 Sources and taxonomic matching**

118 I used freely accessible secondary sources in my compilation (Table 2.1), selected for their broad taxonomic  
119 coverage and/or for their frequent use in macroecological studies. Across sources, similar species could  
120 appear under synonymous names. This was a potential problem for matching sources by binomial names.  
121 Indeed, synonymy can artefactually decrease trait coverage, when trait information is not available across  
122 all synonyms. Notably, difficulties arise when species have been divided into several subspecies or when  
123 different subspecies are clumped together. Systematic manual checks could not be applied considering the  
124 scale of the data collection (there were >39,000 unique binomial names across sources). I developed a  
125 procedure aiming at identifying one accepted name for each of the binomial names found across sources.  
126 When I could not find an accepted name, I used the original name. Figure 2.1 summarizes the main steps;  
127 similar solutions have been used in other large-scale studies (Cooke et al., 2019a).

128 Briefly, the procedure consisted of extracting synonyms from the IUCN (IUCN, 2020) or from the Inte-  
129 grated Taxonomic Information System (ITIS; <https://www.itis.gov/>), using the rredlist (Chamber-  
130 lain, 2018) and taxize (Chamberlain and Szöcs, 2013) R packages. One accepted name was assigned to each  
131 synonym. I produced a “Synonym” dataset that I have also made available. I then normalized taxonomy

132 across sources by replacing binomial names with their identified accepted name where applicable.  
 133 Given that different taxonomic backbones could be used to correct for taxonomy, I make two versions  
 134 of the trait compilations available (corrected and not corrected for taxonomy), meaning that users are free to  
 135 apply their own corrections; for example, taxonomy could be aligned to that of class-specific sources, such  
 136 as The Reptile Database, the American Museum of Natural History's Amphibian Species of the World, the  
 137 Mammal Diversity Database or the International Ornithological Congress World Bird List. Datasets cor-  
 138 rected for taxonomy contain 11,634 species of birds, 5,381 mammals, 10,612 reptiles and 6,990 amphibians.  
 139 Where no taxonomic correction was applied when matching sources, the compiled datasets contain 13,501  
 140 birds, 5,791 mammals, 11,012 reptiles and 8,583 amphibians. For more information, see Appendix 2 (S2.1;  
 141 Figure S2.1).



**Figure 2.1: Procedure used to identify the accepted names of species.** I extracted, where possible, the accepted names of species from either the IUCN Red List or the Integrated Taxonomic Information System (ITIS). Figure reproduced from Etard et al. (2020).

## 142 Compilation methods

143 For continuous traits, I took the median value within species when multiple estimates were available from  
 144 different sources, after removal of any repeated values, which were assumed to represent estimates duplicated  
 145 across secondary compilations and derived from the same underlying primary sources. Although intraspe-  
 146 cific variation is increasingly being recognized to have important effects on ecological systems (Bolnick  
 147 et al., 2011; Des Roches et al., 2018; González-Suárez and Revilla, 2012; Siefert et al., 2015), it was not  
 148 feasible to obtain measures of intraspecific variability from all sources; therefore, estimates were provided

**Table 2.1: Data sources for each trait.** Abbreviations: A = amphibians; B = birds; BL = body length; BM = body mass; DA = diel activity time; GL = generation length; H = habitat data; LCS = litter or clutch size; L/ML = longevity or maximum longevity; M = mammals; MA = age at sexual maturity; R = reptiles; RS = range size; TL = trophic level. Note. Data sources may contain more traits than shown here. Tick marks in parentheses indicate that the trait was present in the data source but that another closely related trait with a better coverage was used instead. The tilde character (~) before a tick mark indicates that I derived trophic levels from species diet. <sup>1</sup> <http://datazone.birdlife.org/home>; <sup>2</sup> <https://www.iucnredlist.org/resources/spatial-data-download>; <sup>3</sup> <http://apiv3.iucnredlist.org/api/v3/docs##general>. Table reproduced from Etard et al. (2020).

Sources	Taxa	Traits									
		BM	BL	L/ML	MA	GL	LCS	TL	DA	RS	H
Oliveira et al. (2017)	Amphibians	(✓)	✓	(✓)	✓		✓	~✓	✓		
Cooper et al. (2008)							✓				
Sodhi et al. (2008)			✓								
Wilman et al. (2014)	Birds	✓						~✓	✓		
BirdLife <sup>1</sup>		✓				✓					✓
Jones et al. (2009)	Mammals	✓	(✓)	(✓)	(✓)		✓		✓		
Kissling et al. (2014)								✓			
Gainsbury et al. (2018)								✓			
Wilman et al. (2014)		✓							✓		
Pacifici et al. (2015)		✓				✓					
Scharf et al. (2015)	Reptiles	✓		✓	(✓)		✓	✓	✓		
Vidan et al. (2017)										✓	
Stark et al. (2018)		✓		✓			✓		✓		
Schwarz and Meiri (2017)								✓			
Novosolov et al. (2017)		✓						✓			
Novosolov et al. (2013)							✓				
Slavenko et al. (2016)		✓									
Feldman et al. (2016)		✓									
Meiri (2018)					✓		✓	✓	✓		
Meiri et al. (2015)								✓	✓		
Roll et al. (2017)										✓	
Myhrvold et al. (2015)	B, M, R	✓	✓	✓	(✓)		✓				
IUCN (2020) <sup>2</sup>	A, B, M										✓
IUCN (2020) <sup>3</sup>	All										✓

<sup>149</sup> as a single measure for each species. For some species and some traits, measures were provided separately  
<sup>150</sup> for females and males. In such cases, I first obtained the mean of these two measures.

<sup>151</sup> Across sources, there were multiple traits related to each of body size and life span. For instance, body  
<sup>152</sup> mass and/or body length information could be provided. Different proxies were also available for life span,  
<sup>153</sup> such as the age at sexual maturity or generation length. In such cases, I focused on the trait presenting the

154 highest coverage.

155 • **Body size**

156 Adult body mass estimates were compiled for mammals, birds and reptiles. Body length information  
157 was compiled for amphibians, because the coverage for body length was higher than that for body  
158 mass. Body mass and body length are known to scale allometrically, although the allometric rela-  
159 tionship differs across amphibian clades (Santini et al., 2018). In the amphibian dataset, Pearson’s  
160 correlation coefficient between log(Body mass) and log(Body length) was 0.71 (data points shown in  
161 Appendix 2, S2.2, Figure S2.2).

162 • **Longevity**

163 I defined longevity as the life span of an individual and maximum longevity as the longest life span  
164 reported. I used closely related traits when longevity/maximum longevity was not available or when  
165 longevity/maximum longevity had a poorer coverage than a related trait. I selected the age at sexual  
166 maturity for amphibians; Pearson’s correlation coefficient between log(Age at sexual maturity) and  
167 log(Maximum longevity) was 0.55 (Appendix 2, S2.2, Figure S2.2). I compiled the generation length  
168 for mammals and birds. The correlation between log(Generation length) and log(Longevity) was 0.74  
169 for mammals and 0.70 for birds (data points shown in Appendix 2, S2.2, Figure S2.3). Finally, I used  
170 maximum longevity directly for reptiles.

171 • **Litter or clutch size**

172 The number of offspring (litter size) or eggs (clutch size) was compiled directly from the sources and  
173 treated as equivalent across classes. I reported measures of central tendencies provided by the sources  
174 where applicable; otherwise, I calculated range midpoints (mean of smallest and largest reported lit-  
175 ter/clutch sizes).

176 • **Trophic level**

177 In all classes, species were described as omnivores, carnivores or herbivores. For reptiles and mam-  
178 mals, this information was compiled directly from the sources. For amphibians and birds, trophic  
179 levels were not provided. For these two classes, I inferred trophic levels from dietary information  
180 (Table 2.1). For birds, I used the primary diet (based on food items recorded as composing  $\geq 50\%$   
181 of the diet of a species). Diet for amphibians was described without respect to the percentage use of  
182 food items; simply as a binary record of whether or not food items were used. In both cases, species  
183 recorded to only consume plant-based resources (seeds, nectar, fruit or other plant material) were

184 classified as herbivores. Species consuming only animal resources (invertebrates or vertebrates) were  
185 classified as carnivores. Species consuming a mixture of plant and animal resources were classified as  
186 omnivores.

187 • **Activity time**

188 Species were described as being either nocturnal or non-nocturnal. Despite a higher resolution of ac-  
189 tivity time information in some of the sources (e.g., species being described as cathemeral, crepuscular  
190 or diurnal), I adopted the classification of the source with the lowest resolution (EltonTraits: Wilman  
191 et al. (2014), for birds), in order to have consistent information across classes. As such, all species  
192 defined as diurnal, cathemeral or crepuscular were classified as non-nocturnal, as opposed to species  
193 classified as strictly nocturnal.

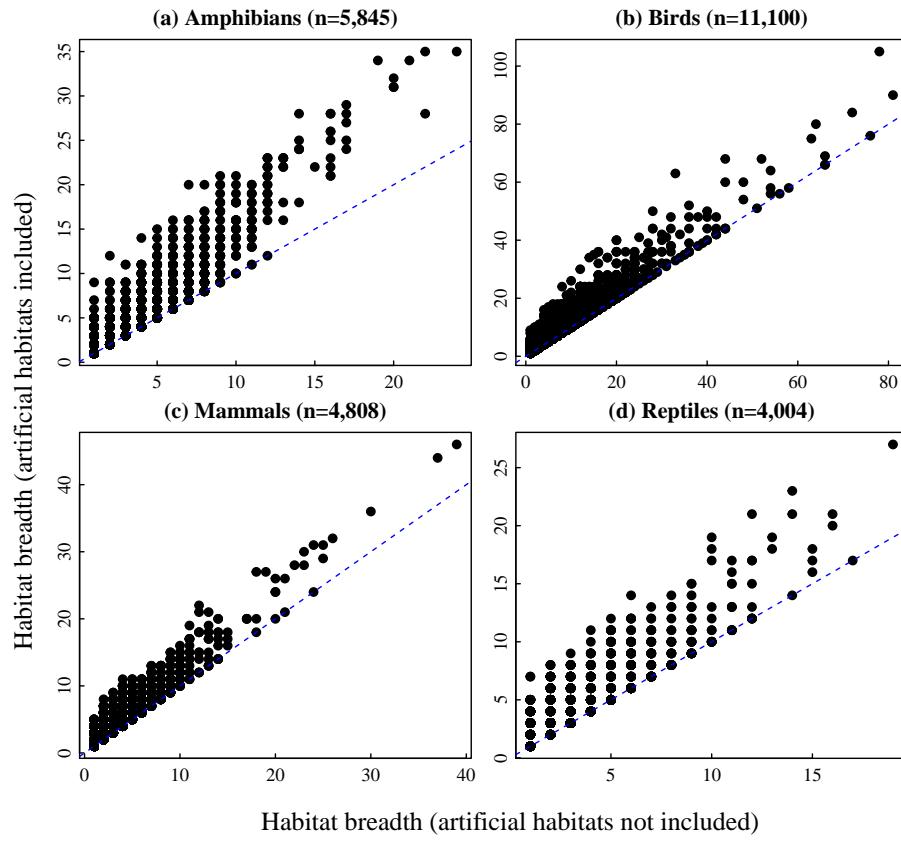
194 • **Habitat breadth**

195 I used IUCN habitat data (IUCN, 2020), which describe species habitat preferences and the suitabil-  
196 ity and importance of each habitat. I defined habitat breadth as the number of habitats a species was  
197 known to use, using level 2 of the IUCN Habitat Classification Scheme for description of habitat  
198 types (divided into: Forest, Savanna, Shrubland, Grassland, Wetland, Rocky areas, Caves and subter-  
199 ranean, Desert, Marine, Marine intertidal or coastal/supratidal, Artificial, Introduced vegetation, and  
200 Other/Unknown.) Note that the total number of habitats, determined by including those that qualify as  
201 artificial, correlates positively with the number of natural habitats used (Figure 2.2).

202 • **Use of artificial habitats**

203 For a species, I recorded whether any artificial habitat was reported to be suitable in the IUCN habitat  
204 data.

205 Finally, the compiled datasets contain an additional column, ‘Note’, where I reported species found to  
206 be extinct or extinct in the wild (EW). I used species Red List status and information from Meiri (2018) to  
207 flag such species. I reported 75 extinct/EW species for mammals, 160 for birds, 34 for amphibians and 53  
208 for reptiles. It is likely that the datasets contain extinct species that I could not flag, because they were not  
209 recorded as extinct in the sources I used.



**Figure 2.2: Number of natural and artificial habitats used by a species against number of strictly natural habitats used by a species.** Pearson's correlation coefficients show a high positive correlation between these two metrics of habitat breadth in all terrestrial vertebrate classes: 0.92 for amphibians (a), 0.95 for birds (b), 0.94 for mammals (c), and 0.90 for reptiles (d).

## 210 Phylogenies

211 I used class-specific phylogenetic trees downloaded on 13 April 2020. For mammals, I used ‘complete’  
 212 trees from Faurby et al. (2018, 2020), downloaded from <https://zenodo.org/record/3690867#.Xyc5wyhKhPZ>. For amphibians, birds and squamates, I obtained trees from <https://data.vertlife.org/>. The original sources were as follows: Jetz et al. (2012) for birds; Jetz and Pyron (2018) for amphibians; and Tonini et al. (2016) for squamates. For each class, a distribution of 1,000 trees was available.  
 213  
 214  
 215  
 216  
 217 For plotting purposes, I obtained consensus trees using the TreeAnnotator program of the BEAST software (Bouckaert et al., 2019).

## 218 Species distributions

219 I obtained extent-of-occurrence distribution maps for reptiles from Roll et al. (2017), available at: <https://datadryad.org/stash/dataset/doi:10.5061/dryad.83s7k> (downloaded 13 April 2020).  
 220

221 For mammals and amphibians, species distribution maps were obtained from the IUCN Red List (IUCN  
222 (2020), downloaded 13 April 2020); for birds, they were obtained from BirdLife International (<http://datazone.birdlife.org/species/requestdis>, downloaded 17 April 2020).

224 For amphibians, mammals and birds, I selected areas of extant or probably extant presence only. Additionally,  
225 I selected areas where species were resident or present during the breeding season, and I excluded  
226 areas occupied during the non-breeding season or where species were considered vagrant.

227 In addition, for all classes, I excluded occupied areas that fell outside the known elevational limits of  
228 species, where such data were available. Lower and upper elevational limits were retrieved from the IUCN  
229 Red List (queried using the rredlist package) and were available for approximately half of the species (Sup-  
230 porting Information, Appendix 2, S2.3, Figure S2.4). Decreases in range sizes were observed after cutting  
231 distribution maps by the known elevational limits (Appendix 2, S2.3, Figure S2.5).

### 232 **2.2.2 Investigating gaps and biases in trait data**

233 I used trait coverage and completeness to investigate taxonomic, phylogenetic and spatial biases in the trait  
234 data. Table 2.2 summarizes the sample sizes (number of species) in each of the following analyses. Note  
235 that species for which completeness was 0% were included in all analyses (for more details, see Figure 2.3).  
236 Also note that I did not filter out species identified as extinct or extinct in the wild, because they represented  
237 a small proportion of the datasets (0.48% for amphibians, 1.4% for both birds and mammals, and 0.50% for  
238 reptiles) and also because I could not exclude such species systematically, because it is likely that I did not  
239 flag them all.

**Table 2.2: Number of species for each analysis.** All species represented in the trait datasets were included in (1). All species from the class-specific phylogenetic trees or from the distribution maps that matched with species in the trait datasets were included in (2) and (3). *Table reproduced from Etard et al. (2020).*

	(1) Taxonomic biases	(2) Phylogenetic biases	(3) Spatial biases
Amphibians	6,990	6,170	5,650
Birds	11,634	8,315	10,802
Mammals	5,381	5,171	5,046
Reptiles	10,612	9,404	9,382

240 **Taxonomic biases**

241 I tested whether completeness varied across taxonomic class using pairwise Wilcoxon rank sum tests. I tested  
242 for the extent and performance of the taxonomic corrections by looking at trait coverage when taxonomic  
243 corrections are applied and when no correction is applied (Appendix 2, S2.4, Figure S2.6).

244 **Phylogenetic biases**

245 Initially, to assess whether more closely related species were more likely to be similar in trait completeness,  
246 I estimated the phylogenetic signal in completeness with Pagel's  $\lambda$  (Pagel, 1999) in each class. I used a  
247 bootstrapping approach, calculating  $\lambda$  for each of 50 trees randomly sampled in each class (using the phylosig  
248 function of the phytools R package; Revell, 2012). I then estimated the mean and 95% confidence intervals  
249 (95% CIs) of  $\lambda$ . Sample sizes for computing  $\lambda$  (number of species represented in both the phylogenies and  
250 trait datasets) are shown in Table 2.2.

251 I then plotted within-family median completeness in phylogenetic trees built at the family level, using  
252 the consensus trees. Within-family median completeness was calculated using taxonomic information in the  
253 trait datasets (sample sizes shown in Table 2.2).

254 **Spatial biases**

255 I first investigated whether wider-ranging species were more likely to be better sampled than narrow-ranging  
256 species. I tested for a relationship between species range size and trait completeness. I fitted a generalized  
257 linear model with a Poisson error distribution (directly using the number of sampled traits, ' $N_{traits}$ ', rather  
258 than the proportion (completeness)). Class was added as a predictor interacting with range size; thus the  
259 model was:

260 
$$N_{traits} \sim \log(\text{Range size}) * \text{Class}.$$

261 Second, I mapped assemblage-level median completeness. Assemblages were characterized at the pixel level  
262 at 50 km<sup>2</sup> resolution. I determined pixel-level composition and richness by stacking species geographical  
263 distributions. I then calculated median completeness across species in each pixel. I show the resulting  
264 maps for herptiles in the main text, and for mammals and birds in Appendix 2 (S2.5, Figure S2.7; median  
265 completeness was very high across most pixels for mammals and birds). In addition, I provide maps of  
266 assemblage-level mean completeness and standard deviation for all classes in Appendix 2 (S2.5; Figures  
267 S2.8 and S2.9 show corresponding maps; Figure S2.10 shows standard deviation against species richness).

I then tested for a spatial correlation between species richness and median completeness. Given that median completeness was very high across most pixels for mammals and birds, I fitted such models for herptiles only. I fitted spatial autoregressive lag models to explain assemblage-level median completeness as a function of species richness (using the function `lagsarlm` of the `spatialreg` package (Bivand and Piras, 2015; Bivand et al., 2013a; Bivand et al., 2013b)). Given that responses could vary geographically, I included the biogeographical realm as an interacting factor (using the World Wide Fund for Nature (WWF) ecoregion shapefile to characterise realms, obtained from <https://www.worldwildlife.org/publications/terrestrial-ecoregions-of-the-world>); the considered realms were Afrotropics, Australasia, Indo-Malayan, Nearctic, Neotropics and Palaearctic. To improve normality, I arc-sin square-root transformed completeness values and log-transformed species richness. The `lagsarlm` function allows for a consideration of spatial autocorrelation in the dependent variable by estimating the autoregressive lag coefficient,  $\rho$ , associated with an n-by-n matrix of spatial weights,  $W$ . The final model was:

$$\text{arcsin}(\sqrt{\text{Completeness}}) \sim \log(\text{Species richness}) * \text{realm} + \rho \cdot W \cdot \text{arcsin}(\sqrt{\text{Completeness}}).$$

The value of  $W$  was estimated using the functions `tri2nb` and `nb2listw` of the `spdep` package (Bivand and Wong, 2018; Bivand et al., 2013b). Fitting the model using all grid cells was computationally intractable; therefore, I randomly sampled cells for this analysis (using 30% of the grid cells in each realm). I selected grid cells where species richness was higher than three to avoid sampling issues. I fitted separate models for amphibians and reptiles, because when adding class as an interacting predictor, the same cells (with the same coordinates) might be sampled for multiple classes, whereas the `tri2nb` function does not tolerate duplicated coordinates.

## 2.3 Results

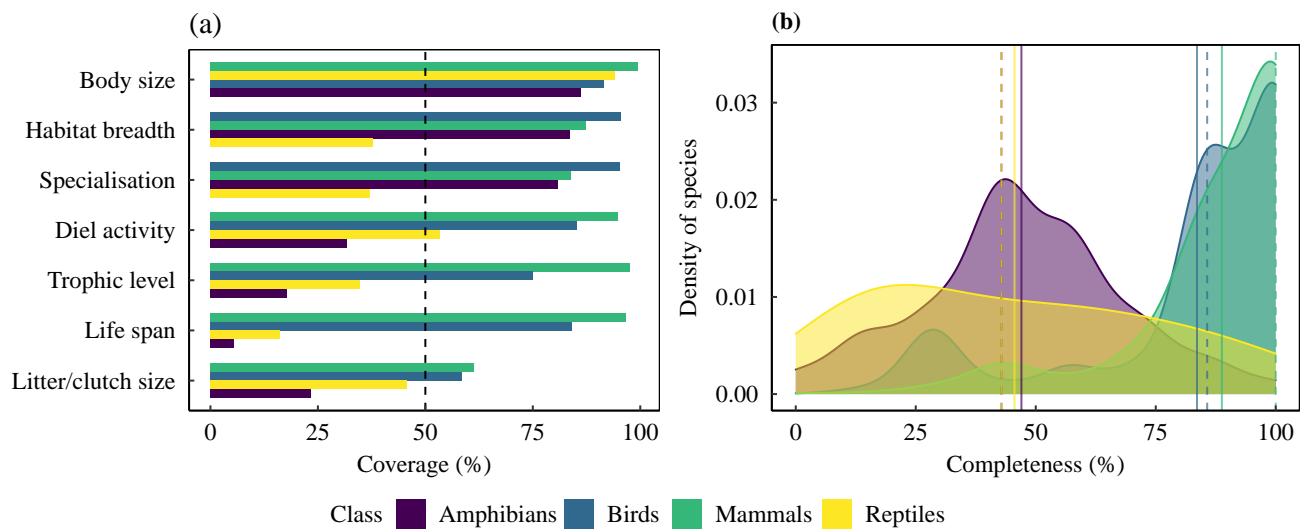
### 2.3.1 Taxonomic biases in trait information

Trait coverage for mammals and birds was overall high (Figure 2.3(a); mean and median coverage across traits: 89% and 95% for mammals; 84% and 85% for birds). In both cases, litter/clutch size was the trait with the poorest coverage (61% for mammals and 59% for birds). Coverage exceeded 80% for all other traits (except trophic level for birds, at 75% coverage).

Conversely, trait coverage was more variable for herptiles, and poorer overall (Figure 2.3(a); mean and median trait coverage: 47% and 32% for amphibians, 46% and 38% for reptiles). Coverage exceeded 80% only for body size in both reptiles and amphibians and for habitat related traits in amphibians only. In all

other cases coverage was <55%, with very little information available for longevity-related traits.

Trait completeness (proportion of non-missing trait values for a species) reflected similar biases (Figure 2.3(b)). The distribution of trait completeness varied significantly among classes (pairwise Wilcoxon rank sum test: p-value<0.0001 in all cases). Distributions were highly left skewed in mammals and birds (skewness: -2 and -1.6). 84% of all mammalian species and 80% of avian species fell in the 80–100% completeness range. Moreover, the completeness distribution was moderately right skewed for reptiles (skewness: 0.4), and slightly right skewed for amphibians (skewness: 0.02). 56% of all reptiles and 57% of amphibians fell in the 0-50% completeness range.



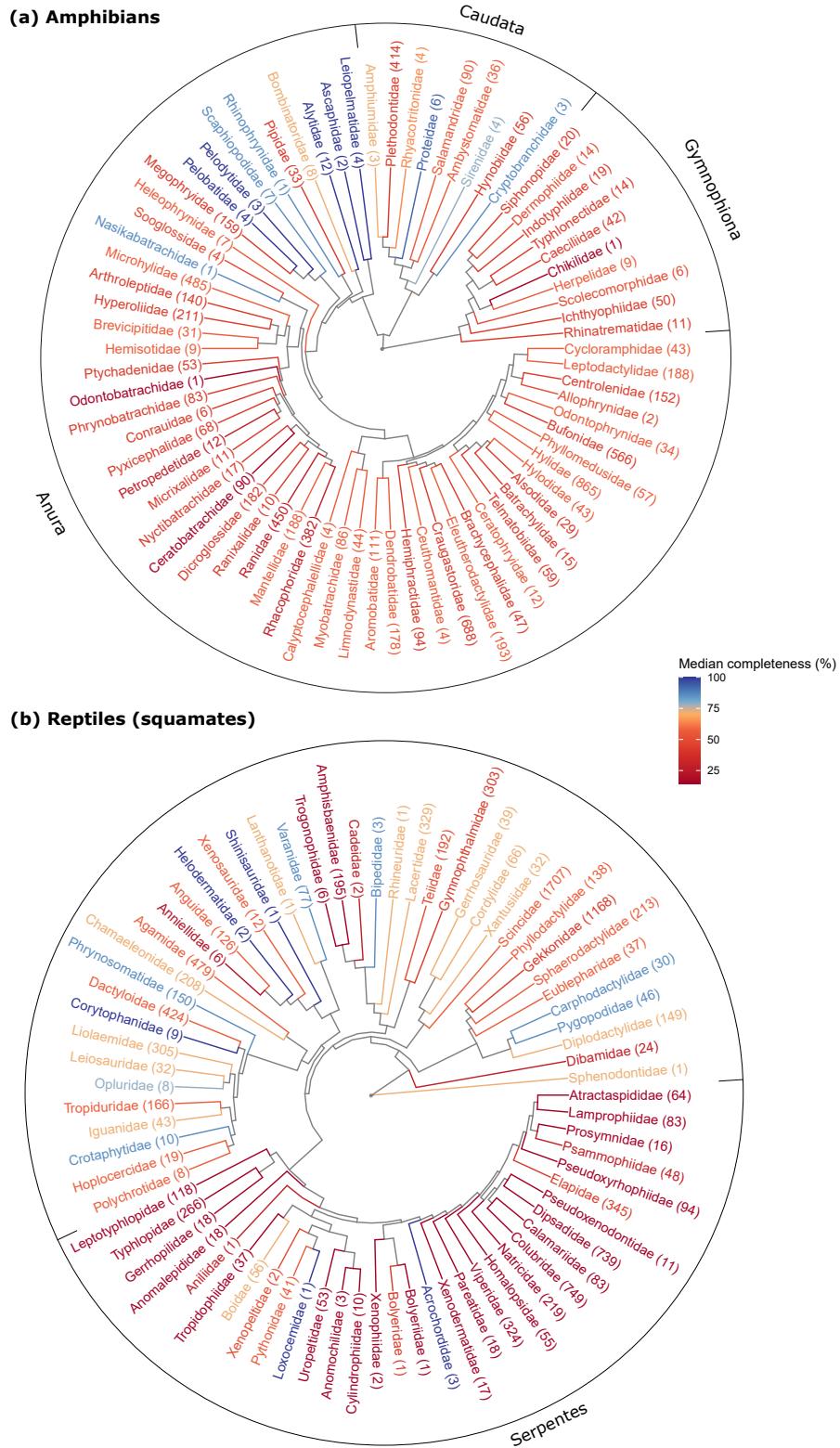
**Figure 2.3: Trait coverage and completeness across species.** (a) I defined coverage as the proportion of species for which an estimate is available for a given trait. The dashed line represents 50% coverage. (b) Trait completeness is the proportion of estimated traits for a species. Here, I show the distribution of completeness. Continuous lines represent the mean trait completeness for each class, whereas dashed lines represent the median trait completeness. Note that there were species with 0% completeness (230 species for amphibians – 3.3% of amphibian species in the trait dataset; 9 for birds – 0.077% of species; 7 for mammals – 0.13% of species; and 161 for reptiles – 1.5% of species). Species with 0% completeness were retained in the datasets when there was information for traits I did not select in the analyses, but no known value for the traits I did select. For instance, the body mass of the amphibian species *Rhinella centralis* was known, but other trait values (including body length) were missing, meaning that *Rhinella centralis* had 0% completeness for the set of traits I considered. *Figure reproduced from Etard et al. (2020).*

### 2.3.2 Phylogenetic biases in trait completeness

As expected from the distribution of trait completeness in mammals and birds (Figure 2.3), within-family median trait completeness was high across most tips of the phylogenetic trees (Appendix 2, Figures S2.11 and S2.12; I present the avian and mammalian phylogenies in the Appendix because there was little variation in completeness across tips). For birds,  $\lambda$  was 0.71 ( $\pm 0.0053$ ). For mammals,  $\lambda$  was 0.78 ( $\pm 0.0035$ ). This indicated that, despite completeness generally being high across tips, the sampling was not evenly distributed across the phylogeny.

312 In herptiles, clusters of families with similar median trait completeness appeared (Figure 2.4). In am-  
313 phibians, groups of families belonging to the order Anura (frogs) showed both the best and worst median  
314 completeness (Figure 2.4(a)). The best-sampled families included the tailed frogs of the family Ascaphidae  
315 (two species) and species of the family Leiopelmatidae (four species endemic to New Zealand). The family  
316 Ceratobatrachidae (containing *c.* 90 species occurring in Southeast Asia and in some Pacific islands), the  
317 family Ranidae (true frogs, 450 species considered here) and the family Rhacophoridae (shrub frogs, 382  
318 species considered here) figured among the worst-sampled families. For amphibians,  $\lambda$  was  $0.63 (\pm 0.0039)$ .  
319 In reptiles, most snakes were poorly sampled, whereas families in other suborders appeared to be sampled  
320 better overall (Figure 2.4(b)). Within snakes, the pythons, boas, the three species of the family Acrochordidi-  
321 dae and the python-like species of the family Loxocemidae were better sampled than other snake families. In  
322 reptiles,  $\lambda$  was  $0.69 (\pm 0.0032)$ . The sampling in herptiles was thus also uneven with regard to the phylogeny.

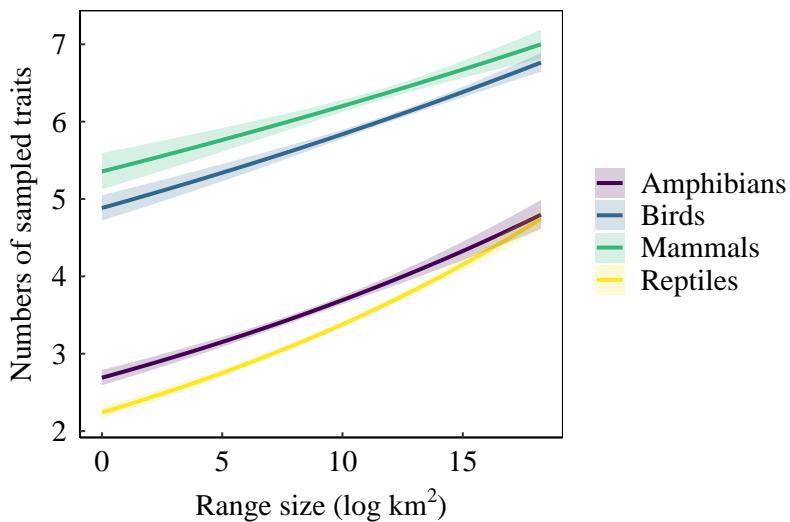
323 It is important to underline that Figure 2.4 shows within-family median completeness, masking the con-  
324 siderable variation in species richness across families, hence masking potential important variation in com-  
325 pleteness across species within families. For example, in the amphibian family Allophrynididae (three recog-  
326 nized species), the within-family median completeness was 50%; but the dataset comprised two species of  
327 completeness 14% and 86%, respectively. I present similar plots to those in Figure 2.4 showing the within-  
328 family standard deviation in completeness in Appendix 2 (Figure S2.13). Within-family standard deviation  
329 tended to increase with within-family species richness (Appendix 2, Figure S2.14).



**Figure 2.4: Within-family median trait completeness in herptiles.** The number next to each family name represents the number of species included in the calculation of the median. *Figure reproduced from Etard et al. (2020).*

330 **2.3.3 Spatial biases in trait completeness**

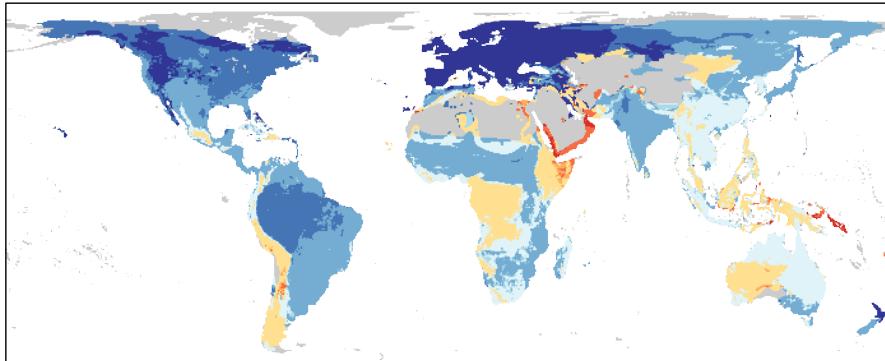
331 Range size was significantly correlated with the number of sampled traits. Larger range sizes were associated  
332 with a higher number of sampled traits (i.e., with higher completeness; Figure 2.5; Appendix 2, S2.7, Table  
333 S2.1). Similar results were obtained when using distribution maps not cut by elevational limits (Appendix 2,  
334 S2.7, Table S2.2; Figure S2.15). The rate of increase was steepest for reptiles, then for amphibians, then for  
335 birds and mammals (slope estimates for birds and mammals were not significantly different from each other;  
336 Appendix 2, Table S2.1).



**Figure 2.5: Relationship between number of sampled traits and geographical range size.** Models were fitted using a Poisson error distribution. Class was added as a predictor interacting with range size. Rates of increase in number of sampled traits with range size were not significantly different for mammals and birds but differed for reptiles and amphibians, with the steepest rates of increase for reptiles. *Figure reproduced from Etard et al. (2020).*

337 There were marked spatial variations in median trait completeness in herptiles (Figure 2.6). North Amer-  
338 ica and Europe were well sampled for both amphibians and reptiles. Moreover, Southeast Asia and the Congo  
339 basin were on average less well sampled. In other regions, contrasting patterns emerged between amphibians  
340 and reptiles. For instance, median completeness was poorer for amphibians than for reptiles in Australia,  
341 but opposite patterns were observed in South America. As in the phylogenetic analyses, assemblage-level  
342 median completeness could mask potential important variation in completeness within species of a given  
343 assemblage. Assemblage-level mean and standard deviation maps are shown in Appendix 2 (S2.5, Figures  
344 S2.8 and S2.9). There was a trend for increasing standard deviation with increasing species richness, with a  
345 larger spread in standard deviation at lower species richness (Appendix 2, section S2.5, Figure S2.10).

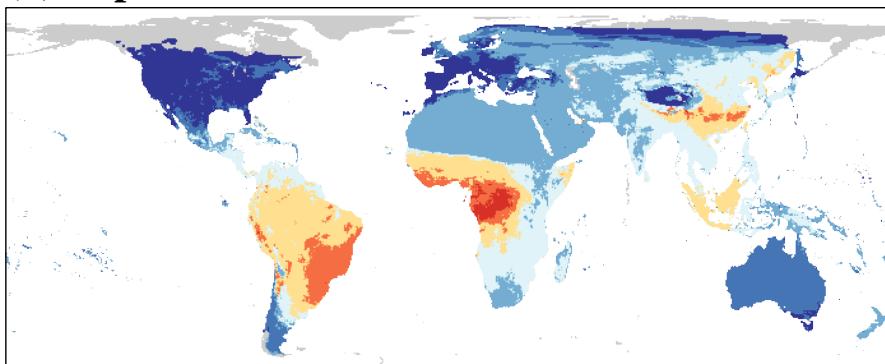
### (a) Amphibians



Median completeness:

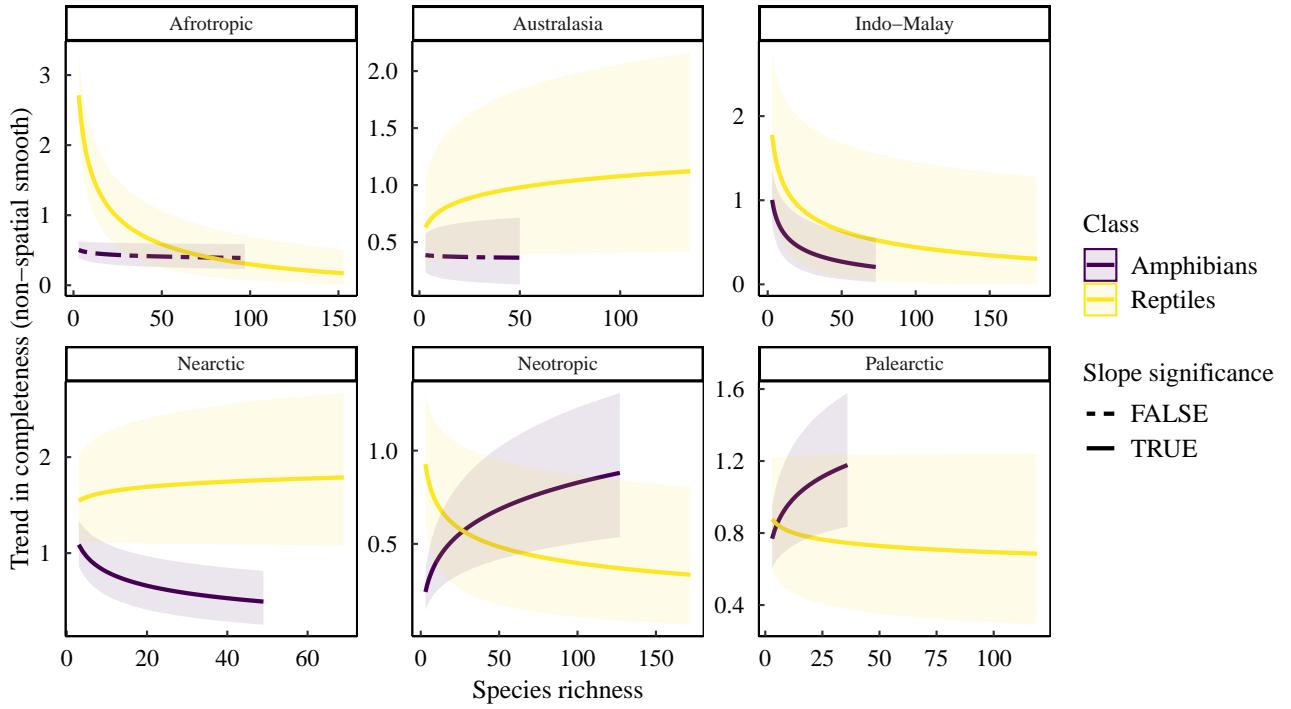
NA
0 – 10%
10 – 20%
20 – 30%
30 – 40%
40 – 50%
50 – 60%
60 – 70%
70 – 80%
80 – 90%
90 – 100%

### (b) Reptiles



**Figure 2.6: Spatial distribution of assemblage-level median trait completeness in herptiles.** Similar maps for birds and mammals are shown in Appendix 2 (S2.5, Figure S2.7). *Figure reproduced from Etard et al. (2020).*

346 Spatial models showed that species richness explained median trait completeness in herptiles in most  
347 realms (Figure 2.7; Appendix 2, S2.8, Tables S2.3 and S2.4); including spatial lags improved the models  
348 (amphibians:  $\rho = 0.92$ , p-value < 0.0001; reptiles:  $\rho = 0.91$ , p-value < 0.0001). For reptiles, completeness  
349 was negatively correlated with species richness in the most species-rich realms (Afrotropics, Indo-Malayan  
350 and Neotropics) and in the Palaearctic; the relationship was steepest in the Afrotropics and shallowest in the  
351 Palaearctic. In the Australasian and Nearctic realms, completeness tended to increase with species richness.  
352 For amphibians, negative relationships were observed in the Indo-Malay and Nearctic realms, whereas pos-  
353 itive trends were observed in the Neotropics and the Palaearctic. The opposite trends between reptiles and  
354 amphibians observed in the Australasian and Neotropical realms reflected patterns observed on the maps.  
355 The Indo-Malayan was the only realm where median completeness tended to decrease with species richness  
356 for both reptiles and amphibians.



**Figure 2.7: Spatial model trends for herptiles.** The lines represent in-sample predictions ( $\pm$  standard error) for the trend components of the spatial models (trends after accounting for spatial autocorrelation). *Figure reproduced from Etard et al. (2020).*

## 357 2.4 Discussion

358 The results of this Chapter illustrate the taxonomic, spatial and phylogenetic dimensions of the knowledge  
 359 gaps in trait data, termed the Raunkiær shortfall by Hortal et al. (2015). To the best of my knowledge,  
 360 this work constitutes the first comparative assessment of global gaps for terrestrial vertebrate trait data,  
 361 despite their use in numerous studies. I showed that the trait data present important taxonomic, spatial and  
 362 phylogenetic biases, with contrasts in the availability of trait information between, on the one hand, herptiles  
 363 and, on the other hand, birds and mammals.

364 Birds and mammals are globally well sampled for the set of traits I considered, even in the most species-  
 365 rich assemblages. Moreover, the availability of trait information for herptiles is lower overall and phyloge-  
 366 netically and geographically biased. Several factors could interplay to shape these patterns. For instance,  
 367 species that are more easily detectable (for example, wider ranging) and more charismatic are likely to be  
 368 better sampled. Diverse socio-economic predictors could also contribute to geographical biases in trait data  
 369 sampling; global biases in primary data collection are likely to be one of the most important contributors to  
 370 the patterns I highlighted. Nevertheless, biases in the data could have been introduced at later stages, notably  
 371 with the selection of sources and traits. The global compilation I obtained in this chapter reflects, in part, the

372 interest and focus of the secondary data sources I used. It is possible that the addition of new sources from  
373 regional journals or other authorities could diminish spatial biases in the data by increasing coverage for  
374 certain areas. Nevertheless, I argue that by focusing on widely used traits, these results are likely to reflect  
375 the “true” availability of the data in primary sources and that the shortfalls for other, less used traits would  
376 be more pronounced.

377 I believe that the results presented here are robust to taxonomic uncertainty, although taxonomic match-  
378 ing might potentially be improved further using class-specific sources, such as the Reptile Database or Am-  
379 phibiaWeb, for identification of synonyms (but see Appendix 2, S2.9, Figure S2.16). I have made two  
380 versions of the data compilations available, one in which my own corrections were applied and one using the  
381 original binomial names of the sources, meaning that users are free to use their own taxonomic backbones  
382 and identify synonyms within the compilations. I believe that taxonomic matching is a recurring issue when  
383 working across thousands of species. Taxonomic synonymy artefactually inflates the numbers of identified  
384 species, potentially lowering trait coverage (whereas clumping subspecies together can have the opposite  
385 effect). Tackling this problem is difficult (Isaac et al., 2004; Jones et al., 2012), notably because there is no  
386 global curated database recording the status of species names, and also because of the nature of taxonomy  
387 and the debates around the species concept (May, 2011). Nevertheless, taxonomic uncertainty can have  
388 important consequences. For instance, Cardoso et al. (2017) showed that inaccuracies and errors in species  
389 checklists contributed to the overestimation of plant diversity in the Amazon (but see Freeman and Pennell  
390 (2021): the relative underdescription of species in tropical areas compared to temperate areas –‘taxonomic  
391 debt’, also referred to as ‘latitudinal taxonomic gradient’ by the authors– may lead to the underestimation of  
392 species richness at low latitudes).

393 Biases in trait data have important implications for conservation planning. Past studies have shown  
394 that narrow-ranged species, for which fewer trait data are available on average, have higher extinction risks  
395 (Collen et al., 2016; Purvis et al., 2000; Ripple et al., 2017) and are more negatively impacted by anthro-  
396 pogenic pressures than wider-ranging species (Newbold et al., 2018). Trait information is also less available  
397 for herptiles in tropical regions such as the Congo basin, Southeast Asia and South America, which are some  
398 of the most diverse areas of crucial importance for worldwide conservation (Barlow et al., 2018). Conse-  
399 quently, trait information is on average less available where potentially more crucial to conservation plan-  
400 ning. Indeed, trait information can be incorporated into vulnerability assessments and, as such, can help to  
401 prioritize conservation efforts. Species traits have been found to mediate species responses to environmental  
402 changes across diverse taxonomic groups, and thus can inform on the sensitivity of species to anthropogenic  
403 pressures (Flynn et al., 2009; Newbold et al., 2013; Nowakowski et al., 2017). Traits are now commonly

404 used to estimate species vulnerability or extinction risks (Pacifici et al., 2015; Ramírez-Bautista et al., 2020).  
405 As opposed to trend-based approaches, which rely on historical population trends (changes in abundance  
406 or shifts in distributions) to predict species' vulnerability and extinction risks, trait-based approaches rely  
407 on species' intrinsic sensitivity to particular threats. The appeal of trait-based approaches to extinction risk  
408 estimation is that, by providing mechanistic insights, they diminish the amount of population information  
409 needed. If the responses of species to a threat consistently relate to certain traits, it is possible to generalize  
410 patterns across species for which population data are less available (Verberk et al., 2013). Integrating traits  
411 into vulnerability assessments is hence of particular interest when field monitoring of species population  
412 sizes or distributions is difficult to achieve, but biases in the data could mean that such information is lacking  
413 for some of the most vulnerable species.

414 Traits that influence species responses to environmental changes have been termed 'response traits' (or  
415 'response-mediating traits'; Luck et al. (2012)), as opposed to 'effect traits' that underpin ecosystem func-  
416 tioning (Lavorel and Garnier, 2002). For instance, relative brain size and longevity have been characterized  
417 as response traits in birds (Newbold et al., 2013; Sayol et al., 2020), whereas dietary characteristics (e.g.,  
418 trophic levels or guilds) are both response and effect traits. Hortal et al. (2015) highlighted that, for plants,  
419 both response and effect traits have been investigated, whereas for vertebrates the research has been more  
420 focused on understanding species responses. This could be because the way vertebrate traits interact to shape  
421 some ecosystem processes has not yet been characterized well.

422 Ecosystem processes sustained by animals might be harder to quantify and might be influenced by a  
423 combination of traits. The traits compiled in this work are likely to have a role in diverse processes. Never-  
424 theless, there was one important omission, in that I did not compile species diet in this chapter, potentially the  
425 most straightforward trait to link with diverse processes, such as grazing, pollination, scavenging and seed  
426 dispersal. From a practical perspective, I chose traits that had been estimated at least for some of the species  
427 in each class, and that were readily available. Diet was excluded because although estimates were available  
428 for amphibians, birds and mammals, there was no readily available database for reptilian diet. Movement or  
429 dispersal abilities were also excluded because information was not readily available for any class. Although  
430 I expect that species diet and dispersal abilities would present similar sampling biases to the ones presented  
431 in this work, the addition of such traits to the compilation would represent a valuable contribution and would  
432 notably facilitate studies looking at the functional roles of reptiles.

433 For practical reasons, I did not consider intraspecific trait variation. Intraspecific variation has been  
434 shown to have important effects on ecological systems, and a growing body of literature encourages trait-  
435 based research to include intraspecific variability (Guralnick et al., 2016). There have been several calls to

436 produce open-access, global trait datasets (Weiss and Ray, 2019), including a representation of intraspecific  
437 trait variation (Kissling et al., 2018). Notably, Schneider et al. (2019) designed a framework to store and  
438 share inter- and intraspecific trait data, accompanied by an R package to standardize the data in a proposed  
439 format. Such a proposition could constitute an important step towards the unification of individual datasets  
440 into a single, comprehensive database for ecological trait data.

441 The current spatial and taxonomic gaps in trait data might limit our ability to scale studies up, whereas  
442 biases in the data can affect the validity of extrapolations to groups or areas that are undersampled. More gen-  
443 erally, biases and gaps in biodiversity data can have important implications for ecological studies. Data gaps  
444 can hinder our ability to draw conclusions on observed macroecological patterns. For example, Chaudhary  
445 et al. (2016) proposed that marine species richness follows a bimodal distribution, peaking at mid-latitudinal  
446 locations, and argued that these patterns were not underpinned by knowledge gaps in species distributions.  
447 Nevertheless, Menegotto and Rangel (2018) attributed the tropical dip in marine species richness to a lack  
448 of species distribution data, explained by lower sampling efforts in tropical areas ('Wallacean' shortfall;  
449 Hortal et al. (2015)). Biases and gaps in trait data could also affect studies in closely related fields, such as  
450 functional ecology – for instance, past studies have shown that functional diversity indices are sensitive to  
451 missing data (Májeková et al., 2016; Pakeman, 2014) – or community assembly (Perronne et al., 2017).

452 Ecologists should, therefore, take particular care when designing trait-based studies, because both data  
453 quality and data gaps are likely to influence the results and the generality of the conclusions. There exist  
454 diverse methods to deal with missing trait values, should data missingness be problematic. Complete removal  
455 of missing values ('case deletion') is commonly used but presents several issues, because it reduces sample  
456 size and statistical power and introduces potential bias in data subsamples (Nakagawa and Freckleton, 2008).  
457 For example, retaining complete cases only from the trait datasets would generate trait data disproportionately  
458 representative of mammals and birds, which would be problematic for conducting cross-taxon analysis on  
459 terrestrial vertebrates. As such, it is recommended that case deletion be applied only when data are missing  
460 completely at random, which is rarely the case (Peugh and Enders, 2004).

461 Alternatives to case deletion consist of filling in the gaps. In recent years, the development of impu-  
462 tation techniques has provided robust methods to handle missing data. Such imputation techniques have  
463 been used to complete trait datasets in recent studies (Cooke et al., 2019a). Penone et al. (2014) used a  
464 simulation approach to evaluate the performance of four of these techniques, namely PhyloPars (Bruggeman  
465 et al., 2009), random forest algorithms as implemented in R with missForest (Stekhoven, 2016; Stekhoven  
466 and Bühlmann, 2012), multivariate imputation by chained equations (MICE; Buuren and Oudshoorn (2007))  
467 and k-nearest neighbour (kNN; Troyanskaya et al. (2001)). Penone et al. (2014) introduced missing values

468 (10%–80%) in a complete trait dataset of carnivorans and measured imputation performance in different  
469 scenarios. Given that phylogenetic non-randomness in missing trait values can impact imputation accuracy,  
470 Penone et al. (2014) removed values in three different ways (completely at random; with a phylogenetic  
471 bias; and with a body mass bias). Out of the four techniques, missForest and PhyloPars performed best when  
472 species phylogenetic position was included as a predictor of missing trait values. Such imputations appeared  
473 to be robust even when trait coverage was as low as 40%, which might be relevant for many reptilian and  
474 amphibian traits. The performance was not significantly affected by phylogenetic non-randomness of the  
475 data. Hence, missForest and PhyloPars appear to be well suited when traits are phylogenetically conserved,  
476 because they allow species phylogenetic position to be included as a predictor of missing trait values. The  
477 study by Penone et al. (2014) highlights that there are robust imputation techniques allowing to deal with  
478 incomplete trait data where biases might otherwise be problematic. Nevertheless, it is important to highlight  
479 that some imputation techniques, such as single or mean imputation, can be problematic because they do  
480 not allow an estimation of uncertainty and suffer from a lack of accuracy (Nakagawa and Freckleton, 2008);  
481 indeed, imputation techniques sometimes perform no better than case deletion. More work should be con-  
482 ducted to assess imputation performance in various contexts (see Johnson 2021), and the datasets compiled  
483 in this chapter might provide an opportunity for such studies.

484 Although robust imputation techniques can be useful for filling gaps in trait datasets, they are no substi-  
485 tute for continued data collection efforts. The results of this chapter show that data are particularly lacking  
486 in herptiles, notably in the Afrotropics, the Neotropics and the Indo-Malayan realms. For these areas, incor-  
487 porating regional databases into existing datasets could contribute to the reduction of global gaps. I believe  
488 that both primary research and subsequent efforts to integrate new data and existing databases are required  
489 if we are to collectively strive towards the unification of trait databases.

490 To conclude, this work constitutes, to my knowledge, the first assessment of the global gaps and biases  
491 in terrestrial vertebrate trait information. I show that herptiles are undersampled compared with mammals  
492 and birds, with important spatial and phylogenetic variability in the availability of trait information. Impu-  
493 tation techniques are one possible solution to these problems. Nevertheless, I believe that primary research,  
494 combined with efforts to complete existing datasets, is the only way to fill the current data gaps genuinely  
495 and robustly. I hope that the compiled trait dataset and these findings can prove useful for guiding further  
496 data collection efforts and for conducting macroecological analyses.

<sup>497</sup> **3 | Intensive human land uses negatively af-**

<sup>498</sup> **fect vertebrate functional diversity**

<sup>499</sup> **Keywords**

<sup>500</sup> Land use; land-use intensity; terrestrial vertebrates; functional diversity; traits.

<sup>501</sup> **Abstract**

<sup>502</sup> Land-use change is the leading driver of global biodiversity loss, thus characterising its impacts on the func-  
<sup>503</sup> tional structure of ecological communities is an urgent challenge. Using a database describing vertebrate  
<sup>504</sup> assemblages in different land uses, I assess how the type and intensity of land use affect the functional di-  
<sup>505</sup> versity of vertebrates globally. I find that human land uses alter local functional structure by driving declines  
<sup>506</sup> in functional diversity, with the strongest effects in the most disturbed land uses (intensely used urban sites,  
<sup>507</sup> cropland and pastures), and among amphibians and birds. Both tropical and temperate areas experience im-  
<sup>508</sup> portant functional losses, which are only partially offset by functional gains. Tropical assemblages are more  
<sup>509</sup> likely to show decreases in functional diversity that exceed those expected from species loss alone. These re-  
<sup>510</sup> sults indicate that land-use change non-randomly reshapes the functional structure of vertebrate assemblages,  
<sup>511</sup> raising concerns about the continuation of ecological processes sustained by vertebrates.

<sup>512</sup> **3.1 Introduction**

<sup>513</sup> Anthropogenic activities are profoundly transforming global biodiversity. Although multiple pressures act  
<sup>514</sup> in combination, land-use change currently poses the greatest threat to biodiversity (Maxwell et al., 2016;  
<sup>515</sup> Newbold et al., 2015). However, not all species respond similarly to land-use change. Traits have been found  
<sup>516</sup> to explain species' sensitivity to land-use change in diverse groups (Newbold et al., 2013; Nowakowski et

517 al., 2017; Quesnelle et al., 2014; Todd et al., 2017). Previous work has also shown that land-use change  
518 leads to non-random modification of assemblage trait composition (or functional diversity) (Chapman et al.,  
519 2018; Colin et al., 2018; Flynn et al., 2009; La Sorte et al., 2018; Newbold et al., 2013; Tinoco et al., 2018).  
520 Since it is widely acknowledged that biodiversity, and in particular trait diversity, may promote ecosystem  
521 functioning and stability, modification to the trait composition of assemblages could have far-reaching and  
522 adverse impacts on ecological processes (Hooper et al., 2012; Magioli et al., 2021; Oliver et al., 2015; Tilman  
523 and Downing, 1994).

524 Terrestrial vertebrates support many processes, ranging from pollination (Ratto et al., 2018), to seed  
525 dispersal to the regulation of lower trophic levels (Barber et al., 2010; Letnic et al., 2012; Salo et al., 2010;  
526 Zhang et al., 2018a). However, we lack a global understanding of how the functional diversity of entire  
527 vertebrate assemblages responds to changes in land use. Most previous studies have been conducted at  
528 regional or local scales (Davison et al., 2021), but these may not be representative of global patterns. Indeed,  
529 recent global syntheses have highlighted how biodiversity responses can differ substantially between regions  
530 and across latitudes, with higher sensitivity reported for the tropics (Matuoka et al., 2020; Millard et al.,  
531 2021; Newbold et al., 2020b). Another key issue is the taxonomic coverage of past work. Few studies  
532 investigating effects of land use on functional diversity have considered several vertebrate classes together,  
533 and comparative studies remain rare. Thus, how land-use change affects the functional diversity of local  
534 vertebrate assemblages at global scales, and the potential geographical and taxonomic variation in the effects,  
535 still largely remains to be explored.

536 Here, I aim to assess how human land use and land-use intensity affect the functional diversity of ver-  
537 tebrate assemblages, across and within taxonomic classes. Building on recent work (Matuoka et al., 2020;  
538 Millard et al., 2021; Newbold et al., 2020b), I investigate differences in response between tropical and tem-  
539 perate regions. I use multiple response metrics to quantify functional diversity. First, functional richness  
540 measures the breadth and variety of trait combinations represented in an assemblage (Legras et al., 2018).  
541 Second, functional dispersion quantifies how similar species in a given assemblage are in terms of their traits  
542 (Laliberte and Legendre, 2010). These metrics can mask important alterations of assemblage composition  
543 if functional losses are compensated for by functional gains. To address this, I consider pairwise measures  
544 between assemblages, to explore levels of functional loss and functional gain across land uses (Figure 3.1).

545 To this end, I combine (1) the trait data across terrestrial vertebrates collected in Chapter 2 (also published  
546 in Etard et al. (2020)), with (2) global records of species occurrence in eight land-use types of differing inten-  
547 sity of use (the PREDICTS database: Hudson et al. (2014, 2017), Figure 3.1, Table S3.1). The PREDICTS  
548 database is currently the most comprehensive database of sampled species occurrence, and for most records

549 also abundance, across multiple land uses of different land-use intensity. Using the PREDICTS database  
550 allows us to contrast biodiversity metrics among intact land uses (primary-vegetation sites, considered to be  
551 the undisturbed reference condition), and all other human land-use types. Specifically, I test the following  
552 hypotheses, both across and within taxonomic classes:

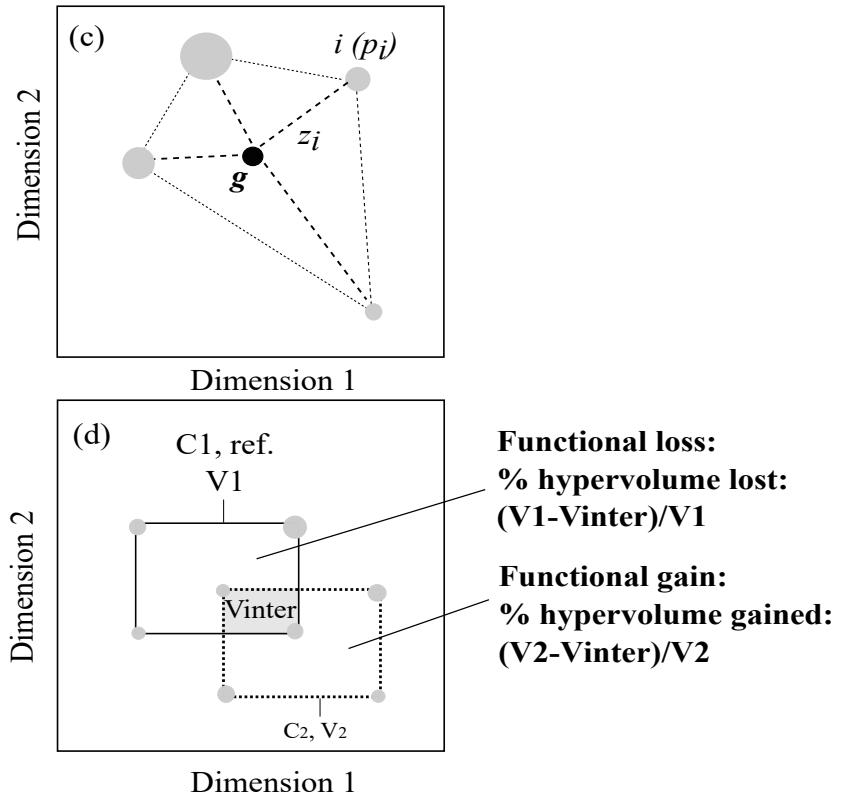
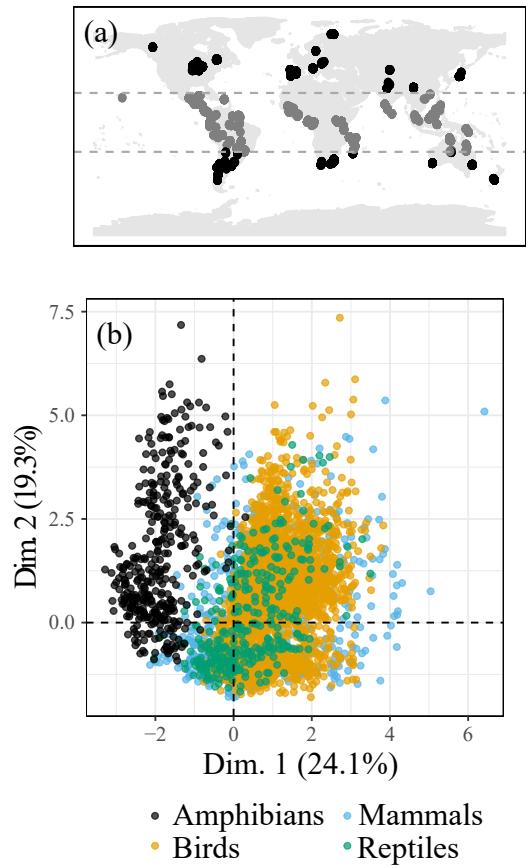
- 553 1. I expect decreases in functional diversity in human land uses compared to primary vegetation, caused  
554 by contractions of occupied trait space. I expect such effects to be more pronounced where land is  
555 used more intensively by humans. This hypothesis builds upon evidence that species with certain  
556 traits are more sensitive to land-use disturbance (Newbold et al., 2013), meaning that disturbed land  
557 uses will retain only disturbance-tolerant species, more functionally similar to one another. Given the  
558 reported higher sensitivity of tropical assemblages to land-use disturbance, I predict that such effects  
559 are stronger in the tropics.
- 560 2. I hypothesise that decreases in functional diversity in disturbed land uses exceed decreases expected  
561 by chance, given local species loss. Thus, I expect disturbed land uses to promote functional under-  
562 dispersion. Functional under-dispersion occurs when species within an assemblage are more similar,  
563 in term of their traits, than expected by chance (Cadotte and Tucker, 2017; Wong et al., 2018) – or,  
564 in other words, when functional dispersion is lower than expected given local species richness. I  
565 predict that under-dispersion is more likely to occur in the highly disturbed sites, in both temperate  
566 and tropical areas. This hypothesis is based on the idea that species are being removed non-randomly  
567 from sensitive areas of the trait space, and increasingly so with higher disturbance level.
- 568 3. Finally, I expect decreases in functional diversity in human land uses to be driven by high functional  
569 loss, whereby species are being removed from previously occupied areas of the trait space; I expect  
570 no functional gain. This hypothesis is based on the idea that the functional trait space in undisturbed  
571 land uses represents all of the possible regional trait combinations and that species with functional  
572 attributes rendering them unable to persist in altered conditions will be filtered out (Cornwell et al.,  
573 2006).

## 574 3.2 Methods

### 575 3.2.1 Vertebrate assemblages

576 I used vertebrate occurrence data from the PREDICTS database (Hudson et al., 2014, 2017), a collection of  
577 studies that recorded species occurrence across multiple land uses and land-use intensities. In PREDICTS,

- Tropical sites (3914)
- Temperate sites (2373)



**Figure 3.1: Overview of the study design and functional metrics.** I used occurrence data for vertebrate species from the PREDICTS database ((Hudson et al., 2014, 2017); 180 studies; 431,170 records; 4,339 species; 6,758 sampled sites). (a) shows the spatial distribution of sites I consider. I combine occurrence data with trait data compiled in Chapter 2 to calculate functional metrics. (b) is a representation of the trait data in two dimensions, plotted across PREDICTS vertebrates. Traits that contributed most to dimension 1 were lifespan (29%) and litter/clutch size (22%), while traits that contributed most to dimension 2 were habitat breadth (47%) and use of artificial habitats (35%). (c) and (d) present the conceptual framework for the calculation of the functional diversity metrics: local measures (c) and pairwise metrics (d). (c) Given a trait space, functional richness is calculated as the hypervolume occupied by the minimum convex hull encompassing all species Villéger et al., 2008. Functional dispersion is calculated as the mean distance of the species to the centroid,  $\mathbf{g}$  Laliberte and Legendre, 2010. (d) I compute functional loss as the proportion of hypervolume lost from the reference assemblage, and I define functional gain as the proportion of hypervolume of the disturbed assemblage that was gained (proportion of novel trait space in the disturbed assemblage). *Figure reproduced from Etard et al. (2021).*

578 each study contains several sites, which may be clustered into spatial blocks. Assemblage and land-use  
579 data are available at the site level: one site is characterised by a unique land use of given land-use intensity  
580 and provides occurrence data for a set of sampled taxa (and the same set of taxa is sought at all other sites  
581 within a study). Sites located between 23.5°N and 23.5°S of latitude were considered tropical, and otherwise  
582 temperate (Figure 3.1).

583 Land uses in PREDICTS were assigned to the following categories, based on the descriptions of the

584 habitat given by the original collectors of the data: primary vegetation (considered to be the undisturbed ref-  
585 erence); secondary vegetation; plantation forest; pasture; cropland; urban (considered human, or disturbed;  
586 Table S3.1; Hudson et al., 2014, 2017). Secondary vegetation is further divided into three categories: ma-  
587 ture, intermediate and young, depending on the stage of recovery of the vegetation. Land-use intensity is  
588 reported as minimal, light or intense, according to criteria that depended on the land-use type in question  
589 (e.g., crop diversity, degree of mechanisation and chemical inputs in cropland, or bushmeat harvesting and  
590 selective logging in primary vegetation; Hudson et al., 2014). I excluded sites for which the land use could  
591 not be characterised or for which the stage of recovery of secondary vegetation was unclear. As the PRE-  
592 DICTS database is a collection of independent studies, the design of this study was not balanced: the sample  
593 size varied across land uses (Figures S3.1, S3.2), and across taxonomic groups (3103 species of birds; 531  
594 mammals; 379 amphibians; 326 reptiles).

### 595 3.2.2 Functional traits and diversity indices

596 Trait choice is a critical step when calculating functional diversity metrics, which are highly sensitive to trait  
597 selection (Mouillot et al., 2021). However, trait selection trades off with data availability. Here, a constraint  
598 was to use similar traits across the different classes. I thus used the seven traits compiled in Chapter 2  
599 across terrestrial vertebrates. Most of these traits were available for at least 50% of the species in each class  
600 (except trophic level in amphibians and lifespan in herptiles; Figure S3.4). In addition, I chose these traits  
601 as they were ecologically relevant, thus I broadened the biological definition of traits (i.e., a characteristic  
602 measurable at the level of an individual) to include measures of habitat breadth and habitat specialisation (still  
603 theoretically measurable at the level of an individual). The final set constituted seven traits that influence  
604 species responses to environmental change: body mass, trophic level, lifespan, litter/clutch size, diel activity,  
605 habitat breadth and use of artificial habitats. These traits related to life-history, habitat specialisation and use  
606 of geographical space (e.g., habitat breadth is a significant predictor of geographical range size in all classes;  
607 Figure S3.3). Here, I did not consider estimations of dispersal abilities or home range size as these were  
608 available for a small fraction of the species (<3%, Alex Smith and M. Green, 2005; Paradis et al., 1998;  
609 Sutherland et al., 2000; Whitmee and Orme, 2013), neither did I include geographical range size which is  
610 measured across many individuals, and hence cannot be considered a trait. As in Chapter 2, I did not consider  
611 intraspecific trait variation, thus assuming no effect of the environment on trait values.

612 Trait coverage was variable among classes and traits, with important gaps for reptiles and amphibians  
613 (Figures S3.4, S3.5; Chapter 2; Etard et al. (2020)). I imputed missing trait values using random forest  
614 algorithms (missForest package: Stekhoven and Bühlmann (2012), Stekhoven (2016)), including traits, tax-

615 onomic order and phylogenetic eigenvectors as predictors (Debastiani et al., 2021; Penone et al., 2014). To  
616 further assess the sensitivity of the results to imputation (see next section), I imputed missing trait values  
617 eight times, thereby obtaining eight sets of imputed traits. I randomly selected one imputed trait set for the  
618 calculation of functional metrics. Imputations of missing trait values & imputation performance are detailed  
619 in Appendix S3.2 and Appendix S3.4 (and see Figures S3.6-S3.8). Post-imputation, continuous traits were  
620 log<sub>10</sub>-transformed (except habitat breadth which was square-rooted) and z-scored (standardised to unit vari-  
621 ance and zero mean). In addition, I also assessed whether the results were robust to imputation error using a  
622 subset of the PREDICTS data considering only species for which I had complete trait information (see next  
623 section).

624 Correlation among traits can be a safeguard against high sensitivity of functional metrics to trait omis-  
625 sion, notably where omitted traits correlate strongly with traits that are already included in the calculation  
626 (Mouillot et al., 2021). Nevertheless, high multicollinearity among traits has been reported as potentially  
627 problematic for the calculation of functional diversity (Cadotte et al., 2011). Thus, I verified that the degree  
628 of multicollinearity among traits was not problematically high (with a threshold of 5 for variance inflation  
629 factors, Table S3.3). Furthermore, I tested the sensitivity of the results to trait omission, by investigating  
630 whether adding geographical range size in the calculation of functional metrics was likely to affect the re-  
631 sults.

### 632 **3.2.3 Effects of land use and land-use intensity on FRic and FDis (Hypothesis 1)**

633 For each assemblage, I measured functional richness using ‘FRic’ (Villéger et al., 2008), and functional  
634 dispersion using ‘FDis’ (Laliberte and Legendre (2010); Figure 3.1), from the FD package (Laliberte and  
635 Legendre (2010); Laliberté et al. (2015)). I assessed the effects of land use, land-use intensity, and region  
636 (temperate versus tropical) on FRic and FDis across and within taxonomic classes using linear mixed-effects  
637 models (lme4 package, Bates et al. (2015)). Land use and land-use intensity were not ranked in the models.  
638 A random intercept of study identity accounted for variation in experimental design across studies, while  
639 a random intercept representing spatial blocks of sampled sites, nested within study, accounted for spatial  
640 structuring within studies. To improve normality and bound predictions between 0 and 1, I transformed FRic  
641 and FDis using an arcsin-square-root transformation. The best-fitting model was sought using backwards  
642 stepwise model selection, starting with the most complex model that included all two-way interactions among  
643 the specified main effects. Model fits were compared using likelihood-ratio tests at each iteration of the  
644 selection procedure.

645 Across vertebrates, the starting models included the effects of land use, land-use intensity and region

646 (temperate versus tropical). The best-fitting model for FRic was:

647  $\text{arcsin}(\sqrt{\text{FRic}}) \sim \text{Land use} + \text{Land-use intensity} + \text{Region} + \text{Land use : Land-use intensity} + \text{Land use : Region.}$

649 (Model 1a)

650 For FDis, the best-fitting model did not include interactions between land use and region, but the main  
651 effect of region was retained:

652  $\text{arcsin}(\sqrt{\text{FDis}}) \sim \text{Land use} + \text{Land-use intensity} + \text{Region} + \text{Land use : Land-use intensity.}$

653 (Model 1b)

654 To investigate differences in responses across classes, I pooled some of the land uses together, because  
655 otherwise, sample sizes would have been too low. Mature, intermediate and young secondary vegetation were  
656 grouped together as ‘Secondary vegetation’, and cropland and pasture were grouped together as ‘Agricultural  
657 land uses’. The starting models included the effects of land use, land-use intensity, region and taxonomic  
658 class. For FRic, the best model was:

659  $\text{arcsin}(\sqrt{\text{FRic}}) \sim \text{Land use} + \text{Land-use intensity} + \text{Region} + \text{Class} + \text{Land use : Land-use intensity} + \text{Land use : Class} + \text{Land-use intensity : Region} + \text{Class : Region.}$

661 (Model 2a)

662 For FDis, regional effects were dropped:

663  $\text{arcsin}(\sqrt{\text{FDis}}) \sim \text{Land use} + \text{Land-use intensity} + \text{Class} + \text{Land use : Land-use intensity} + \text{Land use : Class} + \text{Land-use intensity : Class.}$

665 (Model 2b)

666 To assess whether the results were robust to imputation error, I used a subset of the PREDICTS data  
667 considering only species for which there were complete trait information (6,212 sites; 442 mammals; 1,975  
668 birds; 78 reptiles; 9 amphibians), and I fitted models again to this data subset. I did not have enough complete  
669 trait data among amphibians to be able to consider this class separately, so I first considered amphibians and  
670 reptiles together (herptiles), and reptiles only. In addition, I complemented this validation with a sensitivity  
671 analysis to variation in imputed values. I calculated FDis and FRic using each of the eight imputed trait  
672 datasets and fitted the previous models to each set. I then qualitatively evaluated the congruence of the  
673 estimates from the different models. Finally, because there tended to be more sites sampled in primary  
674 vegetation than in other land uses (Figures S3.1, S3.1), I ran additional sensitivity tests to assess whether the

675 results were robust to resampling primary vegetation sites to a number equal to 50 (a sample size close to the  
676 median number of sites sampled in land uses other than primary vegetation in both regions (median = 37 for  
677 the temperate subset and 57 for the tropical subset, Figure S3.1)).

678 **3.2.4 Investigating functional under-dispersion (Hypothesis 2)**

679 To assess whether effects of land use and land-use intensity on FDis differed from what would be expected  
680 by chance given changes in local species richness, I generated null expectations of FDis at each site. I  
681 randomised assemblage composition 500 times, drawing species from the corresponding study's species  
682 pool while maintaining local species richness. For each site, I thus obtained a null distribution for FDis.  
683 Then, I tested whether FDis differed from null expectations using Wilcoxon signed-rank tests. I created  
684 a binary variable which was assigned 1 if FDis was significantly lower than null expectations at a given  
685 site (significant under-dispersion), and 0 otherwise. I investigated how land use, land-use intensity, region  
686 and taxonomic class affected the probability of occurrence of under-dispersion using a generalised linear  
687 mixed-effects model with a binomial distribution of errors. The best-fitting model did not retain any effect  
688 of taxonomic class:

689  $P_{\text{under-dispersion}} \sim \text{Land use} + \text{Land-use intensity} + \text{Region} + \text{Land use : Land-use intensity} + \text{Land use : Region}$ .

691 (Model 3)

692 **3.2.5 Functional loss and functional gain (Hypothesis 3)**

693 I calculated the proportion of trait space that was lost in disturbed land uses compared to reference land uses  
694 (functional loss) and the proportion of trait space that was gained in disturbed land uses (functional gain)  
695 (Figure 3.1(c)), across and within taxonomic classes. I selected studies where at least one site was sampled in  
696 primary vegetation. I then made within study pairwise comparisons between reference assemblages, sampled  
697 in primary vegetation, and disturbed assemblages. In addition, I considered all comparisons between pairs  
698 of primary-vegetation sites, to create reference pairs. I then investigated how land use, land-use intensity  
699 and region affected functional loss and gain across and within taxonomic classes using linear mixed-effects  
700 models, controlling for study identity in the random effects. Across vertebrates, the best-fitting model for  
701 functional loss was:

702  $\text{arcsin}(\sqrt{\text{loss}}) \sim \text{Land use} + \text{Land-use intensity} + \text{Region} + \text{Land use : Land-use intensity} + \text{Land use : Region}$ .

704 (Model 4a)

705 For functional gain, one interaction term (land use with region) was dropped:

706  $\arcsin(\sqrt{\text{gain}}) \sim \text{Land use} + \text{Land-use intensity} + \text{Region} + \text{Land use : Land-use intensity}.$

707 (Model 4b)

708 When considering the effects of taxonomic class, the best-fitting model for functional loss was:

709  $\arcsin(\sqrt{\text{loss}}) \sim \text{Land use} + \text{Land-use intensity} + \text{Class} + \text{Region} + \text{Land use :}$

710  $\text{Land-use intensity} + \text{Land use : Class} + \text{Land use : Region} + \text{Land-use intensity : Class}.$

711 (Model 5a)

712 For functional gain (Model 5b), the fitted effects were the same as those of Model 2b. More details about  
713 the calculation of functional loss and gain can be found in Appendix S3.5.

714 All data analyses were conducted using R version 3.5.1 (R Core Team, 2018). I made the code available  
715 on figshare (DOIs: <https://doi.org/10.6084/m9.figshare.14161883> and <https://doi.org/10.6084/m9.figshare.15163926>), as well as the main result datasets (<https://doi.org/10.6084/m9.figshare.15163971>).

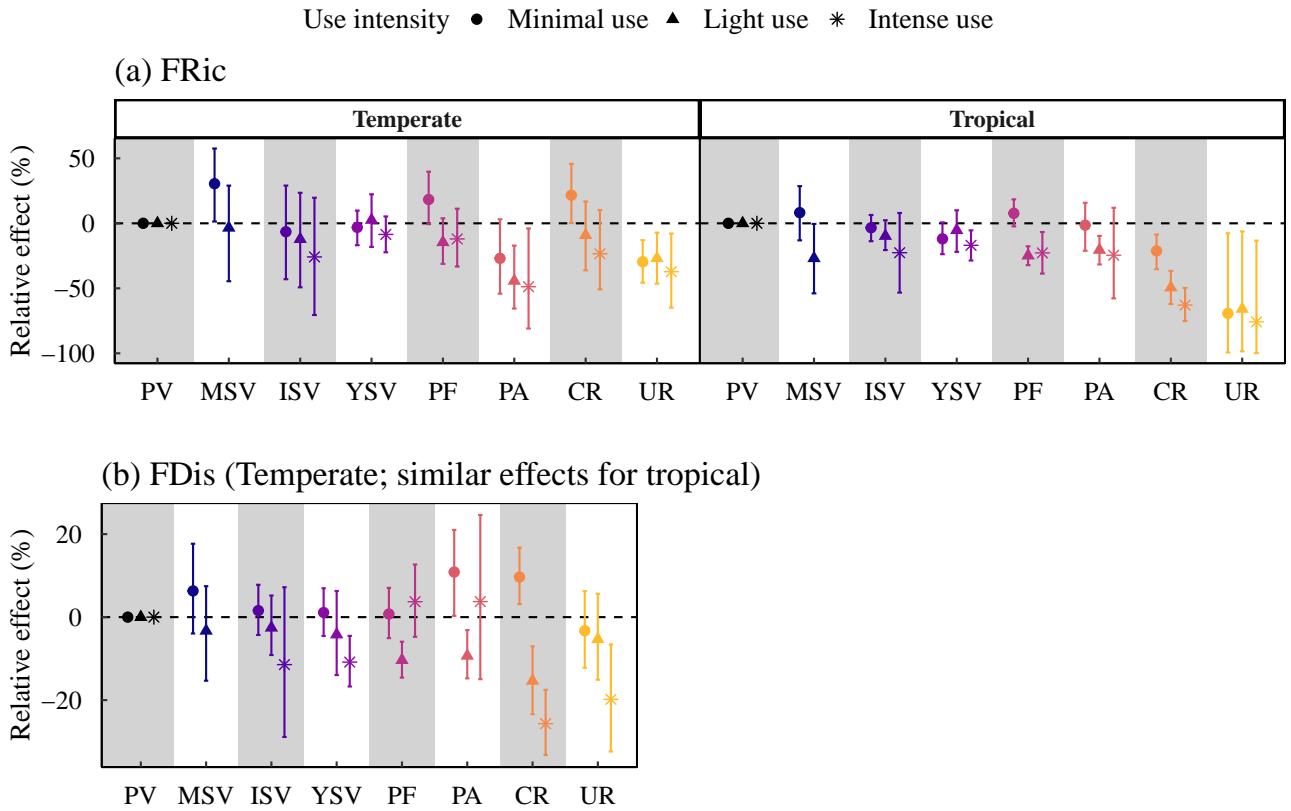
718 

### 3.3 Results

719 

#### 3.3.1 Effects of land use on FRic and FDis

720 Across all vertebrates, land use and land-use intensity significantly affected FRic and FDis (Figure 3.2).  
721 FRic tended to decrease with increasing disturbance level and higher intensity of land use. For FRic, relative  
722 effects differed between regions (Figure 3.2(a)). Although declines were overall more important for dis-  
723 turbed tropical assemblages, significant declines were observed for the temperate assemblages (e.g., a 37%  
724 average decline in intensely used urban areas; a 49% decline in pastoral areas of high land-use intensity).  
725 Nevertheless, tropical assemblages typically showed more important reductions in FRic. For instance, de-  
726 clines averaged 63% for intensely used tropical cropland and 76% for urban areas. For FDis, relative effects  
727 were similar in both regions (Figure 3.2(b)). The most important average declines were observed for urban  
728 assemblages of intense use (20% decline), and for lightly- and intensely used cropland (by 15% and 26%).  
729 Note that confidence intervals around the estimated average declines were large in some cases, highlighting  
730 some heterogeneity in the responses.

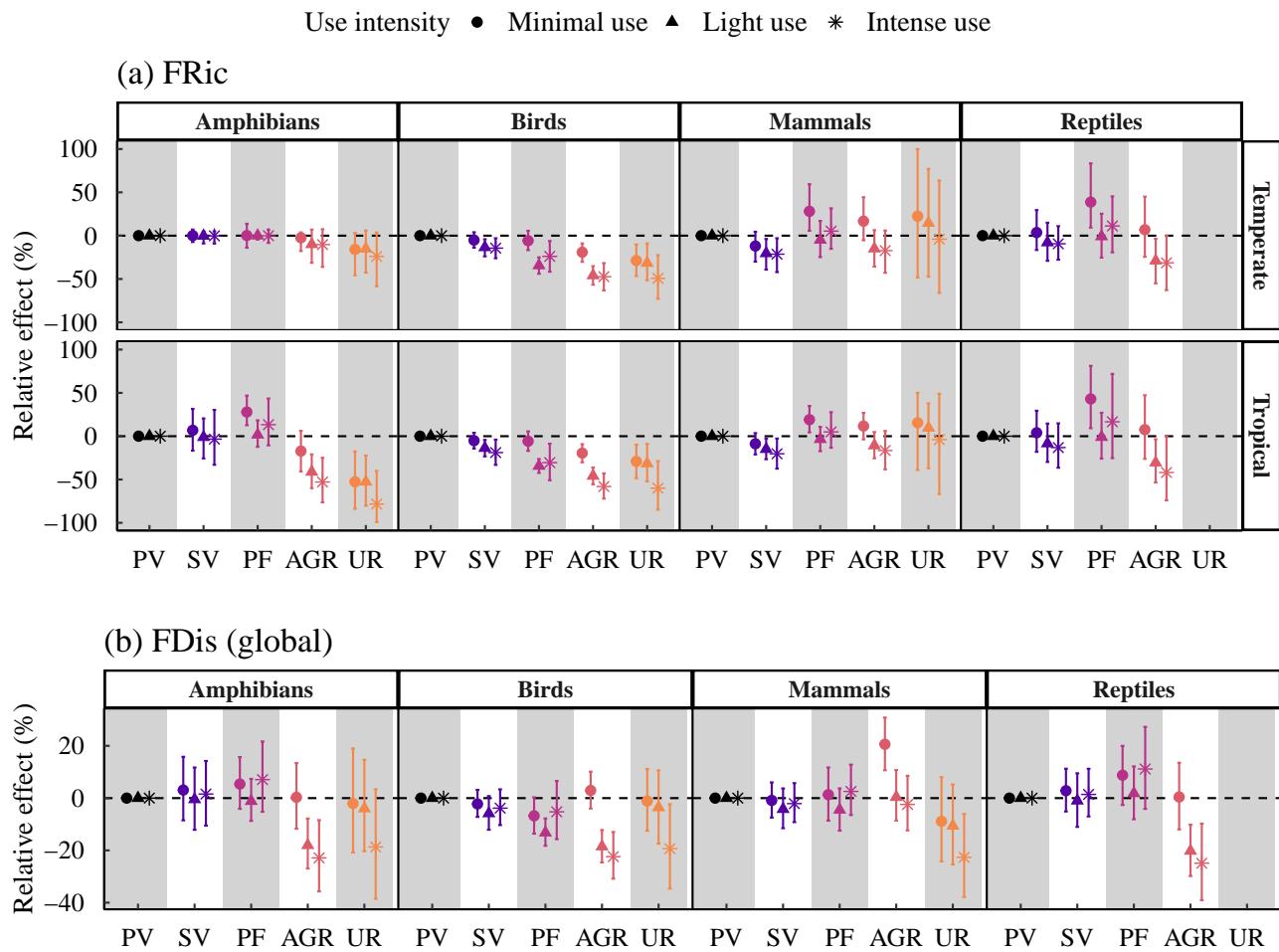


**Figure 3.2: Effects of land use, land-use intensity and region on FRic (a) and on FDis (b) across all PREDICTS vertebrates.** Effects were rescaled with reference to primary vegetation (PV) and are plotted as a % difference relative to PV within each land-use intensity. For FRic, the best-fitting model included interactions between land use and region, while these interactions were dropped for FDis, explaining the similar relative effects in both regions. Error bars represent 95% confidence intervals. MSV: mature secondary vegetation; ISV, intermediate secondary vegetation; YSV, young secondary vegetation; PF, plantation forest; PA, pasture; CR, cropland; UR, urban. Effects for intense use in MSV could not be estimated as there were not enough sampled sites. *Figure reproduced from Etard et al. (2021).*

731 Fitting the same models to the subset of species with complete trait data, I detected important declines in  
 732 functional diversity in a number of land uses, showing that the conclusions are robust to trait imputation un-  
 733 certainty (for example, FRic declined on average by 75% in intensely used temperate pastoral assemblages;  
 734 by 48% for intensely used tropical cropland; and FDis declined by an average 37% in intensely used tropical  
 735 urban assemblages; Figure S3.18). Furthermore, using the subset of species with complete trait data, I found  
 736 that the results were not sensitive to the inclusion of geographical range size as an additional trait (Figure  
 737 S3.19). Finally, the results were not sensitive to variation across imputed trait values (Figure S3.20) and were  
 738 also robust to resampling in primary-vegetation sites (Figure S3.21).

739 Responses of FRic and FDis to land use and land-use intensity differed among taxonomic classes (Figure  
 740 3.3). Within-class effects for FDis were similar between regions. The most notable decreases were observed  
 741 in lightly- and intensely used agricultural land uses in amphibians, birds and reptiles; and in intensely used

742 urban land uses for birds and mammals. For FRic, the effects in tropical and temperate regions were qualitatively similar in three out of four classes (birds, mammals and reptiles), although effect sizes tended to be  
 743 bigger for tropical assemblages. Birds and reptiles showed reductions in disturbed land uses in both tropical  
 744 and temperate regions, whereas I detected few significant effects for mammals. For birds, the most important  
 745 average decline, of 50%, was observed in intensely used tropical urban land uses, while for reptiles I detected  
 746 significant decreases in lightly- and intensely used agricultural sites (but I could not estimate effects for urban  
 747 land uses due to the small sample size). Finally, the effects differed between tropical and temperate regions  
 748 for amphibians, with no significant effects detected across temperate assemblages, but important reductions  
 749 across tropical agricultural and urban assemblages.  
 750



**Figure 3.3: Effects of land use, land-use intensity, region and taxonomic class on FRic (a) and FDis (b).** Effects were rescaled with reference to primary vegetation (PV) and are plotted as a % difference relative to PV within each land-use intensity. Error bars represent 95% confidence intervals. Effects for FRic were estimated from Model 2a, and from Model 2b for FDis. SV, secondary vegetation; PF, plantation forest; AGR, agricultural land uses (pasture and cropland); UR, urban. Effects for reptiles in urban land uses could not be estimated as there were not enough sampled sites. *Figure reproduced from Etard et al. (2021).*

751 Fitting similar models only for species with complete trait data showed that these patterns are unlikely  
752 to be affected by imputation uncertainty for birds; for mammals and reptiles, the main results could even be  
753 conservative (Figure S3.22, S3.23). Indeed, although confidence intervals around the estimates were large,  
754 I typically observed larger decreases in functional diversity when using the complete data subset, including  
755 an 86% decline in FRic for mammals in intensely used tropical agricultural areas. The results were also  
756 unaffected by variation across replicate sets of imputed trait values (Figure S3.24).

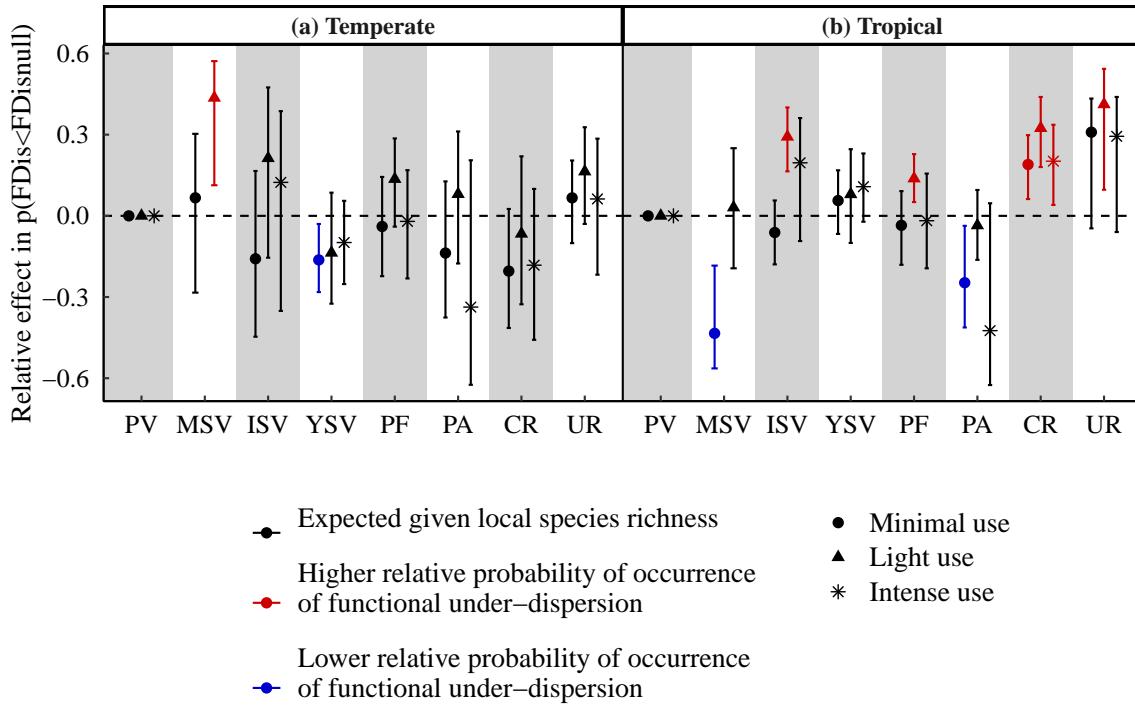
### 757 **3.3.2 Changes in the probability of occurrence of functional under-dispersion**

758 Land use, land-use intensity and region significantly affected the probability of occurrence of functional  
759 under-dispersion across vertebrates. Functional under-dispersion was more likely to occur in tropical crop-  
760 land of all land-use intensities (Figure 3.4(b)), as well as in some of the lightly-used land uses (notably urban  
761 and plantation forest). Contrary to my expectations, and with the exception of tropical cropland, functional  
762 under-dispersion was not more likely to occur in intensely-used land uses. For minimally-used sites, changes  
763 in FDis were mostly consistent with that expected given changes in species richness.

### 764 **3.3.3 Functional loss and gain**

765 Across and within vertebrate classes, I detected high levels of functional loss, exceeding the natural turnover  
766 between primary-vegetation sites, both in temperate and tropical regions. Across vertebrates (Figure 3.5(a)),  
767 functional loss was notably high in temperate pastures (+27% above reference for minimal use; +73% for  
768 intense use), temperate urban sites (+27% for light use; +50% for intense use; effects for tropical urban sites  
769 could not be estimated), temperate and tropical cropland (+44% and +56% respectively for light use; effects  
770 for intense use could not be estimated). Important levels of functional loss were also observed in tropical  
771 plantation forest of light use intensity (+51%; effects for the intense use could not be estimated). High levels  
772 of functional loss were also observed within each class (Figure 3.6(a)) (although not all effects could be  
773 estimated because of limited sample sizes, Table S3.5). The highest losses were observed in agricultural  
774 areas for amphibians and reptiles, with important losses also observed in temperate urban areas for both  
775 birds and amphibians (+35% for minimal use; effects for tropical urban areas could not be estimated).

776 Across vertebrates (Figure 3.5(b)), average functional gain (average proportion of novel trait space in  
777 the disturbed assemblage) was moderate and on average did not exceed 20% in any disturbed land uses.  
778 Patterns of functional gain were similar in both regions. The highest functional gains were observed for  
779 minimally-used urban sites and plantation forest (range: +16% to +20%). On the other hand, important  
780 levels of functional gain were observed in some classes (Figure 3.6(b)), with the highest functional gain for



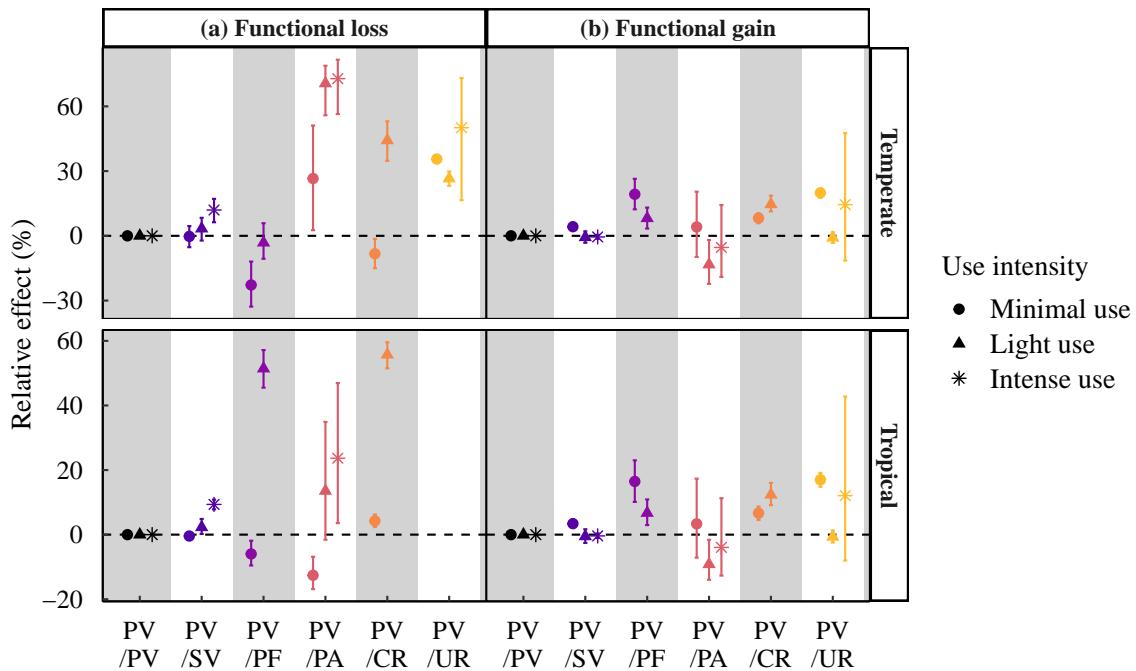
**Figure 3.4: Effects of land use, land-use intensity and region on the probability of occurrence of functional under-dispersion.** Error bars represent 95% confidence intervals. PV: primary vegetation; MSV, mature secondary vegetation; ISV, intermediate secondary vegetation; YSV, young secondary vegetation; PF, plantation forest; PA, pasture; CR, cropland; UR, urban. Effects are rescaled and represent the average difference in the probability of occurrence of functional under-dispersion between the reference (PV, probability of functional under-dispersion set at 0 within each land-use intensity) and the disturbed land uses. *Figure reproduced from Etard et al. (2021).*

781 mammals (+80% in intensely used urban sites).

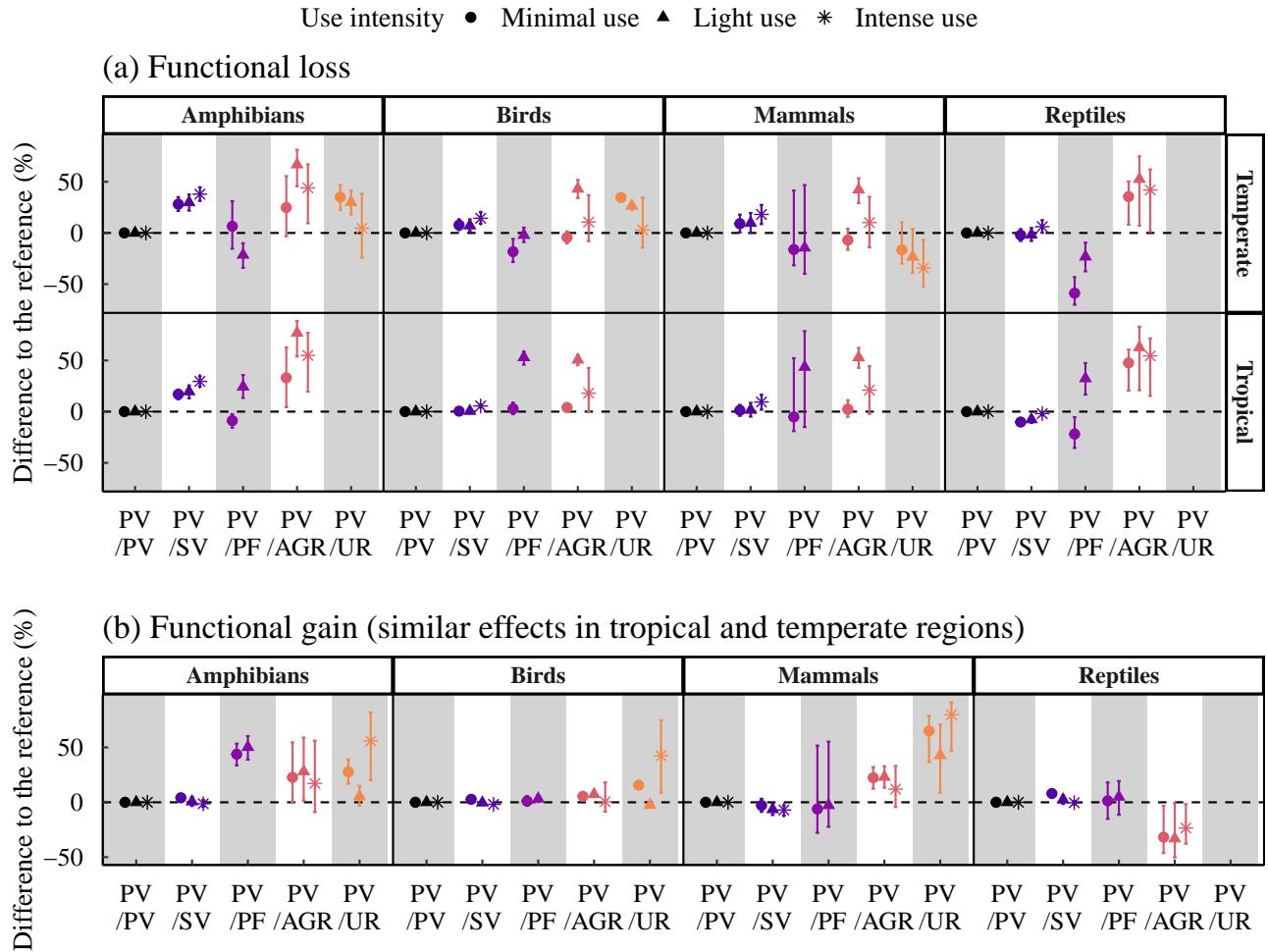
782 Diagnostic plots (qq-plots and residual distributions) for the models are shown in Figures S3.9–S3.17.

783 Overall, the model residuals were appropriately distributed (but with some leptokurtic residual distributions,

784 to which mixed-effect models are generally robust (Schielzeth et al., 2020)).



**Figure 3.5: Effects of land use and land-use intensity on (a) functional loss and (b) functional gain across pairs of sites in temperate areas and tropical areas.** PV: primary vegetation; SV, secondary vegetation; PF, plantation forest; PA, pasture; CR, cropland; UR, urban. Error bars represent 95% confidence intervals. Effects are rescaled and represent the average difference in functional loss and gain between the reference pairs (PV/PV, with loss and gain set at 0% within each land-use intensity), and pairs of primary vegetation with disturbed land uses. Negative effects mean that, on average, levels of functional loss (or gain) were lower than functional loss (or gain) observed between pairs of primary-vegetation sites, and not that there were absolute ‘negative losses’ or ‘negative gains’. *Figure reproduced from Etard et al. (2021).*



**Figure 3.6: Effects of land use, land-use intensity, region and taxonomic class on functional loss and functional gain across pairs of sites.** PV: primary vegetation; SV, secondary vegetation; PF, plantation forest; AGR, agricultural land uses (pasture and cropland); UR, urban. Error bars represent 95% confidence intervals. Effects are rescaled and represent the average difference in functional loss and gain between the reference pairs (PV/PV, with loss and gain set at 0% within each land-use intensity), and pairs of primary vegetation with disturbed land uses. Negative effects mean that, on average, levels of functional loss (or gain) were lower than functional loss (or gain) observed between pairs of primary-vegetation sites, and not that there were absolute ‘negative losses’ or ‘negative gains’. *Figure reproduced from Etard et al. (2021).*

785 **3.4 Discussion**

786 Here, I showed that the functional diversity of vertebrate assemblages is negatively impacted in human land  
787 uses, particularly in the most intensely used land types. The results of this Chapter extend previous studies  
788 that have been more taxonomically or geographically restricted (Flynn et al., 2009; Matuoka et al., 2020).  
789 Matuoka et al. (2020) found that the functional diversity of tropical bird assemblages was negatively affected  
790 by human disturbance, a pattern that did not appear in temperate assemblages. Yet, I found that functional  
791 diversity was negatively affected in both tropical and temperate, with important functional losses in all four  
792 vertebrate classes.

793 Using multiple metrics allowed me to explore different facets of functional diversity. For instance, func-  
794 tional gain could locally offset functional loss in some disturbed land uses. This could indicate that despite  
795 no apparent negative effect on FRic, some disturbed land uses (e.g. lightly-used temperate cropland) could  
796 experience important functional loss, and highlights the importance of using a variety of indicators. This  
797 mechanism could be at play in mammalian assemblages, for which important levels of functional gain were  
798 observed in agricultural and urban sites. Further, functional gain in disturbed land uses could indicate that  
799 disturbances facilitate the introduction of functionally novel species, falling into previously unoccupied parts  
800 of the trait space. This may be because non-native species are more likely to become established in disturbed  
801 assemblages. Previous work has shown that land-use disturbance facilitates biological invasions in island  
802 ecosystems (Jesse et al., 2018; Sanchez-Ortiz et al., 2019), but to my knowledge, this has not been tested  
803 specifically across continental areas for invasive vertebrates (but see Pyšek et al. (2010)). It is also pos-  
804 sible that disturbed areas harbour synanthropic species that do not occur in primary vegetation, leading to  
805 substantial functional gain.

806 Overall, the negative effects of land use on functional richness tended to be more pronounced in the  
807 tropics. This is congruent with past studies that have found tropical biodiversity to be disproportionately  
808 sensitive to human pressures (Martins and Pereira, 2017; Newbold et al., 2020b). There are a number of  
809 potential explanations for this. First, it could be that a long history of intense land-use disturbance at large  
810 scales in many temperate regions (e.g. Western Europe; Stephens et al. (2019)) means that biodiversity is  
811 now less sensitive to new disturbances, because the most sensitive species have been filtered out (Balmford,  
812 1996; Krauss et al., 2010; Le Provost et al., 2020; Munteanu et al., 2020). Species unable to cope with such  
813 disturbances may have gone extinct in the past, while the remaining species would be more disturbance-  
814 tolerant (Betts et al., 2019). Tropical regions, historically less disturbed at large scales, would then contain  
815 a higher proportion of disturbance-sensitive species than temperate regions. Consequently, the functional

richness in undisturbed tropical sites could be less resilient to new disturbances. This also highlights that time since land-use conversion could have important impacts on local functional diversity. Although I did not consider the effects of time since land-use conversion in this work (notably because PREDICTS contained data only for about 22% of the sites), I expect that time since land-use conversion may affect assemblage composition, and thus, functional diversity, with potentially land-use-specific relationships between time since conversion and functional diversity (e.g., a positive relationship for recovering secondary vegetation or a negative relationship for urban areas; but I did not detect such effects when using the data subset for which I have information on time since land-use conversion [see Appendix 3, section S3.8]).

Second, it could be that tropical species are intrinsically more sensitive to disturbances than temperate species because of their evolutionary history. Natural climatic variability experienced by species as well as species history of exposure to disturbances have been proposed to influence sensitivity to disturbance. For instance, tropical species are, on average, nearer to their climatic limits than temperate species (Deutsch et al., 2008; Sunday et al., 2014). Tropical species could therefore experience more deleterious effects from interacting drivers of change, with land-use change bringing about novel climatic conditions pushing them beyond their tolerance limits (Frishkoff et al., 2016; Williams and Newbold, 2020).

In addition to filtering out sensitive species, land-use change is also expected to modify interactions among species, thereby influencing species persistence (Tylianakis et al., 2008; Valiente-Banuet et al., 2015). Although I detected a signal of functional under-dispersion (particularly in tropical cropland), which indicates that assemblages may be locally structured by environment filtering (Bregman et al., 2015), it is likely that several assembly rules underpin assemblage composition (Fournier et al., 2016). For instance, land-use changes could enhance competition among species, promoting over-dispersion by removing species that share similar resources. Such opposite signatures of environmental filtering and enhanced competition on functional dispersion could explain why I did not detect stronger effects of land use on functional under-dispersion occurrence.

Studies looking at impacts of global land use on functional diversity computed with species from all four terrestrial vertebrate classes remain rare. Lack of availability of standardised trait data across terrestrial vertebrates may have hindered such studies from being conducted in the past. To overcome this problem, I based the analyses on a large-scale collation of trait data (Chapter 2; Etard et al. (2020)), and I imputed missing trait values to obtain complete trait datasets in each class. I used random forest algorithms, currently thought to be one of the most robust technique for missing value imputations in trait datasets (Debastiani et al., 2021; Johnson et al., 2021; Penone et al., 2014). Replicating the analyses on complete trait data subsets showed that imputation uncertainty did not affect the main conclusions of this work and that the

negative effects of human land uses were in some cases even stronger when using the complete data subsets. Furthermore, the results were highly consistent across imputed datasets and so insensitive to variation across imputed values. Although missing value imputation can offer a robust filling of missing entries, this study highlights the existing taxonomic biases both in trait data availability and in PREDICTS studies, and thus stresses the need to pursue data compilation efforts, particularly for the least-sampled classes (reptiles and amphibians).

Another implication of trait data availability for vertebrates is that the choice of traits was constrained. Mouillot et al. (2021) showed that functional diversity metrics are sensitive to trait omission and that the sensitivity to trait omission decreases with increasing levels of correlation among traits. Here, I chose seven traits that were available across all classes at least for a subset of the species and that have been implicated in shaping species responses to environmental change. A notable omission was any metric of dispersal ability, which is likely to influence species' ability to respond to land-use change but is difficult to obtain for most species. In fact, past studies have shown that dispersal abilities can be predicted from ecological correlates, such as body mass, diet or geographical range size (Schloss et al., 2012; Sutherland et al., 2000). Since the results were robust to the omission of geographical range size, I am confident that the omission of dispersal abilities also does not affect the conclusions of this work.

Functional diversity metrics are often used as a proxy for ecosystem functioning because of the conceptual and mechanistic link between functional 'effect' traits and ecosystem processes (Lavorel and Garnier, 2002; Violle et al., 2007). In many studies focused on vertebrates, however, functional diversity metrics do not correlate with a given ecosystem function (Hatfield et al., 2018). Here, I did not explicitly target given ecosystem functions, but I argue that evidence of functional loss of vertebrate assemblages indicates that processes sustained by vertebrates are put at risk by land-use change. My results further show that some disturbed land uses are more likely to experience functional under-dispersion, particularly tropical cropland and tropical urban areas, which again indicates a potential imperilment of ecological processes. Indeed, in such cases, decreases in functional dispersion exceed changes expected from the chance removal of species; such non-random modifications indicate that certain areas of the functional trait space are more sensitive to land-use disturbance. Future work could investigate the impacts of land-use change on particular ecosystem functions. The integration of trophic information (beyond the trophic levels I used here) to the species-trait dataset could be an interesting step in that direction, as dietary traits relate to resource use and are, as such, probably the most straightforward traits to link with ecosystem functions. Furthermore, my results suggest that the functional loss experienced within a class is unlikely to be compensated for by the persistence of functionally similar species in other classes. Indeed, I detected negative effects of human land use on func-

880 tional richness in at least three out of four vertebrate classes (amphibians, birds, and reptiles), in accordance  
881 with past studies focusing on each of these groups (Gallmetzer and Schulze, 2015; Marcacci et al., 2021;  
882 Riemann et al., 2017; Sol et al., 2020). Although overall mammalian functional richness was less affected,  
883 high levels of functional gain suggest that the functional composition of mammalian assemblages is heavily  
884 modified in disturbed land uses.

885 To conclude, the results of this third chapter highlight the negative impacts of human land uses on mul-  
886 tiple dimensions of functional diversity, within and across terrestrial vertebrate classes, at a global scale.  
887 In many disturbed sites, decreases in functional diversity exceed changes expected from species loss alone,  
888 showing that human activities non-randomly reshape ecological assemblages. By intensifying functional  
889 loss and promoting functional under-dispersion, land-use change could have deleterious effects on ecosys-  
890 tem functioning, highlighting the necessity of putting into place effective conservation measures in the face  
891 of anthropogenic change.

892 **4 | Geographical range area, habitat breadth**

893 **and specialisation on natural habitats**

894 **predict land-use responses and climate-**

895 **change sensitivity more consistently than**

896 **life-history and dietary traits in terres-**

897 **trial vertebrates**

898 **Keywords**

899 **Abstract**

900 **4.1 Introduction**

901 **4.2 Methods**

902 **4.3 Results**

903 **4.4 Discussion**

904 **5 | Energetic constraints and trophic group**

905 **explain species persistence in disturbed**

906 **land uses**

907 **Keywords**

908 Land use; land-use intensity; metabolic rates; energetic constraints; energetic requirements; terrestrial verte-  
909 brates; trophic group; occurrence.

910 **Abstract**

911 Land-use change is the primary driver of global biodiversity loss. In terrestrial vertebrates, previous work  
912 has shown that sensitivity to land-use change depends on species traits, but the extent to which energetic  
913 constraints explain species responses to disturbed land uses remains largely unexplored. Here, I investigate  
914 relationships between the energetic requirements of terrestrial vertebrates (estimated from resting metabolic  
915 rates) and land-use change, at two levels of organisation. First, at the assemblage level I hypothesize that  
916 total energetic requirements in disturbed land uses are lower than in undisturbed land uses, assuming that  
917 there is less energy available in these areas overall. Second, after controlling for the effects of body mass and  
918 taxonomy on metabolic rates, I predict that species with relatively lower energetic expenditure are favoured  
919 over species with relatively higher energetic expenditure in disturbed land uses, as resource efficiency will be  
920 beneficial in these resource-poor environments. Because trophic group influences species ability to assimili-  
921 late various types of food, I investigate whether my predictions are consistent among trophic groups (here,  
922 omnivores, carnivores or herbivores). The results challenged both hypotheses. I found that total assemblage-  
923 level energetic requirements did not systematically decrease in disturbed land uses. For instance, I detected

924 significant increases for urban areas in all trophic groups, highlighting that disturbed areas may not be as  
925 energy-poor as I initially assumed. Second, I found a positive effect of metabolic rates (after controlling  
926 for body mass and taxonomy) on species probability of occurrence across all trophic groups for at least one  
927 of the most disturbed land uses I considered (pasture, cropland and urban). Species for which there are  
928 exploitable resources in disturbed environments may benefit from having larger energetic expenditure: they  
929 may display a set of characteristics rendering them more able to cope with disturbances and more able to  
930 acquire available resources, such as higher activity levels or bigger brain sizes. The findings of this Chapter  
931 highlight that land-use change has substantial impacts on vertebrate community metabolism.

## 932 **5.1 Introduction**

933 Land-use change is currently the strongest driver of global biodiversity declines (Maxwell et al., 2016;  
934 Newbold et al., 2015), with major and long-lasting impacts on the structure and functioning of ecological  
935 communities (Bregman et al., 2016; Fukasawa and Akasaka, 2019; Magioli et al., 2021; Marcacci et al.,  
936 2021). With land-use change likely to continue to intensify (Stehfest et al., 2019), it is vital to put into  
937 place conservation and mitigation measures to minimise future losses of biodiversity and negative impacts  
938 on ecosystem functioning. To this end, pressing questions remain as to what renders species able or unable  
939 to cope with human disturbance, and how losses of sensitive species might influence ecosystem functioning  
940 (Dirzo et al., 2014; Young et al., 2016).

941 Land-use change acts as an environmental filter affecting species persistence (Edwards et al., 2021;  
942 Evans et al., 2018). Past studies have shown that sensitivity to land-use change is distributed unevenly across  
943 the tree of life (Nowakowski et al., 2018), and across behavioural (Lowry et al., 2013; Samia et al., 2015)  
944 and ecological strategies (De Palma et al., 2015; Møller, 2009; Newbold et al., 2013). For instance, long-  
945 lived and large forest specialist birds respond more negatively to land-use change than generalist species  
946 (Newbold et al., 2013). In addition, human land uses impose energetic constraints on species, by modifying  
947 the amount and type of available resources (Inger et al., 2016; Zuñiga-Palacios et al., 2021). Consequently,  
948 land-use change brings about modifications in dietary and foraging strategies (Møller, 2009; Ramesh et al.,  
949 2017; Sévèque et al., 2020), which can have cascading effects on local trophic structure (e.g., mesopredator  
950 release; Crooks and Soule (2010)). The impacts of land-use change on biodiversity also depend on land-  
951 use intensity (Davison et al., 2021; Newbold et al., 2015), that is, the magnitude of human disturbance in  
952 a given land-use type, which can vary importantly according to local practices and management (Dullinger  
953 et al., 2021). In particular, land-use intensity affects the types and availability of resources (by homogenising

resources for example), which in turn may affect biodiversity patterns (Weiner et al., 2011).

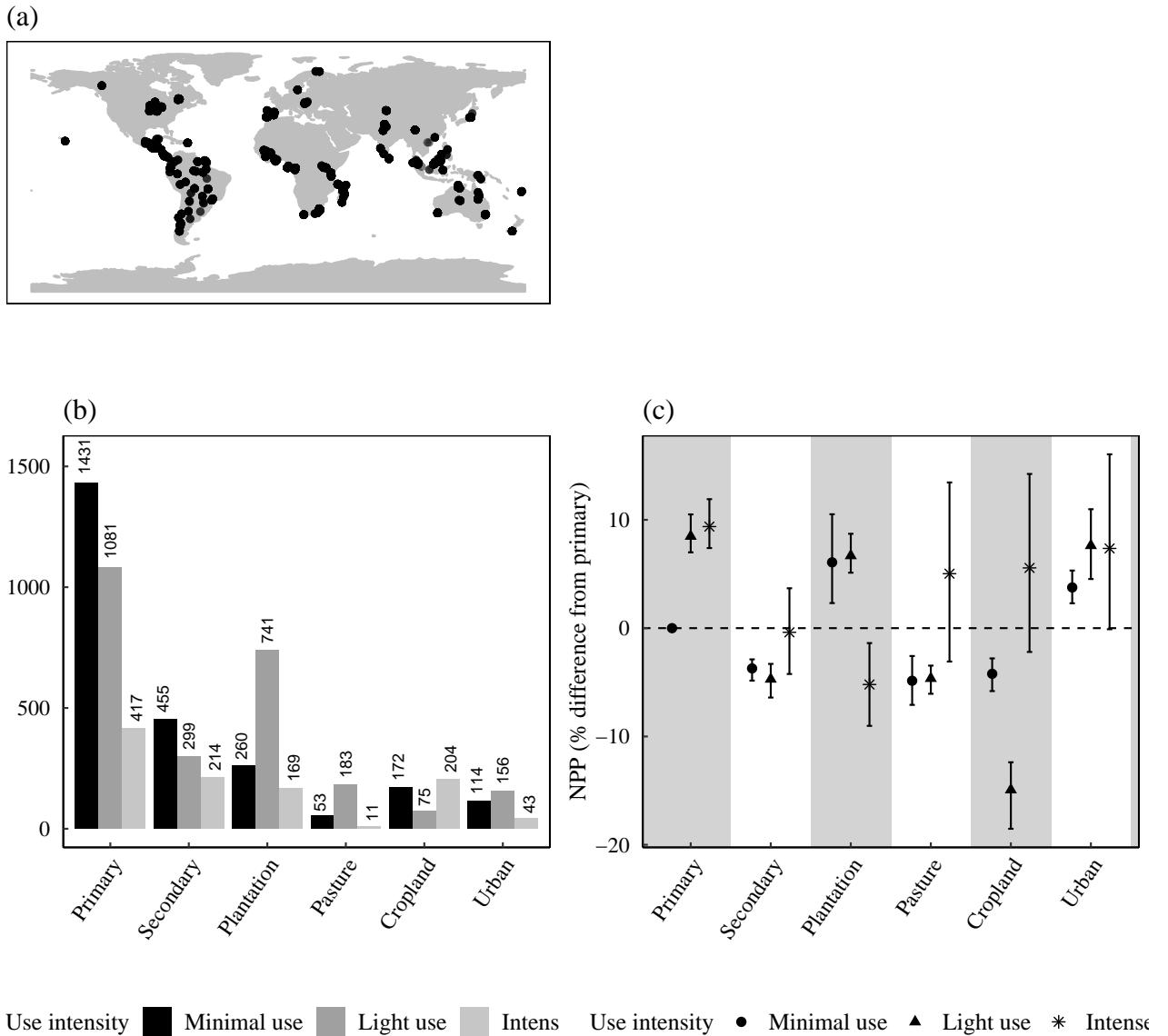
Energetic requirements are likely to be important predictors of species' ability to cope with a disturbed environment where resource availability and resource types are strongly impacted (Auer et al., 2020). Past research has shown that metabolic rates (the amount of energy required for organismal maintenance and for other processes such as growth and reproduction) are heritable and subject to selection, for instance with species in energy-rich environments having evolved faster metabolic rates than species in less energy-rich environments (Mueller and Diamond, 2001). However, metabolic rates have received less attention than other life-history traits in studies aiming to understand different species' responses to land-use change. Body mass, known to explain much of the variation in metabolic rates (Bushuev et al., 2018; Hudson et al., 2013; White and Seymour, 2003), has been considered in past studies examining effects of land-use disturbance on assemblage composition (Hevia et al., 2017). Since metabolic rates correlate positively with body mass, the effects of land-use change on local energetic requirements have been indirectly assessed through investigations of changes in body mass composition along land-use gradients (Newbold et al., 2020a; Tinoco et al., 2018). However, to the best of our knowledge, no study has yet directly reported the effects of land-use change on assemblage-level energetic requirements.

Thus, the first objective of this Chapter is to investigate how land-use change impacts assemblage-level energetic requirements in terrestrial vertebrates. To this end, I use metabolic rates as a proxy for maintenance energetic requirements at the species level. I predict that, at the assemblage level, total energetic requirements in disturbed land uses are lower than in undisturbed land uses. This prediction relies on the assumption that there is less energy available overall in disturbed land uses than in undisturbed land uses (Figure 5.1c), with disturbed land uses being characterised by human appropriation of net primary productivity (Krausmann et al., 2013). Because species' ability to exploit particular resources and meet energetic demands also depends on food types and diet (McNab, 1986; Mendoza and Araújo, 2019), I test whether this prediction holds true within trophic groups (i.e., within herbivores, omnivores, and carnivores). Here, I consider energetic requirements at the assemblage level, not controlling for species body mass, to obtain estimates of total energy expenditure in different land-use types (Figure 5.2a). Thus, any changes in assemblage-level energetic requirements could be ultimately driven by shifts in the size-spectrum of local assemblages.

Body mass explains much of the variation in metabolic rates, but species with similar body masses still display important variation in metabolic rates (Mueller and Diamond, 2001). Some of this variation can be explained by taxonomic position (which also accounts for differences in thermoregulatory strategy between endotherms and ectotherms). However, there remains substantial interspecific variation in metabolic rates even after accounting for body mass and taxonomic affiliation (Mueller and Diamond, 2001; White

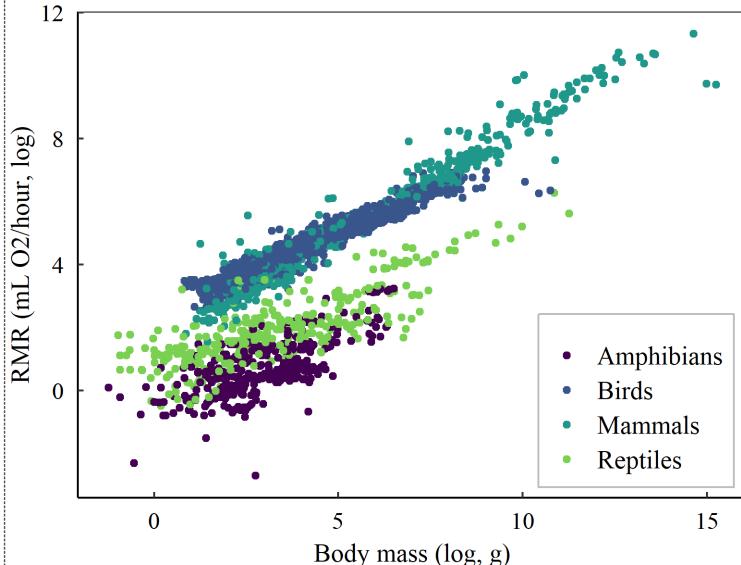
and Kearney, 2013). Food availability and food type might explain this residual variation in metabolic rates among species (Mueller and Diamond, 2001). Species that evolved in areas with abundant resources assimilate and burn calories at high rates, regardless of their body mass – while species that evolved with limiting resources could process resources at lower rates, irrespectively of their body mass (Mueller and Diamond, 2001). Building upon this idea and the ‘allocation’ principle (i.e., the trade-offs in resource allocation; Auer et al. (2020)), species with lower energetic expenditure (than expected from body mass and taxonomy) may be able to cope with food scarcity better than those with larger energy expenditure (than expected from body mass and taxonomy). Indeed, regardless of their body mass, such species should need fewer resources to meet energetic demands (Clarke and Fraser, 2004). Thus, as land-use disturbance modifies the types and abundance of resources, we expect species energetic requirements –after removing the effects of body mass and taxonomic position– to influence responses to land-use change. Yet, whether energetic constraints explain species responses to land-use change in terrestrial vertebrates has not been investigated before (Hevia et al., 2017). Thus, the second objective of this Chapter is to characterise the influence of energetic constraints on species occurrence probability in disturbed land uses, compared to undisturbed land uses. I approximate energetic constraints at the species level from residual variation in metabolic rates, that is, the variation not explained by body mass or taxonomy. I predict that species with negative deviations (lower rates than expected from body mass and taxonomy) are more likely to persist in disturbed land uses than species with positive deviations (higher rates than expected from body mass and taxonomy; Figure 5.2b). Given that trophic group can influence resource acquisition and species ability to assimilate various types of food, I investigate whether this prediction is consistent among trophic groups.

To test both predictions, I use a space-for-time substitution approach, obtaining vertebrate assemblage composition data from the PREDICTS database (Hudson et al., 2014, 2017). I combine this database with trophic-group information (characterising species as either omnivores, herbivores, or carnivores), and with species-level estimates of resting metabolic rates and body mass.



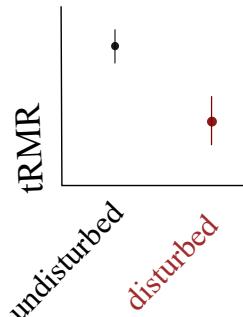
**Figure 5.1:** (a) **Spatial distribution of the sampled sites** from the PREDICTS database for terrestrial vertebrates (6,484 sites); (b) **Number of sites in each land-use and land-use-intensity category**; (c) **Net primary productivity by land use and land-use intensity** (derived from MODIS satellite imagery), with model predictions plotted relative to minimally used primary vegetation (and rescaled with reference to that land-use type). Primary: primary vegetation; secondary: secondary vegetation; plantation: plantation forest.

**(a) Assemblage-level:** there is less energy available to utilize by vertebrate assemblages in disturbed land uses than in undisturbed land uses.



**Prediction 1:** within all trophic groups, total assemblage-level **mass-dependent** metabolic rates decrease in disturbed land uses compared to undisturbed land uses.

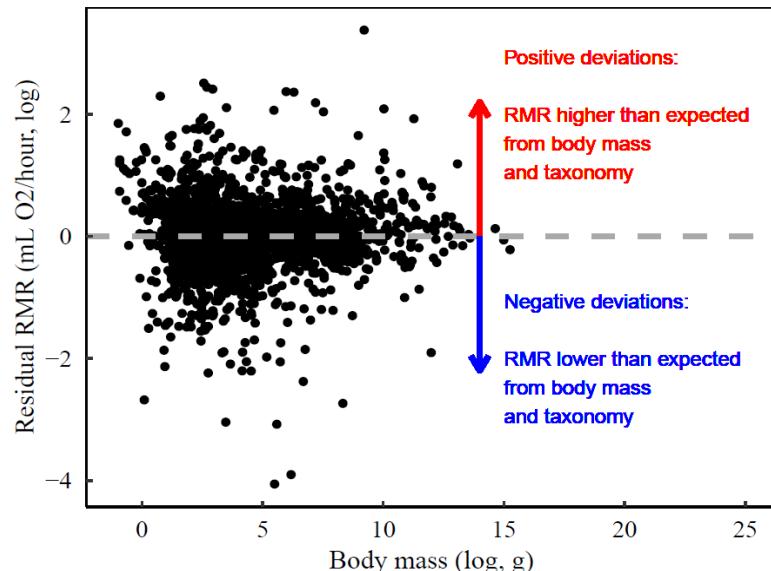
$$\text{Assemblage level:} \\ t\text{RMR} = \sum_i (\text{RMR}_i * a_i)$$



$$\text{Across assemblages - model 1:} \\ \log(t\text{RMR}) \sim \text{LU} + \text{LUI} + \text{TG} + \\ \text{LU:LUI} + \text{LU:TG} + \text{LUI:TG} + \\ \text{LU:LUI:TG} + \text{RE}$$

**(b) Species-level:** species with lower energetic expenditure than expected from body mass and taxonomy (negative deviations) do better in disturbed land uses than species with higher energetic expenditure than expected from body mass and taxonomy (positive deviations).

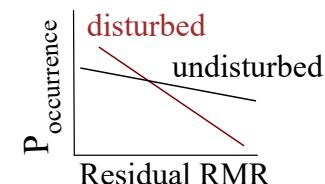
$$\text{residuals of} \\ \log(\text{RMR}) \sim \log(\text{BM}) + \\ (1|\text{Class/Order/Family}))$$



**Prediction 2:** within all trophic groups, species occurrence probability in disturbed land uses is negatively affected by **residual RMR**. The slope of the relationship between occurrence probability and residual RMR is significantly lower than the slope estimate for the undisturbed land uses.

Across species - model 2:

$$\text{P}_{\text{occurrence}} \sim \text{LU} + \text{LUI} + \text{resRMR} + \text{TG} + \\ \text{LU:LUI} + \text{LU:TG} + \text{LU:resRMR} + \\ \text{LUI:TG} + \text{LUI:resRMR} + \text{TG:resRMR} + \\ \text{LU:TG:resRMR} + \text{LUI:TG:resRMR} + \text{RE}$$



**Figure 5.2: Framework for the predictions and models.** I use resting metabolic rates (RMR) as a proxy for the amount of energy processed by vertebrate species. (a) I hypothesize that there is less energy overall that can be utilized by vertebrate species in disturbed land uses compared to undisturbed land uses. As such, I expect decreases in assemblage-level total metabolic rates ( $t\text{RMR}$ ).  $t\text{RMR}$  is calculated as the sum, over all species occurring at a site, of species RMR ( $\text{RMR}_i$ ) weighted by species abundance ( $a_i$ ). (b) After removing the effects of body mass and taxonomic position, I expect species with lower energy expenditure than expected to do better than species with higher energy expenditure than expected. I predict that the slope of the relationship between residual RMR and occurrence probability is more negative in disturbed land uses than in undisturbed land uses. For both (a) and (b), I test whether the predictions are consistent among trophic groups. LU: Land use; LUI: land-use intensity; TG: trophic group; RE: random effects. In model 1, random effects include the identity of the PREDICTS studies from which the assemblages are derived. In model 2, random effects include study, site, and species identity.

1010 **5.2 Methods**

1011 **5.2.1 Vertebrate assemblage composition**

1012 I obtained vertebrate assemblage composition in different land uses from the PREDICTS database (Hudson  
1013 et al., 2014, 2017). The PREDICTS database is a large collection of published studies that measure bio-  
1014 diversity across different land uses and is one of the most comprehensive global databases of its type. In  
1015 each PREDICTS study, species occurrence and often abundance were recorded across different sites. Each  
1016 site was assigned to one of the following land-use types: primary vegetation (natural, undisturbed vegeta-  
1017 tion), secondary vegetation (recovering after complete destruction of primary vegetation), plantation forest  
1018 (woody crops), pasture (areas grazed by livestock), cropland (herbaceous crops) and urban (built-up areas).  
1019 The land-use categories were assigned based on habitat descriptions from the original studies (Hudson et al.,  
1020 2014), sometimes in consultation with the original study authors. Each site was also classified in terms of  
1021 land-use intensity as either minimal, light or intense. The land-use-intensity assignment was also made on  
1022 the basis of the habitat description in the original studies, and depended on criteria specific to each land use  
1023 (such as degree of mechanisation, yield or chemical inputs for cropland; or the amount of green space in  
1024 urban areas; Hudson et al. (2014)).

1025 I subset the PREDICTS database for studies that sampled terrestrial vertebrates, and for which both land  
1026 use and land-use intensity had been characterised. I thus obtained 181 studies for 4,238 species sampled  
1027 across 6,484 sites (Figure 5.1a). Sample sizes varied across land uses and land-use intensities (Figure 5.1b).

1028 **5.2.2 Energy availability by land-use type and land-use intensity**

1029 The predictions of this Chapter rely on the assumption that resource types and abundance are modified in dis-  
1030 turbed environments, with less energy available in disturbed compared to undisturbed land uses overall. To  
1031 test this assumption, I used terrestrial net primary productivity (NPP) across land uses as a proxy for available  
1032 energy. NPP quantifies the amount of atmospheric carbon fixed by plants and accumulated as biomass. NPP  
1033 estimates were derived using imagery from the Moderate Resolution Imaging Spectroradiometer (MODIS)  
1034 on board NASA's Terra satellite. NPP estimates were based on a yearly composite of measures made at 8-day  
1035 intervals, captured at 500-m spatial resolution (Running and Zhao, 2015). I obtained NPP for 4,062 of the  
1036 PREDICTS sites used in the analysis (matching the sites to the NPP data using the sampling year available in  
1037 PREDICTS). I fit a linear mixed-effects model (lme4 package, version 1.1-23, Bates et al. (2015)) explaining  
1038 site-level NPP by land use and land-use intensity, with a random intercept accounting for study identity, to

control for differences in experimental design across studies. Model predictions showed that NPP decreased significantly in several land uses (e.g., pasture and cropland) compared with the primary vegetation reference level, although the strength and in some cases direction of the difference varied among land-use and land-use intensity combinations (e.g., increases in urban land uses; Figure 5.1c).

### 5.2.3 Resting Metabolic Rates (RMR) & imputations of missing RMR values

As a proxy for species-level energetic expenditure, I used estimates of the minimum amount of energy required for organismal maintenance, i.e., basal metabolic rates (BMR) for endotherms, and resting metabolic rates (RMR) for ectotherms. From the literature, I obtained estimates of BMR for 719 species of birds and 685 mammals, and estimates of RMR for 126 amphibians and 173 reptiles (Appendix 4, Table S5.1). For endotherms, BMR are measured when species are in their thermoneutral zone, that is, when there is little to no energy expenditure allocated to thermoregulation. Thus, BMR estimates were derived from lab studies that mostly measured oxygen consumption of the organisms at rest under controlled conditions and in the thermoneutral zone of the species. For an ectotherm, there is no ‘basal’ metabolic rate, as body temperature mainly depends on environmental temperature. Their metabolic rates follow a hump-shaped relationship with environmental temperature, highest at an optimal temperature which corresponds to a performance peak.

To be able to compare endotherms’ BMR with ectotherms’ RMR, Stark et al. (2020) used the metabolic rates that correspond to a performance peak for both groups (i.e., BMR in the thermoneutral zone for endotherms, and metabolic rates at optimal temperature for ectotherms). Thus, I used the data compiled in Stark et al. (2020) for ectotherms, and from the sources specified in Table S5.1 for endotherms. The units for BMR and RMR were standardized to mL of dioxygen consumed per hour (mLO<sub>2</sub>/h). As in Stark et al. (2020), I henceforth refer to both basal and resting metabolic rates as RMR.

For the species occurring in PREDICTS, initial data coverage for RMR was poor (Table S5.1), necessitating imputation of missing values. To do so, I first measured the phylogenetic signal in BMR and RMR ( $\log_e$ -transformed), using Pagel’s  $\lambda$  (Pagel, 1999), to assess whether metabolic rates were sufficiently phylogenetically conserved to be estimated from species phylogenetic position. I obtained class-specific phylogenetic trees from Jetz et al. (2012) for birds, from Faurby et al. (2018, 2020) for mammals, from (Tonini et al., 2016) for reptiles (squamates), and from Jetz and Pyron (2018) for amphibians (all downloaded in April 2020). For each class, I randomly sampled 100 trees. To account for phylogenetic uncertainty, I calculated Pagel’s  $\lambda$  for each sampled tree and reported the median value, as well as the 2.5th and 97.5th percentiles (Table S5.1).

In addition to being highly phylogenetically conserved (Table S5.1), RMR correlate strongly with body

mass (Figure 5.2a). Thus, I imputed missing values using body-mass information (see next section), phylogenetic relationships and taxonomic orders as predictors (Penone et al., 2014). For each class, I used a consensus phylogenetic tree from which I summarised phylogenetic relationships in the form of five phylogenetic eigenvectors. Including more eigenvectors had little impact on the imputed values (results not shown). Consensus trees were obtained with the TreeAnnotator programme of the BEAST software (Bouckaert et al., 2014). Missing RMR values were imputed using random forests algorithms implemented in R using the ‘missForest’ package (version 1.4; Stekhoven (2016) and Stekhoven and Bühlmann (2012)).

#### 5.2.4 Trophic group and body mass information

I used body mass and trophic group information for terrestrial vertebrates compiled in Chapter 2. Body mass was compiled as a single measure at the species level, meaning I was unable to consider intraspecific variation. Trophic group described species as either carnivores, omnivores, or herbivores. Because there were gaps in the availability of the data, more so for trophic group than for body mass (see Chapter 2), I imputed the missing trait values (independently of RMR imputations), then used both imputed and empirical body mass values for imputations of missing RMR values. To impute missing body mass and trophic groups, I used random forests algorithms (again, using the missForest R package), including as additional predictors phylogenetic information, added in the form of 10 phylogenetic eigenvectors (Diniz-Filho et al., 2012) following Penone et al. (2014), and also taxonomic order. I considered a wider set of life-history traits in the missing values imputations: lifespan, litter/clutch size, habitat breadth and use of artificial habitats (compiled in Chapter 2). Phylogenetic eigenvectors were extracted from the class-specific phylogenies using the PVR package (Santos, 2018).

#### 5.2.5 Effects of land use, land-use intensity and trophic group on assemblage-level total RMR (prediction 1; Figure 5.2a)

Assemblage-level total RMR (tRMR) was obtained by summing abundance-weighted RMR for the species occurring in each site; abundance data were available for 125 of the 181 PREDICTS studies I considered (sampling 3,487 species across 4,644 sites). I fitted a linear mixed-effects model to explain  $\log_e$ -tRMR as a function of land use, land-use intensity and trophic level, with a random intercept accounting for study identity to control for differences in experimental design across studies. I started with a model allowing all two-way interactions among the predictors. I then tested whether adding the three-way interaction among land use, land-use intensity and trophic level improved the fit of the model, using a likelihood-ratio test. The model that included the three-way interaction was retained ( $P \ll 0.01$ ; model 1, Figure 5.2). In addi-

1100 tion, because it is well established that resting metabolic rates are influenced by temperature (Clarke and  
1101 Fraser, 2004), I checked whether including annual mean temperature in the model affected the conclusions.  
1102 Annual mean temperature at each PREDICTS site was estimated from WorldClim version 2.1 (Fick and  
1103 Hijmans, 2017), using a 2.5 arc-minute resolution. Adding annual mean temperature did not improve model  
1104 fit (likelihood-ratio test:  $P=0.113$ ), thus I did not consider its effects any further.

1105 **Model validation.**

1106 To ensure that imputation uncertainty did not affect the conclusions, I refitted model 1 using the subset of  
1107 species ( $n = 426$ ) from PREDICTS for which there were empirical RMR information (i.e., excluding imputed  
1108 RMR values).

1109 **Disentangling the effects of body mass and abundance on tRMR.**

1110 Since RMR correlates strongly with body mass, changes in tRMR are likely to be driven in part by changes  
1111 in the size-spectrum of ecological assemblages. I fitted an additional model to explain changes in species'  
1112 abundance (given presence) by land use, land-use intensity, trophic group, body mass and their interactions,  
1113 to understand the role of shifts in the body mass of species on observed changes in tRMR (see Appendix 4,  
1114 Figure S5.1).

1115 **5.2.6 Effects of land use, land-use intensity, trophic group and residual RMR on species  
1116 occurrence probability (prediction 2; Figure 5.2b)**

1117 To control for the effects of body mass and taxonomy on RMR, I used the residual variation in RMR after  
1118 accounting for these variables, from a linear mixed-effects model fitting  $\log_e$ -RMR as a function of  $\log_e$ -  
1119 body mass with nested random taxonomic effects (1|Class/Order/Family; Figure 5.2). Hence, I used a metric  
1120 that describes how much more energy (positive deviations) or less energy (negative deviations) than expected  
1121 from body mass and taxonomic position a species spends for organismal maintenance. Similar approaches  
1122 have been used in previous papers (Furness and Speakman, 2008; Naya et al., 2013). As detailed earlier, I  
1123 expect species with lower residual RMR to do better in disturbed land uses than species with higher residual  
1124 RMR (prediction 2; Figure 5.2b) because, given any body mass, investing less energy in maintenance could  
1125 contribute to persistence in a context of resource scarcity.

1126 To test the second prediction, I fitted a binomial mixed-effects model explaining species occurrence  
1127 with land use, land-use intensity, trophic group and residual RMR. I started with a complete model that  
1128 included all two-way interactions among the main effects. Because I wanted to test whether the second

prediction was valid for each trophic group, I needed to account for potential differences in the slope of the relationships between occurrence probability and residual RMR among trophic groups. Thus, I performed a forward stepwise selection procedure to test whether adding three-way interactions among (1) land use, trophic group and residual RMR and (2) among land-use intensity, trophic group and residual RMR improved model fit, using likelihood-ratio tests. The final model included both three-way interactions (Figure 5.2b; model 2). I fitted random effects that accounted for species identity, as well as for study and site identity within PREDICTS.

#### Model validation.

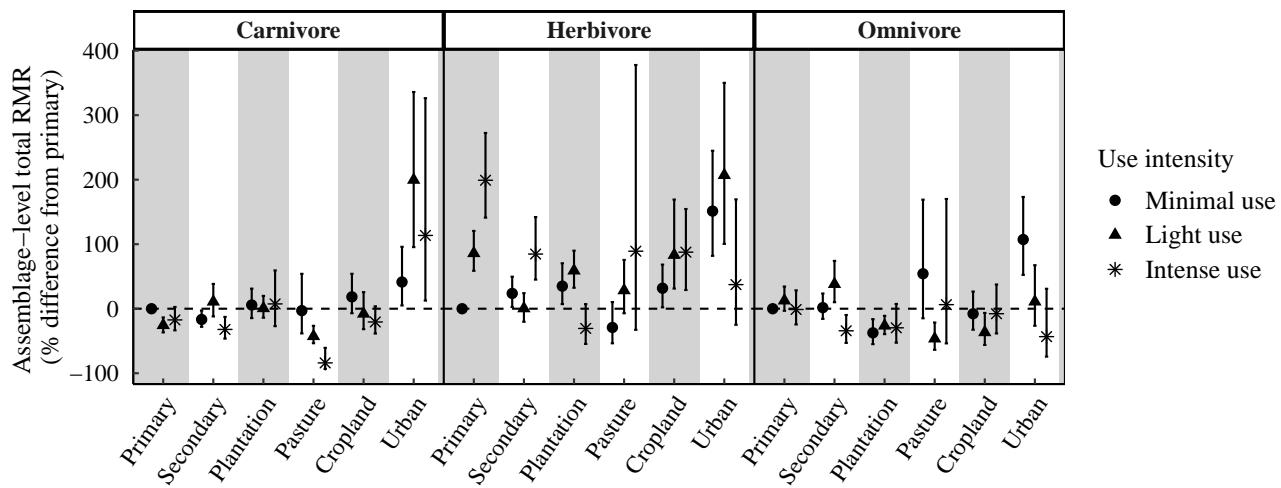
I checked the phylogenetic signal in the model residuals using Pagel's  $\lambda$  (Pagel, 1999). Non-significant phylogenetic signal in the residuals would indicate that fitting species identity in the model's random effects was sufficient to account for residual phylogenetic variation in RMR. Further, to assess the potential effects of imputation uncertainty on the results, I again fitted model 2 on the data subset for the 489 species with collected empirical RMR values, across 5,948 sites in 151 studies (i.e., excluding imputed values).

## 5.3 Results

### 5.3.1 Effects of land use, land-use intensity and trophic group on assemblage-level total RMR

Land use, land-use intensity, trophic group and their interactions had significant effects on assemblage-level total RMR (Figure 5.3). Overall, and contrary to our expectations, assemblage-level total RMR did not show systematic decreases in disturbed land uses. In fact, urban land uses were associated with strong significant increases in tRMR in all trophic groups (e.g., a 200% average increase in tRMR in lightly-used urban areas for carnivores, compared with primary vegetation levels; +207% on average in lightly-used urban areas for herbivores; +107% for minimally-used urban areas for omnivores). In other land uses, responses depended on trophic group and land-use intensity. Whilst for herbivores, disturbed land uses were typically associated with increases in tRMR, we detected decreases in tRMR for omnivores and carnivores in several land uses, most notably in intensely-used pasture for carnivores (-84%). Such effects could reflect changes in the size-spectrum of local assemblages (see Appendix 4, Figure S5.2). For instance, in minimally-used urban areas, larger herbivores tended to occur at higher abundances compared to primary vegetation level; and in intensely-used pastures, carnivores tended to occur at lower abundances overall (Figure S5.2).

1157 The model residuals were appropriately distributed (see diagnostic plots in Appendix 4, Figure S5.3).  
 1158 Investigating the sensitivity of our results to imputation uncertainty showed that our results and conclusions  
 1159 were robust to the removal of all imputed estimates of RMR (the correlation coefficient was 0.72 between the  
 1160 two sets of model coefficients; Figure S5.4). Comparing model predictions showed that effects were mostly  
 1161 congruent, although there were some differences (Figure S5.5). In particular, for herbivores, effect sizes  
 1162 tended to be bigger for the model fitted on empirical data compared with the model that included imputed  
 1163 data. Thus, our main results appear to be conservative if anything. The model fitted on empirical data had  
 1164 larger standard errors, likely due to the reduction in sample size.

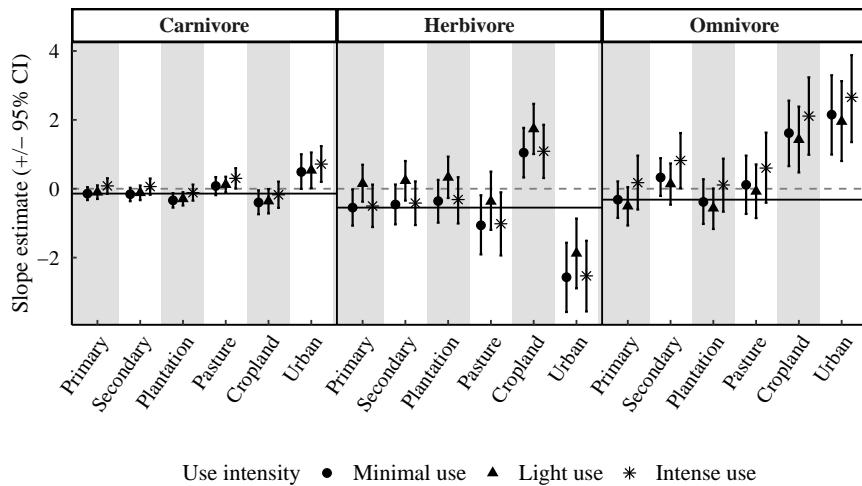


**Figure 5.3: Effects of land use, land-use intensity and trophic group on assemblage-level total RMR.** Model predictions are rescaled with reference to minimally-used primary vegetation, considered to be the undisturbed baseline. Primary: primary vegetation; secondary: secondary vegetation; plantation: plantation forest.

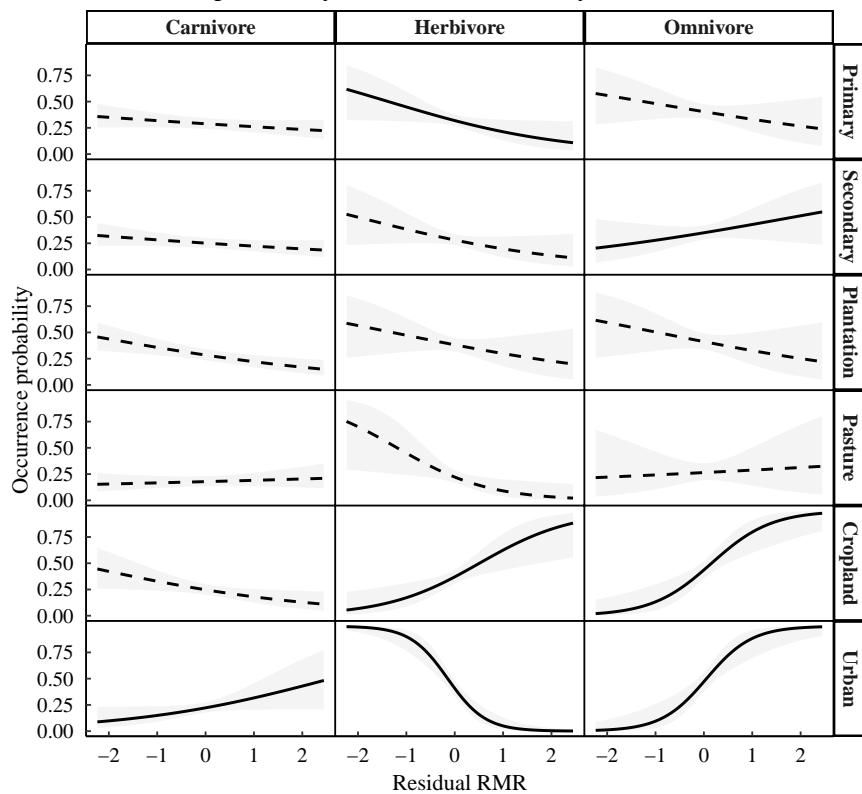
### 1165 5.3.2 Effects of land use, land-use intensity, trophic group and residual RMR on species' 1166 occurrence probability

1167 Species' occurrence probability was significantly affected by land use, land-use intensity, trophic level, resid-  
 1168 ual RMR and their interactions (Figures 5.4, 5.5). Contrary to our expectations, species with higher residual  
 1169 RMR (relative to their body mass and taxonomic position) tended to do better than species with lower resid-  
 1170 ual RMR in a number of disturbed land uses. Overall, land-use type was more important for determining the  
 1171 relationship between occurrence probability and residual RMR than land-use intensity (Figure 5.4a).

(a) Slope of the relationships between occurrence probability (log-odds) and residual RMR



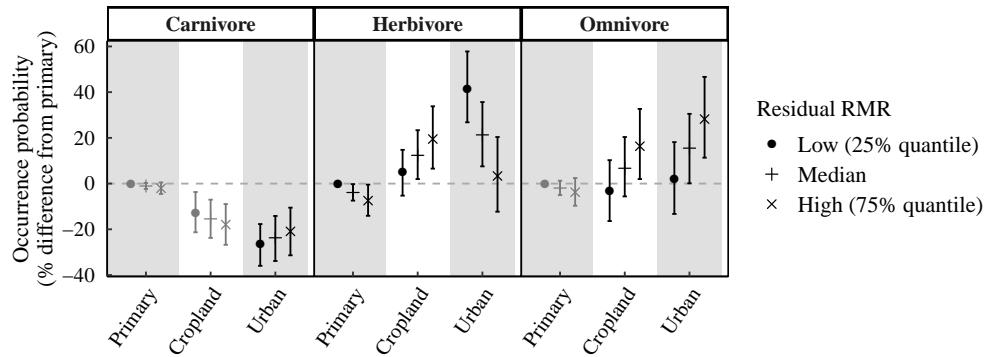
(b) Predicted effects of residual RMR on occurrence probability (minimal use intensity)



**Figure 5.4:** (a) Slope estimates for the relationship between residual RMR and occurrence probability in each land-use type and for the three levels of land-use intensity. The black horizontal line indicates the slope for the reference level (primary vegetation) for minimal land-use intensity. The grey dashed line marks 0. Error bars are 95% confidence intervals. (b) Effect of residual RMR on species probability of occurrence within each trophic level and for each land use-type. I plotted the predictions for minimal land-use intensity only. Solid lines represent significant relationships. Primary: primary vegetation; secondary: secondary vegetation; plantation: plantation forest.

1172 For minimally-used primary vegetation (reference), the model predicted negative effects of residual RMR  
 1173 on species occurrence probability in all trophic levels (but with a significant slope for herbivores only; Figure  
 1174 **5.4a**). However, the directionality of this relationship was reversed in some disturbed land uses in all trophic  
 1175 groups (secondary vegetation, cropland and urban for omnivores; cropland for herbivores; urban for carni-  
 1176 vores), with significant positive slopes, also significantly higher than those observed for primary vegetation  
 1177 (Figure **5.4a**). The only exception was the opposite pattern for urban herbivores (Figure **5.4b**), where residual  
 1178 RMR had a more negative effect on occurrence probability than in minimally-used primary vegetation.

1179 I would like to emphasize that positive effects of residual RMR on occurrence probability in some of the  
 1180 most disturbed land uses (e.g., urban for carnivores) do not mean that there were absolute increases in species  
 1181 occurrence probability in disturbed land uses compared to primary vegetation (and vice-versa). I illustrate  
 1182 this point in Figure **5.5**. For carnivores with a median value for residual RMR, occurrence probability was  
 1183 reduced by an average 24% in urban land uses compared to primary vegetation.



**Figure 5.5: Predicted occurrence probabilities (+/- 95% confidence interval) in primary vegetation (primary) and for the most disturbed land uses where residual RMR was found to have a significant effect on occurrence probability.** For visualisation purposes, I discretised residual RMR in three levels. The predicted probabilities of occurrence were rescaled with reference to primary vegetation for the lowest value of residual RMR (25% quantile). Here, the predictions are plotted for minimal land-use intensity (effects would be similar for light and intense land-use intensities). Black points and error bars are plotted where the relationship between occurrence probability and residual RMR was significant (and dark grey points and error bars represent non-significant trends).

1184 Finally, the model showed some degree of deviation from distributional assumptions (diagnostic plots,  
 1185 Appendix 4, Figure **S5.6**). Nevertheless, the model's coefficients were similar when estimated with a  
 1186 Bayesian framework, suggesting that the estimates were robust (Figure **S5.7**). The phylogenetic signal in  
 1187 the model residuals was weak and non-significant ( $\lambda=0.004$ ,  $P=0.24$ ). Re-fitting the model using the com-  
 1188 plete data subset (i.e., excluding imputed RMR estimates) showed that our conclusions are likely robust to  
 1189 imputation uncertainty (Figure **S5.8**), with congruent results overall, although there were a few differences  
 1190 in the predictions between the two models – notably, for herbivores in urban land uses (Figure **S5.8**).

1191 **5.4 Discussion**

1192 The results of this Chapter provide insights into the relationship between land-use change and energetic  
1193 requirements at two levels of organisation (at the assemblage level and the species level), and contradict  
1194 both my initial predictions. I found that total energetic requirements did not show systematic decreases in  
1195 disturbed land uses. On the contrary, there were strong increases in all trophic groups, most notably in urban  
1196 land uses. Further, in all trophic groups, species with higher energetic expenditure (than expected from body  
1197 mass and taxonomy) tended to be better in some of the most disturbed land uses (that is, cropland and urban)  
1198 than species with lower energetic expenditure (than expected from body mass and taxonomy).

1199 At the assemblage level, the results first highlight the effects of land-use change on vertebrate com-  
1200 munity metabolism. Contrary to my expectations, I found differing effects of land use on total energetic  
1201 requirements among trophic groups, reflecting changes in the size-spectrum of ecological assemblages (i.e.,  
1202 changes in the distribution of abundance along the body mass gradient). On the one hand, decreases in total  
1203 energetic requirements, such as observed for carnivores in intensely used pastures, are likely due to overall  
1204 reductions in local abundance (with carnivorous species in pastures perceived as a threat to livestock; Eeden  
1205 et al., 2018; or because of increased competition for fewer resources). On the other hand, increases in total  
1206 energetic requirements could reflect higher levels of resource intake in some disturbed land-use types. On the  
1207 basis of net primary productivity, I hypothesized that there would be fewer resources available in disturbed  
1208 areas. However, it could be that the carrying capacity of some disturbed land-use types actually increases  
1209 for some trophic groups compared to that of primary vegetation, because of the presence of novel or more  
1210 abundant food sources. For instance, in urban areas, both wildlife feeding and the presence of anthropogenic  
1211 food sources, such as human refusal or pet food, could lead to an increase in resource availability all year-  
1212 round for species with a non-specialised diet that are able to exploit such food sources (Fischer et al., 2012).  
1213 Past research on urban carnivores has shown that some species have adapted to urban environments by ex-  
1214 ploiting anthropogenic food sources, and also benefit from physical protection (with some human structures  
1215 providing shelter) and from reduced occurrence of natural enemies (Bateman and Fleming, 2012). Herbi-  
1216 vores and omnivores could also benefit from increased resources in urban areas, notably where the urban  
1217 matrix includes semi-natural habitats, which is congruent with observed increases in occurrence probability  
1218 in minimally used urban areas for these trophic groups. Further, anthropogenic ‘bottom-up food forcing’ has  
1219 been shown to affect food-web dynamics, with decreases in predation rates, loss of anti-predator abilities  
1220 (Geffroy et al., 2020) and mesopredator release (Fischer et al., 2012). Thus, in urban areas, increases in  
1221 anthropogenic food sources could provide support for a higher abundance of medium-sized species released

1222 from predation, increasing total energetic requirements.

1223 I hypothesized that resource availability would be the main driver behind changes in energetic require-  
1224 ments. However, resource availability likely interacts with other factors to explain the patterns. In particular,  
1225 past work has shown that brain size is an important predictor of species ability to cope with disturbance,  
1226 notably in urban settings (Sayol et al., 2020). Since brain size scales allometrically with body mass, it is  
1227 possible that I observe increases in total energetic requirements where having larger brains is advantageous  
1228 for resource acquisition, and in turn where resources are both exploitable and abundant enough to sustain  
1229 the larger energetic expenditures associated with bigger brain sizes. In addition, body size, which explains  
1230 an important proportion of the interspecific variation in mobility, likely interacts with characteristics of the  
1231 landscape (such as fragmentation) to predict species survival in disturbed land uses (Merckx et al., 2018).  
1232 Hence, the patterns I observe in this Chapter could emerge from interactions and trade-offs among resource  
1233 availability, diet, body size and human management of the landscape.

1234 At the species level, the results underline the role of energetic constraints on species responses to land-  
1235 use change. After controlling for the effects of body mass and taxonomy, I found that residual energetic  
1236 expenditure was a significant predictor of species occurrence probability in disturbed land uses. Contrary  
1237 to my expectations, in several disturbed land uses and in all trophic groups, species with higher residual ener-  
1238 getic expenditure tended to have a higher occurrence probability than species with lower residual ener-  
1239 getic expenditure. It could be that species with lower residual energetic requirements are less well equipped  
1240 than species with higher residual energetic requirements at making use of the available food sources in dis-  
1241 turbed land uses. Species in disturbed land uses may need to display higher levels of feeding innovation  
1242 (Coogan et al., 2018), have larger brains (Sayol et al., 2020), or be bolder and more active to make use of  
1243 the available resources. In mammals, past research has shown that larger brains are associated with larger  
1244 residual energetic expenditure (Isler and Van Schaik, 2006). At the individual level, past research suggests  
1245 that metabolic rates are linked to differences in behaviour, with bolder and more active individuals exhibiting  
1246 higher metabolic rates than less active individuals, with consequences for food acquisition (Biro and Stamps,  
1247 2010). Although I am not aware of similar evidence at the species level, I propose that residual metabolic  
1248 rates interact with species' ecological traits, behavioural traits and foraging strategy in influencing species'  
1249 responses to land use. Among species with an adaptable diet, able to make use of the resources available in  
1250 a disturbed landscape, those with higher residual metabolic rates could present a set of behavioural charac-  
1251 teristics that render them better at acquiring the available resources, hence more able to cope with land-use  
1252 disturbance. This could be the case for urban carnivores and omnivores, as well as herbivores in cropland.  
1253 Thus, it is possible that disturbed land uses favour species with higher residual resting metabolic rates be-

cause such species are overall better competitors when faced with disturbance. In addition, release from predation in some disturbed land uses – most notably in urban areas – could favour bolder species, as they are less exposed to predation risks. Conversely, and in accordance with my initial hypothesis, the results suggest that when there are no or few exploitable resources in disturbed areas, spending less energy than expected from body mass and taxonomy might be beneficial for persistence (this could be the case for urban herbivores, which include species specialised on fruit, nectar, and other plant materials that likely become less abundant and less substitutable in disturbed areas).

To summarise, I propose that, for species with adaptable diets, increases in exploitable resources coupled with decreases in predation rates leads to increases in overall energetic expenditure, associated with shifts towards assemblages composed of a higher abundance of medium-sized, bolder, more active and larger-brained species. On the other hand, species with specialised diets might be better able to persist and cope with resource scarcity when their energetic expenditure are lower than expected from body mass. These ideas could be further tested by considering more refined dietary groups (as those compiled in Chapter 4) rather than trophic groups, as diet should be easier to link with resource availability in a given land-use type. This could constitute a development for this work.

Finally, I would like to emphasize that energetic budget allocation is a fundamental aspect of organismal fitness (Robert Burger et al., 2021). Past studies suggest that individuals exposed to food scarcity can lower their intrinsic energy expenditure to increase their chances of survival. For instance, Zhang et al. (2018b) showed that, subjected to food restrictions, the Chinese bulbul (*Pycnonotus sinensis*) was able to lower its basal metabolic rate not only passively through body and organ mass reductions, but also by modifying enzymatic activity and metabolism in the muscles and liver. Other studies have also found that bird nestlings downregulate energetic costs allocated to maintenance when exposed to food shortage (Brzék and Konarzewski, 2001; Moe et al., 2004), or to disturbances such as urban noise (Brischoux et al., 2017). Intraspecific variation in energetic expenditure could affect individual responses to land-use change, which in turn could affect species' ability to adapt to human pressures. However, the lack of spatially explicit estimates of metabolic rates precluded its consideration in this Chapter.

To conclude, the findings of this Chapter suggest important effects of land-use change on energetic fluxes, and thus potentially on ecosystem functioning. Further research is warranted as to the effects of these energetic patterns for ecosystem processes. Indeed, assemblage-level energy fluxes may serve as important indicators of change for ecosystem processes, such as decomposition (Barnes et al., 2014). Understanding these impacts is particularly important given increasing land-use change and biodiversity loss across the globe.

1286 **6 | General discussion**

<sup>1287</sup> 7 | Conclusion

<sup>1288</sup>

# Bibliography

- <sup>1289</sup> Alex Smith, M. and M. Green, D. (2005). Dispersal and the metapopulation paradigm in amphibian ecology  
<sup>1290</sup> and conservation: are all amphibian populations metapopulations? *Ecography*, 28.1, 110–128.
- <sup>1291</sup> Auer, S. K., Solowey, J. R., Rajesh, S., and Rezende, E. L. (2020). Energetic mechanisms for coping with  
<sup>1292</sup> changes in resource availability. *Biology Letters*, 16.11, 1–8. DOI: 10.1098/rsbl.2020.0580.
- <sup>1293</sup> Balmford, A. (1996). Extinction filters and current resilience: The significance of past selection pressures for  
<sup>1294</sup> conservation biology. *Trends in Ecology and Evolution*. DOI: 10.1016/0169-5347(96)10026-4.
- <sup>1295</sup> Baraldi, A. N. and Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School  
1296 Psychology*. DOI: 10.1016/j.jsp.2009.10.001.
- <sup>1297</sup> Barber, N. A., Mooney, K. A., Greenberg, R., Philpott, S. M., Van Bael, S. A., and Gruner, D. S. (2010). Inter-  
<sup>1298</sup> actions among predators and the cascading effects of vertebrate insectivores on arthropod communities  
<sup>1299</sup> and plants. *Proceedings of the National Academy of Sciences*. DOI: 10.1073/pnas.1001934107.
- <sup>1300</sup> Barlow, J., França, F., Gardner, T. A., Hicks, C. C., Lennox, G. D., Berenguer, E., Castello, L., Economo,  
<sup>1301</sup> E. P., Ferreira, J., Guénard, B., Gontijo Leal, C., Isaac, V., Lees, A. C., Parr, C. L., Wilson, S. K., Young,  
<sup>1302</sup> P. J., and Graham, N. A. (2018). The future of hyperdiverse tropical ecosystems. *Nature*. DOI: 10.  
<sup>1303</sup> 1038/s41586-018-0301-1.
- <sup>1304</sup> Barnes, A. D., Jochum, M., Mumme, S., Haneda, N. F., Farajallah, A., Widarto, T. H., and Brose, U. (2014).  
<sup>1305</sup> Consequences of tropical land use for multitrophic biodiversity and ecosystem functioning. *Nature Communications*, 5, 1–7. DOI: 10.1038/ncomms6351.
- <sup>1307</sup> Bateman, P. W. and Fleming, P. A. (2012). Big city life: Carnivores in urban environments. *Journal of Zool-  
1308 ogy*, 287.1, 1–23. DOI: 10.1111/j.1469-7998.2011.00887.x.
- <sup>1309</sup> Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using  
<sup>1310</sup> lme4. *Journal of Statistical Software*. DOI: 10.18637/jss.v067.i01.
- <sup>1311</sup> Betts, M. G. et al. (2019). Extinction filters mediate the global effects of habitat fragmentation on animals.  
<sup>1312</sup> *Science*. DOI: 10.1126/science.aax9387.
- <sup>1313</sup> Biro, P. A. and Stamps, J. A. (2010). Do consistent individual differences in metabolic rate promote con-  
<sup>1314</sup> sistent individual differences in behavior? *Trends in Ecology and Evolution*, 25.11, 653–659. DOI: 10.  
<sup>1315</sup> 1016/j.tree.2010.08.003.
- <sup>1316</sup> Bivand, R. and Piras, G. (2015). Comparing Implementations of Estimation Methods for Spatial Economet-  
<sup>1317</sup> rics. *Journal of Statistical Software*, 63.18, 1–36.

- 1318 Bivand, R. and Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial  
1319 association. *TEST*, 27.3, 716–748.
- 1320 Bivand, R., Hauke, J., and Kossowski, T. (2013a). Computing the Jacobian in Gaussian spatial autoregressive  
1321 models: An illustrated comparison of available methods. *Geographical Analysis*, 45.2, 150–179.
- 1322 Bivand, R. S., Pebesma, E., and Gomez-Rubio, V. (2013b). *Applied spatial data analysis with R, Second  
1323 edition*. Springer, NY.
- 1324 Böhm, M., Cook, D., Ma, H., Davidson, A. D., García, A., Tapley, B., Pearce-Kelly, P., and Carr, J. (2016).  
1325 Hot and bothered: Using trait-based approaches to assess climate change vulnerability in reptiles. *Bio-  
1326 logical Conservation*. DOI: 10.1016/j.biocon.2016.06.002.
- 1327 Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., Rudolf, V. H., Schreiber,  
1328 S. J., Urban, M. C., and Vasseur, D. A. (2011). Why intraspecific trait variation matters in community  
1329 ecology. *Trends in Ecology and Evolution*. DOI: 10.1016/j.tree.2011.01.009.
- 1330 Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A., and  
1331 Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS  
1332 Computational Biology*. DOI: 10.1371/journal.pcbi.1003537.
- 1333 Bouckaert, R. et al. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis.  
1334 *PLoS Computational Biology*. DOI: 10.1371/journal.pcbi.1006650.
- 1335 Bregman, T. P., Lees, A. C., Seddon, N., Macgregor, H. E., Darski, B., Aleixo, A., Bonsall, M. B., and  
1336 Tobias, J. A. (2015). Species interactions regulate the collapse of biodiversity and ecosystem function in  
1337 tropical forest fragments. *Ecology*, 96.10, 2692–2704. DOI: 10.1890/14-1731.1.
- 1338 Bregman, T. P., Lees, A. C., MacGregor, H. E., Darski, B., Moura, N. G. de, Aleixo, A., Barlow, J., and  
1339 Tobias, J. A. (2016). Using avian functional traits to assess the impact of land-cover change on ecosystem  
1340 processes linked to resilience in tropical forests. *Proceedings of the Royal Society B: Biological Sciences*,  
1341 283.1844. DOI: 10.1098/rspb.2016.1289.
- 1342 Brischoux, F., Meillère, A., Dupoué, A., Lourdais, O., and Angelier, F. (2017). Traffic noise decreases  
1343 nestlings' metabolic rates in an urban exploiter. *Journal of Avian Biology*, 48.7, 905–909. DOI: 10.  
1344 1111/jav.01139.
- 1345 Bruggeman, J., Heringa, J., and Brandt, B. W. (2009). PhyloPars: Estimation of missing parameter values  
1346 using phylogeny. *Nucleic Acids Research*. DOI: 10.1093/nar/gkp370.
- 1347 Brzék, P. and Konarzewski, M. (2001). Effect of food shortage on the physiology and competitive abilities  
1348 of sand martin (*Riparia riparia*) nestlings. *Journal of Experimental Biology*, 204.17, 3065–3074. DOI:  
1349 10.1242/jeb.204.17.3065.
- 1350 Bushuev, A., Tolstenkov, O., Zubkova, E., Solovyeva, E., and Kerimov, A. (2018). Basal metabolic rate  
1351 in free-living tropical birds: The influence of phylogenetic, behavioral, and ecological factors. *Current  
1352 Zoology*, 64.1, 33–43. DOI: 10.1093/cz/zox018.
- 1353 Buuren, S van and Oudshoorn, C. G. M. (2007). MICE: multivariate imputation by chained equations. *R  
1354 package version*.
- 1355 Cadotte, M. W. and Tucker, C. M. (2017). Should Environmental Filtering be Abandoned? *Trends in Ecology  
1356 and Evolution*. DOI: 10.1016/j.tree.2017.03.004.

- 1357 Cadotte, M. W., Carscadden, K., and Mirochnick, N. (2011). Beyond species: Functional diversity and the  
1358 maintenance of ecological processes and services. *Journal of Applied Ecology*. DOI: 10.1111/j.  
1359 1365–2664.2011.02048.x.
- 1360 Cardoso, D. et al. (2017). Amazon plant diversity revealed by a taxonomically verified species list. *Proceedings of the National Academy of Sciences of the United States of America*, 114.40, 10695–10700. DOI:  
1361 10.1073/pnas.1706756114.
- 1362 Chamberlain, S. (2018). *rredlist: 'IUCN' Red List Client*. R package version 0.5.0.
- 1363 Chamberlain, S. A. and Szöcs, E. (2013). taxize : taxonomic search and retrieval in R [version 2; referees: 3  
1364 approved]. *F1000Research*. DOI: 10.12688/f1000research.2-191.v2.
- 1365 Chapman, P. M., Tobias, J. A., Edwards, D. P., and Davies, R. G. (2018). Contrasting impacts of land-use  
1366 change on phylogenetic and functional diversity of tropical forest birds. *Journal of Applied Ecology*.  
1367 DOI: 10.1111/1365–2664.13073.
- 1368 Chaudhary, C., Saeedi, H., and Costello, M. J. (2016). Bimodality of Latitudinal Gradients in Marine Species  
1369 Richness. *Trends in Ecology & Evolution*, 31.9, 670–676. DOI: <https://doi.org/10.1016/j.tree.2016.06.001>.
- 1370 Clarke, A. and Fraser, K. P. (2004). Why does metabolism scale with temperature? *Functional Ecology*, 18.2,  
1371 243–251. DOI: 10.1111/j.0269-8463.2004.00841.x.
- 1372 Colin, N., Villéger, S., Wilkes, M., Sostoa, A. de, and Maceda-Veiga, A. (2018). Functional diversity mea-  
1373 sures revealed impacts of non-native species and habitat degradation on species-poor freshwater fish  
1374 assemblages. *Science of the Total Environment*. DOI: 10.1016/j.scitotenv.2017.12.316.
- 1375 Collen, B., Ram, M., Zamin, T., and McRae, L. (2008). The Tropical Biodiversity Data Gap: Address-  
1376 ing Disparity in Global Monitoring. *Tropical Conservation Science*, 1.2, 75–88. DOI: 10.1177/  
1377 194008290800100202.
- 1378 Collen, B., Dulvy, N. K., Gaston, K. J., Gärdenfors, U., Keith, D. A., Punt, A. E., Regan, H. M., Böhm,  
1379 M., Hedges, S., Seddon, M., Butchart, S. H., Hilton-Taylor, C., Hoffmann, M., Bachman, S. P., and  
1380 Akçakaya, H. R. (2016). Clarifying misconceptions of extinction risk assessment with the IUCN Red  
1381 List. *Biology Letters*. DOI: 10.1098/rsbl.2015.0843.
- 1382 Coogan, S. C., Raubenheimer, D., Zantis, S. P., and Machovsky-Capuska, G. E. (2018). Multidimensional  
1383 nutritional ecology and urban birds. *Ecosphere*, 9.4. DOI: 10.1002/ecs2.2177.
- 1384 Cooke, R. S., Bates, A. E., and Eigenbrod, F. (2019a). Global trade-offs of functional redundancy and func-  
1385 tional dispersion for birds and mammals. *Global Ecology and Biogeography*, October 2018, 1–12. DOI:  
1386 10.1111/geb.12869.
- 1387 Cooke, R. S., Eigenbrod, F., and Bates, A. E. (2019b). Projected losses of global mammal and bird ecological  
1388 strategies. *Nature Communications*. DOI: 10.1038/s41467-019-10284-z.
- 1389 Cooper, N., Bielby, J., Thomas, G. H., and Purvis, A. (2008). Macroecology and extinction risk correlates of  
1390 frogs. *Global Ecology and Biogeography*. DOI: 10.1111/j.1466-8238.2007.00355.x.
- 1391 Cornwell, W. K., Schwilk, D. W., and Ackerly, D. D. (2006). A trait-based test for habitat filtering: Convex  
1392 hull volume. *Ecology*. DOI: 10.1890/0012-9658(2006)87[1465:ATTFHF]2.0.CO;2.
- 1393 Crooks, K. and Soule, M. (2010). Mesopredator release and avifaunal extinctions in a fragmented system.  
1394 *Nature*, 563–566.

- 1397 Davison, C. W., Rahbek, C., and Morueta-Holme, N. (2021). Land-use change and biodiversity: Challenges  
1398 for assembling evidence on the greatest threat to nature. *Global Change Biology*, 27.21, 5414–5429. DOI:  
1399 10.1111/gcb.15846.
- 1400 De Palma, A., Kuhlmann, M., Roberts, S. P., Potts, S. G., Börger, L., Hudson, L. N., Lysenko, I., Newbold,  
1401 T., and Purvis, A. (2015). Ecological traits affect the sensitivity of bees to land-use pressures in European  
1402 agricultural landscapes. *Journal of Applied Ecology*, 52.6, 1567–1577. DOI: 10.1111/1365–2664.  
1403 12524.
- 1404 Debastiani, V. J., Bastazini, V. A. G., and Pillar, D (2021). Ecological Informatics Using phylogenetic infor-  
1405 mation to impute missing functional trait values in ecological databases. 63.April. DOI: 10.1016/j.  
1406 ecoinf.2021.101315.
- 1407 Des Roches, S., Post, D. M., Turley, N. E., Bailey, J. K., Hendry, A. P., Kinnison, M. T., Schweitzer, J. A.,  
1408 and Palkovacs, E. P. (2018). The ecological importance of intraspecific variation. *Nature Ecology and*  
1409 *Evolution*. DOI: 10.1038/s41559-017-0402-5.
- 1410 Deutsch, C. A., Tewksbury, J. J., Huey, R. B., Sheldon, K. S., Ghilambor, C. K., Haak, D. C., and Martin,  
1411 P. R. (2008). Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the*  
1412 *National Academy of Sciences of the United States of America*, 105.18, 6668–6672. DOI: 10.1073/  
1413 pnas.0709472105.
- 1414 Díaz, S. et al. (2016). The global spectrum of plant form and function. *Nature*. DOI: 10.1038/nature16489.
- 1415 Diniz-Filho, J. A. F., Bini, L. M., Rangel, T. F., Morales-Castilla, I., Olalla-Tárraga, M. Á., Rodríguez,  
1416 M. Á., and Hawkins, B. A. (2012). On the selection of phylogenetic eigenvectors for ecological analyses.  
1417 *Ecography*. DOI: 10.1111/j.1600-0587.2011.06949.x.
- 1418 Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J., and Collen, B. (2014). Defaunation in the  
1419 Anthropocene. *Science*, 345.6195, 401–406. DOI: 10.1126/science.1251817.
- 1420 Dullinger, I., Essl, F., Moser, D., Erb, K., Haberl, H., and Dullinger, S. (2021). Biodiversity models need to  
1421 represent land-use intensity more comprehensively. *Global Ecology and Biogeography*, 30.5, 924–932.  
1422 DOI: 10.1111/geb.13289.
- 1423 Edwards, F. A., Edwards, D. P., Hamer, K. C., and Fayle, T. M. (2021). Tropical land-use change alters  
1424 trait-based community assembly rules for dung beetles and birds. *Oecologia*, 195.3, 705–717. DOI: 10.  
1425 1007/s00442-020-04829-z.
- 1426 Eeden, L. M. van et al. (2018). Carnivore conservation needs evidence-based livestock protection. *PLoS  
1427 Biology*, 16.9, 1–8. DOI: 10.1371/journal.pbio.2005577.
- 1428 Escudero, A. and Valladares, F. (2016). Trait-based plant ecology: moving towards a unifying species coex-  
1429 istence theory: Features of the Special Section. *Oecologia*. DOI: 10.1007/s00442-016-3578-5.
- 1430 ESRI (2011). ArcGIS desktop: Release 10. *Environmental Systems Research Institute*.
- 1431 Etard, A., Morrill, S., and Newbold, T. (2020). Global gaps in trait data for terrestrial vertebrates. *Global  
1432 Ecology and Biogeography*, November 2019, 1–16. DOI: 10.1111/geb.13184.
- 1433 Etard, A., Pigot, A. L., and Newbold, T. (2021). Intensive human land uses negatively affect vertebrate  
1434 functional diversity. *Ecology Letters*. DOI: <https://doi.org/10.1111/ele.13926>.

- 1435 Evans, B. S., Reitsma, R., Hurlbert, A. H., and Marra, P. P. (2018). Environmental filtering of avian com-  
1436 munities along a rural-to-urban gradient in Greater Washington, D.C., USA. *Ecosphere*, 9.11. DOI: 10 .  
1437 1002/ecs2.2402.
- 1438 Faurby, S., Davis, M., Pedersen, R., Schowanek, S. D., Antonelli1, A., and Svenning, J.-C. (2018). PHY-  
1439 LACINE 1.2: The Phylogenetic Atlas of Mammal Macroecology. *Ecology*, 99.11, 2626–2626. DOI:  
1440 10 .1002/ecy.2443.
- 1441 Faurby, S., Pedersen, R., Davis, M., Schowanek, S. D., Jarvie, S., Antonelli, A., and Svenning, J.-C. (2020).  
1442 PHYLACINE 1.2.1: An update to the Phylogenetic Atlas of Mammal Macroecology. DOI: 10 .5281/  
1443 zenodo.3690867.
- 1444 Feldman, A., Sabath, N., Pyron, R. A., Mayrose, I., and Meiri, S. (2016). Body sizes and diversification rates  
1445 of lizards, snakes, amphisbaenians and the tuatara. *Global Ecology and Biogeography*, 25.2, 187–197.  
1446 DOI: 10 .1111/geb .12398.
- 1447 Fick, S. E. and Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global  
1448 land areas. *International Journal of Climatology*, 37.12, 4302–4315. DOI: 10 .1002/joc .5086.
- 1449 Fischer, J. D., Cleton, S. H., Lyons, T. P., and Miller, J. R. (2012). Urbanization and the predation paradox:  
1450 The role of trophic dynamics in structuring vertebrate communities. *BioScience*, 62.9, 809–818. DOI:  
1451 10 .1525/bio .2012 .62 .9 .6.
- 1452 Flynn, D. F., Gogol-Prokurat, M., Nogeire, T., Molinari, N., Richers, B. T., Lin, B. B., Simpson, N., May-  
1453 field, M. M., and DeClerck, F. (2009). Loss of functional diversity under land use intensification across  
1454 multiple taxa. *Ecology Letters*, 12.1, 22–33. DOI: 10 .1111/j .1461-0248 .2008 .01255 .x.
- 1455 Fournier, B., Mouly, A., and Gillet, F. (2016). Multiple assembly rules drive the co-occurrence of orthopteran  
1456 and plant species in grasslands: Combining network, functional and phylogenetic approaches. *Frontiers  
1457 in Plant Science*, 7.AUG2016, 1–12. DOI: 10 .3389/fpls .2016 .01224.
- 1458 Freeman, B. G. and Pennell, M. W. (2021). The latitudinal taxonomy gradient. *Trends in Ecology and Evo-  
1459 lution*, 36.9, 778–786. DOI: 10 .1016/j.tree .2021 .05 .003.
- 1460 Frishkoff, L. O., Karp, D. S., Flanders, J. R., Zook, J., Hadly, E. A., Daily, G. C., and M’Gonigle, L. K.  
1461 (2016). Climate change and habitat conversion favour the same species. *Ecology letters*. DOI: 10 .1111/  
1462 ele .12645.
- 1463 Fukasawa, K. and Akasaka, T. (2019). Long-lasting effects of historical land use on the current distribution  
1464 of mammals revealed by ecological and archaeological patterns. *Scientific Reports*, 9.1, 1–11. DOI: 10 .  
1465 1038/s41598-019-46809-1.
- 1466 Furness, L. J. and Speakman, J. R. (2008). Energetics and longevity in birds. *Age*, 30.2-3, 75–87. DOI:  
1467 10 .1007/s11357-008-9054-3.
- 1468 Gainsbury, A. M., Tallowin, O. J., and Meiri, S. (2018). An updated global data set for diet preferences  
1469 in terrestrial mammals: testing the validity of extrapolation. *Mammal Review*. DOI: 10 .1111/mam .  
1470 12119.
- 1471 Galán-Acedo, C., Arroyo-Rodríguez, V., Andresen, E., and Arasa-Gisbert, R. (2019). Ecological traits of the  
1472 world’s primates. *Scientific data*. DOI: 10 .1038/s41597-019-0059-9.
- 1473 Gallmetzer, N. and Schulze, C. H. (2015). Impact of oil palm agriculture on understory amphibians and  
1474 reptiles: A Mesoamerican perspective. *Global Ecology and Conservation*, 4, 95–109. DOI: 10 .1016/  
1475 j.gecco .2015 .05 .008.

- 1476 Geffroy, B., Sadoul, B., Putman, B. J., Berger-Tal, O., Garamszegi, L. Z., Møller, A. P., and Blumstein, D. T.  
1477 (2020). Evolutionary dynamics in the anthropocene: Life history and intensity of human contact shape  
1478 antipredator responses. *PLoS Biology*, 18.9, 1–17. DOI: 10.1371/journal.pbio.3000818.
- 1479 González-Suárez, M. and Revilla, E. (2012). Variability in life-history and ecological traits is a buffer against  
1480 extinction in mammals. *Ecology Letters*, 16.2, 242–251. DOI: 10.1111/ele.12035.
- 1481 González-Suárez, M., Lucas, P. M., and Revilla, E. (2012). Biases in comparative analyses of extinction risk:  
1482 Mind the gap. *Journal of Animal Ecology*. DOI: 10.1111/j.1365-2656.2012.01999.x.
- 1483 González-Suárez, M., Zanchetta Ferreira, F., and Grilo, C. (2018). Spatial and species-level predictions of  
1484 road mortality risk using trait data. *Global Ecology and Biogeography*, 27.9, 1093–1105. DOI: 10 .  
1485 1111/geb.12769.
- 1486 Guralnick, R. P., Zermoglio, P. F., Wieczorek, J., LaFrance, R., Bloom, D., and Russell, L. (2016). The  
1487 importance of digitized biocollections as a source of trait data and a new VertNet resource. *Database*.  
1488 DOI: 10.1093/database/baw158.
- 1489 Hatfield, J. H., Harrison, M. L. K., and Banks-Leite, C. (2018). Functional Diversity Metrics: How They Are  
1490 Affected by Landscape Change and How They Represent Ecosystem Functioning in the Tropics. *Current  
1491 Landscape Ecology Reports*. DOI: 10.1007/s40823-018-0032-x.
- 1492 Hevia, V., Martín-López, B., Palomo, S., García-Llorente, M., Bello, F. de, and González, J. A. (2017). Trait-  
1493 based approaches to analyze links between the drivers of change and ecosystem services: Synthesizing  
1494 existing evidence and future challenges. *Ecology and Evolution*, 7.3, 831–844. DOI: https://doi .  
1495 org/10.1002/ece3.2692.
- 1496 Hooper, D. U., Adair, E. C., Cardinale, B. J., Byrnes, J. E., Hungate, B. A., Matulich, K. L., Gonzalez, A.,  
1497 Duffy, J. E., Gamfeldt, L., and Connor, M. I. (2012). A global synthesis reveals biodiversity loss as a  
1498 major driver of ecosystem change. *Nature*. DOI: 10.1038/nature11118.
- 1499 Hortal, J., Bello, F. de, Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., and Ladle, R. J. (2015). Seven  
1500 Shortfalls that Beset Large-Scale Knowledge of Biodiversity. *Annual Review of Ecology, Evolution, and  
1501 Systematics*. DOI: 10.1146/annurev-ecolsys-112414-054400.
- 1502 Hudson, L. N., Isaac, N. J., and Reuman, D. C. (2013). The relationship between body mass and field  
1503 metabolic rate among individual birds and mammals. *Journal of Animal Ecology*, 82.5, 1009–1020. DOI:  
1504 10.1111/1365-2656.12086.
- 1505 Hudson, L. N. et al. (2014). The PREDICTS database: A global database of how local terrestrial biodiversity  
1506 responds to human impacts. *Ecology and Evolution*. DOI: 10.1002/ece3.1303.
- 1507 — (2017). The database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing  
1508 Terrestrial Systems) project. *Ecology and Evolution*. DOI: 10.1002/ece3.2579.
- 1509 Inger, R., Cox, D. T., Per, E., Norton, B. A., and Gaston, K. J. (2016). Ecological role of vertebrate scavengers  
1510 in urban ecosystems in the UK. *Ecology and Evolution*. DOI: 10.1002/ece3.2414.
- 1511 Isaac, N. J., Mallet, J., and Mace, G. M. (2004). Taxonomic inflation: Its influence on macroecology and  
1512 conservation. *Trends in Ecology and Evolution*. DOI: 10.1016/j.tree.2004.06.004.
- 1513 Isler, K. and Van Schaik, C. P. (2006). Metabolic costs of brain size evolution. *Biology Letters*, 2.4, 557–560.  
1514 DOI: 10.1098/rsbl.2006.0538.
- 1515 IUCN (2020). The IUCN Red List of Threatened Species. <http://www.iucnredlist.org>. Version 2020.1.

- 1516 Jesse, W. A., Behm, J. E., Helmus, M. R., and Ellers, J. (2018). Human land use promotes the abundance  
1517 and diversity of exotic species on Caribbean islands. *Global Change Biology*, 24.10, 4784–4796. DOI:  
1518 10.1111/gcb.14334.
- 1519 Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., and Mooers, A. O. (2012). The global diversity of birds in  
1520 space and time. *Nature*. DOI: 10.1038/nature11631.
- 1521 Jetz, W. and Pyron, R. A. (2018). The interplay of past diversification and evolutionary isolation with  
1522 present imperilment across the amphibian tree of life. *Nature Ecology and Evolution*. DOI: 10.1038/  
1523 s41559-018-0515-5.
- 1524 Johnson, T. F., Isaac, N. J., Paviolo, A., and González-Suárez, M. (2021). Handling missing values in trait  
1525 data. *Global Ecology and Biogeography*. DOI: 10.1111/geb.13185.
- 1526 Jones, K. E. et al. (2009). PanTHERIA: a species-level database of life history, ecology, and geography of  
1527 extant and recently extinct mammals. *Ecology*. DOI: 10.1890/08-1494.1.
- 1528 Jones, O. R., Purvis, A., and Quicke, D. L. (2012). Latitudinal gradients in taxonomic overdescription rate  
1529 affect macroecological inferences using species list data. *Ecography*. DOI: 10.1111/j.1600-0587.  
1530 2011.06956.x.
- 1531 Kattge, J. et al. (2011). TRY - a global database of plant traits. *Global Change Biology*. DOI: 10.1111/j.  
1532 1365-2486.2011.02451.x.
- 1533 Kissling, W. D. et al. (2018). Towards global data products of Essential Biodiversity Variables on species  
1534 traits. *Nature Ecology and Evolution*. DOI: 10.1038/s41559-018-0667-3.
- 1535 Kissling, W. D., Dalby, L., Fløjgaard, C., Lenoir, J., Sandel, B., Sandom, C., Trøjelsgaard, K., and Svenning,  
1536 J. C. (2014). Establishing macroecological trait datasets: Digitalization, extrapolation, and validation of  
1537 diet preferences in terrestrial mammals worldwide. *Ecology and Evolution*. DOI: 10.1002/ece3.  
1538 1136.
- 1539 Krausmann, F., Erb, K. H., Gingrich, S., Haberl, H., Bondeau, A., Gaube, V., Lauk, C., Plutzar, C., and  
1540 Searchinger, T. D. (2013). Global human appropriation of net primary production doubled in the 20th  
1541 century. *Proceedings of the National Academy of Sciences of the United States of America*, 110.25,  
1542 10324–10329. DOI: 10.1073/pnas.1211349110.
- 1543 Krauss, J., Bommarco, R., Guardiola, M., Heikkinen, R. K., Helm, A., Kuussaari, M., Lindborg, R., Öckinger,  
1544 E., Pärtel, M., Pino, J., Pöyry, J., Raatikainen, K. M., Sang, A., Stefanescu, C., Teder, T., Zobel, M., and  
1545 Steffan-Dewenter, I. (2010). Habitat fragmentation causes immediate and time-delayed biodiversity loss  
1546 at different trophic levels. *Ecology Letters*. DOI: 10.1111/j.1461-0248.2010.01457.x.
- 1547 La Sorte, F. A., Lepczyk, C. A., Aronson, M. F., Goddard, M. A., Hedblom, M., Katti, M., MacGregor-Fors,  
1548 I., Mörtberg, U., Nilón, C. H., Warren, P. S., Williams, N. S., and Yang, J. (2018). The phylogenetic and  
1549 functional diversity of regional breeding bird assemblages is reduced and constricted through urbanization.  
1550 *Diversity and Distributions*. DOI: 10.1111/ddi.12738.
- 1551 Laliberte, E. and Legendre, P. (2010). A distance-based framework for measuring functional diversity from  
1552 multiple traits. *Ecology*, 91.1, 299–305. DOI: 10.1890/08-2244.1.
- 1553 Laliberté, E., Legendre, P., and Shipley, B. (2015). FD: measuring functional diversity from multiple traits,  
1554 and other tools for functional ecology. *R Package*.

- 1555 Lavorel, S. and Garnier, E. (2002). Predicting changes in community composition and ecosystem functioning  
1556 from plant traits: Revisiting the Holy Grail. *Functional Ecology*. DOI: 10.1046/j.1365-2435.  
1557 2002.00664.x.
- 1558 Le Provost, G., Badenhausen, I., Le Bagousse-Pinguet, Y., Clough, Y., Henckel, L., Violle, C., Bretagnolle,  
1559 V., Roncoroni, M., Manning, P., and Gross, N. (2020). Land-use history impacts functional diversity  
1560 across multiple trophic groups. *Proceedings of the National Academy of Sciences of the United States of  
1561 America*. DOI: 10.1073/pnas.1910023117.
- 1562 Legras, G., Loiseau, N., and Gaertner, J. C. (2018). Functional richness: Overview of indices and underlying  
1563 concepts. *Acta Oecologica*. DOI: 10.1016/j.actao.2018.02.007.
- 1564 Letnic, M., Ritchie, E. G., and Dickman, C. R. (2012). Top predators as biodiversity regulators: The dingo  
1565 *Canis lupus dingo* as a case study. *Biological Reviews*. DOI: 10.1111/j.1469-185X.2011.  
1566 00203.x.
- 1567 Lowry, H., Lill, A., and Wong, B. B. (2013). Behavioural responses of wildlife to urban environments.  
1568 *Biological Reviews*, 88.3, 537–549. DOI: 10.1111/brv.12012.
- 1569 Luck, G. W., Lavorel, S., McIntyre, S., and Lumb, K. (2012). Improving the application of vertebrate trait-  
1570 based frameworks to the study of ecosystem services. *Journal of Animal Ecology*. DOI: 10.1111/j.  
1571 1365-2656.2012.01974.x.
- 1572 Magioli, M., Ferraz, K. M. P. M. d. B., Chiarello, A. G., Galetti, M., Setz, E. Z. F., Paglia, A. P., Abrego,  
1573 N., Ribeiro, M. C., and Ovaskainen, O. (2021). Land-use changes lead to functional loss of terrestrial  
1574 mammals in a Neotropical rainforest. *Perspectives in Ecology and Conservation*, 19.2, 161–170. DOI:  
1575 10.1016/j.pecon.2021.02.006.
- 1576 Májeková, M., Paal, T., Plowman, N. S., Bryndová, M., Kasari, L., Norberg, A., Weiss, M., Bishop, T. R.,  
1577 Luke, S. H., Sam, K., Le Bagousse-Pinguet, Y., Lepš, J., Götzenberger, L., and De Bello, F. (2016).  
1578 Evaluating Functional diversity: Missing trait data and the importance of species abundance structure  
1579 and data transformation. *PLoS ONE*. DOI: 10.1371/journal.pone.0149270.
- 1580 Marcacci, G., Westphal, C., Wenzel, A., Raj, V., Nölke, N., Tscharntke, T., and Grass, I. (2021). Taxonomic  
1581 and functional homogenization of farmland birds along an urbanization gradient in a tropical megacity.  
1582 *Global Change Biology*, June, 1–15. DOI: 10.1111/gcb.15755.
- 1583 Martin, L. J., Blossey, B., and Ellis, E. (2012). Mapping where ecologists work: biases in the global distri-  
1584 bution of terrestrial ecological observations. *Frontiers in Ecology and the Environment*, 10.4, 195–201.  
1585 DOI: 10.1890/110154.
- 1586 Martins, I. S. and Pereira, H. M. (2017). Improving extinction projections across scales and habitats using the  
1587 countryside species-area relationship. *Scientific Reports*. DOI: 10.1038/s41598-017-13059-y.
- 1588 Matuoka, M. A., Benchimol, M., Almeida-Rocha, J. M. de, and Morante-Filho, J. C. (2020). Effects of  
1589 anthropogenic disturbances on bird functional diversity: A global meta-analysis. *Ecological Indicators*.  
1590 DOI: 10.1016/j.ecolind.2020.106471.
- 1591 Maxwell, S. L., Fuller, R. A., Brooks, T. M., and Watson, J. E. M. (2016). Biodiversity: The ravages of guns,  
1592 nets and bulldozers. *Nature*. DOI: 10.1038/536143a.
- 1593 May, R. M. (2011). Why worry about how many species and their loss? *PLoS Biology*. DOI: 10.1371/  
1594 journal.pbio.1001130.

- 1595 McNab, B. K. (1986). The Influence of Food Habits on the Energetics of Eutherian Mammals. *Ecological  
1596 Monographs*, 56.1, 1–19.
- 1597 Meiri, S., Feldman, A., and Kratochvíl, L. (2015). Squamate hatchling size and the evolutionary causes of  
1598 negative offspring size allometry. *Journal of Evolutionary Biology*, 28.2, 438–446. DOI: 10.1111/  
1599 jeb.12580.
- 1600 Meiri, S. (2018). Traits of lizards of the world: Variation around a successful evolutionary design. *Global  
1601 Ecology and Biogeography*, 27.10, 1168–1172. DOI: 10.1111/geb.12773.
- 1602 Mendoza, M. and Araújo, M. B. (2019). Climate shapes mammal community trophic structures and humans  
1603 simplify them. *Nature Communications*, 10.1. DOI: 10.1038/s41467-019-12995-9.
- 1604 Menegotto, A. and Rangel, T. F. (2018). Mapping knowledge gaps in marine diversity reveals a latitudi-  
1605 nal gradient of missing species richness. *Nature Communications*. DOI: 10.1038/s41467-018-  
1606 07217-7.
- 1607 Merckx, T. et al. (2018). Body-size shifts in aquatic and terrestrial urban communities. *Nature*. DOI: 10.  
1608 1038/s41586-018-0140-0.
- 1609 Millard, J., Outhwaite, C. L., Kinnersley, R., Freeman, R., Gregory, R. D., Adedoja, O., Gavini, S., Kioko, E.,  
1610 Kuhlmann, M., Ollerton, J., Ren, Z.-x., and Newbold, T. (2021). Global effects of land-use intensity on  
1611 local pollinator biodiversity. *Nature Communications*, 1–11. DOI: 10.1038/s41467-021-23228-  
1612 3.
- 1613 Moe, B., Brunvoll, S., Mork, D., Brobak, T. E., and Bech, C. (2004). Developmental plasticity of physi-  
1614 ology and morphology in diet-restricted European shag nestlings (*Phalacrocorax aristotelis*). *Journal of  
1615 Experimental Biology*, 207.23, 4067–4076. DOI: 10.1242/jeb.01226.
- 1616 Møller, A. P. (2009). Successful city dwellers: A comparative study of the ecological characteristics of urban  
1617 birds in the Western Palearctic. *Oecologia*, 159.4, 849–858. DOI: 10.1007/s00442-008-1259-8.
- 1618 Mouillot, D., Loiseau, N., Grenié, M., Algar, A. C., Allegra, M., Cadotte, M. W., Casajus, N., Denelle, P.,  
1619 Guéguen, M., Maire, A., Maitner, B., McGill, B. J., McLean, M., Mouquet, N., Munoz, F., Thuiller, W.,  
1620 Villéger, S., Violle, C., and Auber, A. (2021). The dimensionality and structure of species trait spaces.  
1621 *Ecology Letters*, March, 1–22. DOI: 10.1111/ele.13778.
- 1622 Mueller, P. and Diamond, J. (2001). Metabolic rate and environmental productivity: Well-provisioned ani-  
1623 mals evolved to run and idle fast. *Proceedings of the National Academy of Sciences of the United States  
1624 of America*, 98.22, 12550–12554. DOI: 10.1073/pnas.221456698.
- 1625 Munteanu, C., Kamp, J., Nita, M. D., Klein, N., Kraemer, B. M., Müller, D., Koshkina, A., Prishchepov,  
1626 A. V., and Kuemmerle, T. (2020). Cold War spy satellite images reveal long-term declines of a philopatric  
1627 keystone species in response to cropland expansion. *Proceedings of the Royal Society B: Biological  
1628 Sciences*, 287.1927. DOI: 10.1098/rspb.2019.2897.
- 1629 Myhrvold, N. P., Baldridge, E., Chan, B., Sivam, D., Freeman, D. L., and Ernest, S. K. M. (2015). An amniote  
1630 life-history database to perform comparative analyses with birds, mammals, and reptiles. *Ecology*. DOI:  
1631 10.1890/15-0846R.1.
- 1632 Nakagawa, S. and Freckleton, R. P. (2008). Missing inaction: the dangers of ignoring missing data. *Trends  
1633 in Ecology and Evolution*. DOI: 10.1016/j.tree.2008.06.014.

- 1634 Naya, D. E., Spangenberg, L., Naya, H., and Bozinovic, F. (2013). How does evolutionary variation in basal  
1635 metabolic rates arise? A statistical assessment and a mechanistic model. *Evolution*, 67.5, 1463–1476.  
1636 DOI: 10.1111/evo.12042.
- 1637 Newbold, T., Scharlemann, J. P. W., Butchart, S. H. M., Sekercioğlu, C. H., Alkemade, R., Booth, H., and  
1638 Purves, D. W. (2013). Ecological traits affect the response of tropical forest bird species to land-use  
1639 intensity. *Proceedings. Biological sciences / The Royal Society*. DOI: 10.1098/rspb.2012.2131.
- 1640 Newbold, T. et al. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*. DOI: 10 .  
1641 1038/nature14324.
- 1642 Newbold, T., Hudson, L. N., Contu, S., Hill, S. L., Beck, J., Liu, Y., Meyer, C., Phillips, H. R., Scharlemann,  
1643 J. P., and Purvis, A. (2018). Widespread winners and narrow-ranged losers: Land use homogenizes biodi-  
1644 versity in local assemblages worldwide. *PLoS Biology*. DOI: 10.1371/journal.pbio.2006841.
- 1645 Newbold, T., Bentley, L. F., Hill, S. L., Edgar, M. J., Horton, M., Su, G., Şekercioğlu, Ç. H., Collen, B., and  
1646 Purvis, A. (2020a). Global effects of land use on biodiversity differ among functional groups. *Functional  
1647 Ecology*, 34.3, 684–693. DOI: 10.1111/1365-2435.13500.
- 1648 Newbold, T., Oppenheimer, P., Etard, A., and Williams, J. J. (2020b). Tropical and Mediterranean biodiver-  
1649 sity is disproportionately sensitive to land-use and climate change. *Nature Ecology and Evolution*. DOI:  
1650 10.1038/s41559-020-01303-0.
- 1651 Novosolov, M., Raia, P., and Meiri, S. (2013). The island syndrome in lizards. *Global Ecology and Biogeog-  
1652 raphy*. DOI: 10.1111/j.1466-8238.2012.00791.x.
- 1653 Novosolov, M., Rodda, G. H., North, A. C., Butchart, S. H., Tallowin, O. J., Gainsbury, A. M., and Meiri,  
1654 S. (2017). Population density–range size relationship revisited. *Global Ecology and Biogeography*. DOI:  
1655 10.1111/geb.12617.
- 1656 Nowakowski, A. J., Thompson, M. E., Donnelly, M. A., and Todd, B. D. (2017). Amphibian sensitivity to  
1657 habitat modification is associated with population trends and species traits. *Global Ecology and Biogeog-  
1658 raphy*. DOI: 10.1111/geb.12571.
- 1659 Nowakowski, A. J., Frishkoff, L. O., Thompson, M. E., Smith, T. M., and Todd, B. D. (2018). Phylogenetic  
1660 homogenization of amphibian assemblages in human-altered habitats across the globe. *Proceedings of  
1661 the National Academy of Sciences of the United States of America*, 115.15, E3454–E3462. DOI: 10 .  
1662 1073/pnas.1714891115.
- 1663 Oliveira, B. F., São-Pedro, V. A., Santos-Barrera, G., Penone, C., and Costa, G. C. (2017). AmphiBIO, a  
1664 global database for amphibian ecological traits. *Scientific Data*. DOI: 10.1038/sdata.2017.123.
- 1665 Oliver, T. H., Heard, M. S., Isaac, N. J., Roy, D. B., Procter, D., Eigenbrod, F., Freckleton, R., Hector, A.,  
1666 Orme, C. D. L., Petachey, O. L., Proença, V., Raffaelli, D., Suttle, K. B., Mace, G. M., Martín-López,  
1667 B., Woodcock, B. A., and Bullock, J. M. (2015). Biodiversity and Resilience of Ecosystem Functions.  
1668 *Trends in Ecology and Evolution*. DOI: 10.1016/j.tree.2015.08.009.
- 1669 Pacifici, M. et al. (2015). Assessing species vulnerability to climate change. *Nature Climate Change*. DOI:  
1670 10.1038/nclimate2448.
- 1671 Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*. DOI: 10.1038/44766.
- 1672 Pakeman, R. J. (2014). Functional trait metrics are sensitive to the completeness of the species' trait data?  
1673 *Methods in Ecology and Evolution*. DOI: 10.1111/2041-210X.12136.

- 1674 Paradis, E., Baillie, S. R., Sutherland, W. J., and Gregory, R. D. (1998). Patterns of natal and breeding  
1675 dispersal in birds. *Journal of Animal Ecology*, 67.4, 518–536. DOI: 10.1046/j.1365-2656.  
1676 1998.00215.x.
- 1677 Pearson, R. G., Stanton, J. C., Shoemaker, K. T., Aiello-Lammens, M. E., Ersts, P. J., Horning, N., Ford-  
1678 ham, D. A., Raxworthy, C. J., Ryu, H. Y., Mcnees, J., and Akçakaya, H. R. (2014). Life history and  
1679 spatial traits predict extinction risk due to climate change. *Nature Climate Change*. DOI: 10.1038/  
1680 nclimate2113.
- 1681 Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M., Young, B. E.,  
1682 Graham, C. H., and Costa, G. C. (2014). Imputation of missing data in life-history trait datasets: Which  
1683 approach performs the best? *Methods in Ecology and Evolution*. DOI: 10.1111/2041-210X.12232.
- 1684 Perronne, R., Munoz, F., Borgy, B., Reboud, X., and Gaba, S. (2017). How to design trait-based analyses  
1685 of community assembly mechanisms: Insights and guidelines from a literature review. *Perspectives in  
1686 Plant Ecology, Evolution and Systematics*. DOI: 10.1016/j.ppees.2017.01.004.
- 1687 Peugh, J. L. and Enders, C. K. (2004). Missing data in educational research: A review of reporting practices  
1688 and suggestions for improvement. *Review of Educational Research*. DOI: 10.3102/00346543074004525.
- 1689 Purvis, A., Gittleman, J. L., Cowlishaw, G., and Mace, G. M. (2000). Predicting extinction risk in declining  
1690 species. *Proceedings of the Royal Society B: Biological Sciences*. DOI: 10.1098/rspb.2000.1234.
- 1691 Pyšek, P., Bacher, S., Chytrý, M., Jarošík, V., Wild, J., Celesti-Grapow, L., Gassó, N., Kenis, M., Lambdon,  
1692 P. W., Nentwig, W., Pergl, J., Roques, A., Sádlo, J., Solarz, W., Vilà, M., and Hulme, P. E. (2010).  
1693 Contrasting patterns in the invasions of European terrestrial and freshwater habitats by alien plants,  
1694 insects and vertebrates. *Global Ecology and Biogeography*, 19.3, 317–331. DOI: 10.1111/j.1466-  
1695 8238.2009.00514.x.
- 1696 Quesnelle, P. E., Lindsay, K. E., and Fahrig, L. (2014). Low reproductive rate predicts species sensitivity to  
1697 habitat loss: A meta-analysis of wetland vertebrates. *PLoS ONE*. DOI: 10.1371/journal.pone.  
1698 0090926.
- 1699 R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical  
1700 Computing. Vienna, Austria.
- 1701 Ramesh, T., Kalle, R., and Downs, C. T. (2017). Space use in a South African agriculture landscape by the  
1702 caracal (Caracal caracal). *European Journal of Wildlife Research*, 63.1. DOI: 10.1007/s10344-  
1703 016-1072-3.
- 1704 Ramírez-Bautista, A., Thorne, J. H., Schwartz, M. W., and Williams, J. N. (2020). Trait-based climate vulner-  
1705 ability of native rodents in southwestern Mexico. *Ecology and Evolution*. DOI: 10.1002/ece3.6323.
- 1706 Ratto, F., Simmons, B. I., Spake, R., Zamora-Gutierrez, V., MacDonald, M. A., Merriman, J. C., Tremlett,  
1707 C. J., Poppy, G. M., Peh, K. S., and Dicks, L. V. (2018). Global importance of vertebrate pollinators for  
1708 plant reproductive success: a meta-analysis. *Frontiers in Ecology and the Environment*. DOI: 10.1002/  
1709 fee.1763.
- 1710 Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Meth-  
1711 ods in Ecology and Evolution*, 3, 217–223.
- 1712 Riemann, J. C., Ndriantsoa, S. H., Rödel, M. O., and Glos, J. (2017). Functional diversity in a fragmented  
1713 landscape — Habitat alterations affect functional trait composition of frog assemblages in Madagascar.  
1714 *Global Ecology and Conservation*, 10, 173–183. DOI: 10.1016/j.gecco.2017.03.005.

- 1715 Ripple, W. J., Wolf, C., Newsome, T. M., Hoffmann, M., Wirsing, A. J., and McCauley, D. J. (2017). Extinction risk is most acute for the world's largest and smallest vertebrates. *Proceedings of the National Academy of Sciences*. DOI: 10.1073/pnas.1702078114.
- 1718 Robert Burger, J., Hou, C., A. S. Hall, C., and Brown, J. H. (2021). Universal rules of life: metabolic rates, biological times and the equal fitness paradigm. *Ecology Letters*, 24.6, 1262–1281. DOI: 10.1111/ele.13715.
- 1721 Roll, U. et al. (2017). The global distribution of tetrapods reveals a need for targeted reptile conservation. *Nature Ecology and Evolution*. DOI: 10.1038/s41559-017-0332-2.
- 1723 Running, S. W. and Zhao, M. (2015). Daily GPP and Annual NPP (MOD17A2/A3) products NASA Earth Observing System MODIS Land Algorithm (User's guide V3). *User Guide*.
- 1725 Salo, P., Banks, P. B., Dickman, C. R., and Korpimäki, E. (2010). Predator manipulation experiments: Impacts on populations of terrestrial vertebrate prey. *Ecological Monographs*. DOI: 10.1890/09-1260.1.
- 1728 Samia, D. S., Nakagawa, S., Nomura, F., Rangel, T. F., and Blumstein, D. T. (2015). Increased tolerance to humans among disturbed wildlife. *Nature Communications*, 6. DOI: 10.1038/ncomms9877.
- 1730 Sanchez-Ortiz, K., Taylor, K. J., Palma, A. de, König, C., Pyšek, P., Weigelt, P., Essl, F., and Purvis, A. (2019). Effects of land-use change and related pressures on alien and native subsets of island communities. DOI: 10.1101/2019.12.16.878041.
- 1733 Santini, L., Benítez-López, A., Ficetola, G. F., and Huijbregts, M. A. J. (2018). Length–mass allometries in amphibians. *Integrative Zoology*, 13.1, 36–45. DOI: 10.1111/1749-4877.12268.
- 1735 Santos, T. (2018). Package ‘PVR’. Phylogenetic Eigenvectors Regression and Phylogenetic Signal-Representation Curve.
- 1737 Sayol, F., Sol, D., and Pigot, A. L. (2020). Brain Size and Life History Interact to Predict Urban Tolerance in Birds. *Frontiers in Ecology and Evolution*, 8, 58. DOI: 10.3389/fevo.2020.00058.
- 1739 Scharf, I., Feldman, A., Novosolov, M., Pincheira-Donoso, D., Das, I., Böhm, M., Uetz, P., Torres-Carvajal, O., Bauer, A., Roll, U., and Meiri, S. (2015). Late bloomers and baby boomers: Ecological drivers of longevity in squamates and the tuatara. *Global Ecology and Biogeography*. DOI: 10.1111/geb.12244.
- 1743 Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., Réale, D., Dochtermann, N. A., Garamszegi, L. Z., and Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution*, 11.9, 1141–1152. DOI: 10.1111/2041-210X.13434.
- 1747 Schloss, C. A., Nunez, T. A., and Lawler, J. J. (2012). Dispersal will limit ability of mammals to track climate change in the Western Hemisphere. *Proceedings of the National Academy of Sciences*. DOI: 10.1073/pnas.1116791109.
- 1750 Schneider, F. D., Fichtmueller, D., Gossner, M. M., Güntsch, A., Jochum, M., König-Ries, B., Le Provost, G., Manning, P., Ostrowski, A., Penone, C., and Simons, N. K. (2019). Towards an ecological trait-data standard. *Methods in Ecology and Evolution*. DOI: 10.1111/2041-210X.13288.
- 1753 Schwarz, R. and Meiri, S. (2017). The fast-slow life-history continuum in insular lizards: a comparison between species with invariant and variable clutch sizes. *Journal of Biogeography*. DOI: 10.1111/jbi.13067.

- 1756 Sévèque, A., Gentle, L. K., López-Bao, J. V., Yarnell, R. W., and Uzal, A. (2020). Human disturbance has  
1757 contrasting effects on niche partitioning within carnivore communities. *Biological Reviews*, 95.6, 1689–  
1758 1705. DOI: 10.1111/brv.12635.
- 1759 Siefert, A. et al. (2015). A global meta-analysis of the relative extent of intraspecific trait variation in plant  
1760 communities. *Ecology Letters*, 18.12, 1406–1419. DOI: 10.1111/ele.12508.
- 1761 Slavenko, A., Tallowin, O. J., Itescu, Y., Raia, P., and Meiri, S. (2016). Late Quaternary reptile extinctions:  
1762 size matters, insularity dominates. *Global Ecology and Biogeography*. DOI: 10.1111/geb.12491.
- 1763 Sodhi, N. S., Bickford, D., Diesmos, A. C., Lee, T. M., Koh, L. P., Brook, B. W., Sekercioglu, C. H., and  
1764 Bradshaw, C. J. (2008). Measuring the meltdown: Drivers of global amphibian extinction and decline.  
1765 *PLoS ONE*. DOI: 10.1371/journal.pone.0001636.
- 1766 Sol, D., Trisos, C., Múrria, C., Jeliazkov, A., González-Lagos, C., Pigot, A. L., Ricotta, C., Swan, C. M.,  
1767 Tobias, J. A., and Pavoine, S. (2020). The worldwide impact of urbanisation on avian functional diversity.  
1768 *Ecology Letters*, 23.6, 962–972. DOI: 10.1111/ele.13495.
- 1769 Stark, G., Tamar, K., Itescu, Y., Feldman, A., and Meiri, S. (2018). Cold and isolated ectotherms: drivers  
1770 of reptilian longevity. *Biological Journal of the Linnean Society*. DOI: 10.1093/biolinnean/  
1771 bly153/5145102.
- 1772 Stark, G., Pincheira-Donoso, D., and Meiri, S. (2020). No evidence for the ‘rate-of-living’ theory across the  
1773 tetrapod tree of life. *Global Ecology and Biogeography*, 29.5, 857–884. DOI: 10.1111/geb.13069.
- 1774 Stehfest, E., Zeist, W. J. van, Valin, H., Havlik, P., Popp, A., Kyle, P., Tabeau, A., Mason-D'Croz, D.,  
1775 Hasegawa, T., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fujimori, S., Humpenöder, F., Lotze-Campen,  
1776 H., Meijl, H. van, and Wiebe, K. (2019). Key determinants of global land-use projections. *Nature Com-*  
1777 *munications*, 10.1, 1–10. DOI: 10.1038/s41467-019-09945-w.
- 1778 Stekhoven, D. J. (2016). Nonparametric Missing Value Imputation using Random Forest. *R Package version*  
1779 *1.4*. DOI: 10.1093/bioinformatics/btr597.
- 1780 Stekhoven, D. J. and Bühlmann, P. (2012). Missforest-Non-parametric missing value imputation for mixed-  
1781 type data. *Bioinformatics*. DOI: 10.1093/bioinformatics/btr597.
- 1782 Stephens, L. et al. (2019). Archaeological assessment reveals Earth’s early transformation through land use.  
1783 *Science*, 365.6456, 897 LP –902. DOI: 10.1126/science.aax1192.
- 1784 Sunday, J. M., Bates, A. E., Kearney, M. R., Colwell, R. K., Dulvy, N. K., Longino, J. T., and Huey,  
1785 R. B. (2014). Thermal-safety margins and the necessity of thermoregulatory behavior across latitude  
1786 and elevation. *Proceedings of the National Academy of Sciences of the United States of America*. DOI:  
1787 10.1073/pnas.1316145111.
- 1788 Sutherland, G. D., Harestad, A. S., Price, K., and Lertzman, K. P. (2000). Scaling of natal dispersal distances  
1789 in terrestrial birds and mammals. *Ecology and Society*, 4.1. DOI: 10.5751/es-00184-040116.
- 1790 Tilman, D. and Downing, J. A. (1994). Biodiversity and stability in grasslands. *Nature*. DOI: 10.1038/  
1791 367363a0.
- 1792 Tinoco, B. A., Santillán, V. E., and Graham, C. H. (2018). Land use change has stronger effects on functional  
1793 diversity than taxonomic diversity in tropical Andean hummingbirds. *Ecology and Evolution*. DOI: 10.  
1794 1002/ece3.3813.

- 1795 Titley, M. A., Snaddon, J. L., and Turner, E. C. (2017). Scientific research on animal biodiversity is sys-  
1796 tematically biased towards vertebrates and temperate regions. *PLoS ONE*. DOI: 10.1371/journal.  
1797 pone.0189577.
- 1798 Todd, B. D., Nowakowski, A. J., Rose, J. P., and Price, S. J. (2017). Species traits explaining sensitivity  
1799 of snakes to human land use estimated from citizen science data. *Biological Conservation*. DOI: 10 .  
1800 1016/j.biocon.2016.12.013.
- 1801 Tonini, J. F. R., Beard, K. H., Ferreira, R. B., Jetz, W., and Pyron, R. A. (2016). Fully-sampled phylogenies  
1802 of squamates reveal evolutionary patterns in threat status. *Biological Conservation*. DOI: 10.1016/j.  
1803 biocon.2016.03.039.
- 1804 Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., and Altman,  
1805 R. B. (2001). Missing value estimation methods for DNA microarrays. *Bioinformatics*. DOI: 10.1093/  
1806 bioinformatics/17.6.520.
- 1807 Tylianakis, J. M., Didham, R. K., Bascompte, J., and Wardle, D. A. (2008). Global change and species  
1808 interactions in terrestrial ecosystems. *Ecology Letters*, 11.12, 1351–1363. DOI: 10.1111/j.1461-  
1809 0248.2008.01250.x.
- 1810 United Nations Educational Scientific and Cultural and Organization (2015). *UNESCO Global Science Re-  
1811 port: Towards 2030*. Tech. rep.
- 1812 Valiente-Banuet, A., Aizen, M. A., Alcántara, J. M., Arroyo, J., Cocucci, A., Galetti, M., García, M. B.,  
1813 García, D., Gómez, J. M., Jordano, P., Medel, R., Navarro, L., Obeso, J. R., Oviedo, R., Ramírez, N., Rey,  
1814 P. J., Traveset, A., Verdú, M., and Zamora, R. (2015). Beyond species loss: The extinction of ecological  
1815 interactions in a changing world. *Functional Ecology*, 29.3, 299–307. DOI: 10.1111/1365-2435.  
1816 12356.
- 1817 van Rossum, G. (1995). *Python tutorial*. Centrum voor Wiskunde en Informatica (CWI).
- 1818 Verberk, W. C. E. P., Noordwijk, C. G. E. van, and Hildrew, A. G. (2013). Delivering on a promise: integrating  
1819 species traits to transform descriptive community ecology into a predictive science. *Freshwater Science*.  
1820 DOI: 10.1899/12-092.1.
- 1821 Vidan, E., Roll, U., Bauer, A., Grismer, L., Guo, P., Maza, E., Novosolov, M., Sindaco, R., Wagner, P.,  
1822 Belmaker, J., and Meiri, S. (2017). The Eurasian hot nightlife: Environmental forces associated with  
1823 nocturnality in lizards. *Global Ecology and Biogeography*. DOI: 10.1111/geb.12643.
- 1824 Villéger, S., Mason, N. W. H., and Mouillot, D. (2008). New multidimensional functional diversity indices  
1825 for a multifaceted framework in functional ecology. *Ecology*.
- 1826 Violle, C., Navas, M. L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., and Garnier, E. (2007). Let the  
1827 concept of trait be functional! *Oikos*. DOI: 10.1111/j.0030-1299.2007.15559.x.
- 1828 Weiner, C. N., Werner, M., Linsenmair, K. E., and Blüthgen, N. (2011). Land use intensity in grasslands:  
1829 Changes in biodiversity, species composition and specialisation in flower visitor networks. *Basic and  
1830 Applied Ecology*, 12.4, 292–299. DOI: 10.1016/j.baae.2010.08.006.
- 1831 Weiss, K. C. and Ray, C. A. (2019). Unifying functional trait approaches to understand the assemblage of  
1832 ecological communities: synthesizing taxonomic divides. *Ecography*. DOI: 10.1111/ecog.04387.
- 1833 White, C. R. and Kearney, M. R. (2013). Determinants of inter-specific variation in basal metabolic rate.  
1834 *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, 183.1,  
1835 1–26. DOI: 10.1007/s00360-012-0676-5.

- 1836 White, C. R. and Seymour, R. S. (2003). Mammalian basal metabolic rate is proportional to body mass $^{2/3}$ .  
1837 *Proceedings of the National Academy of Sciences of the United States of America*, 100.7, 4046–4049.  
1838 DOI: 10.1073/pnas.0436428100.
- 1839 Whitmee, S. and Orme, C. D. L. (2013). Predicting dispersal distance in mammals: A trait-based approach.  
1840 *Journal of Animal Ecology*, 82.1, 211–221. DOI: 10.1111/j.1365-2656.2012.02030.x.
- 1841 Williams, J. J. and Newbold, T. (2020). Local climatic changes affect biodiversity responses to land use: A  
1842 review. *Diversity and Distributions*. DOI: 10.1111/ddi.12999.
- 1843 Wilman, H., Belmaker, J., Simpson, J., Rosa, C. de la, Rivadeneira, M. M., and Jetz, W. (2014). EltonTraits  
1844 1.0: Species-level foraging attributes of the world's birds and mammals. *Ecology*. DOI: 10.1890/13–  
1845 1917.1.
- 1846 Wong, M. K., Guénard, B., and Lewis, O. T. (2018). Trait-based ecology of terrestrial arthropods. *Biological  
1847 Reviews*. DOI: 10.1111/brv.12488.
- 1848 Young, H. S., McCauley, D. J., Galetti, M., and Dirzo, R. (2016). Patterns, Causes, and Consequences of  
1849 Anthropocene Defaunation. *Annual Review of Ecology, Evolution, and Systematics*, 47. December, 333–  
1850 358. DOI: 10.1146/annurev-ecolsys-112414-054142.
- 1851 Zamudio, K. R., Bell, R. C., and Mason, N. A. (2016). Phenotypes in phylogeography: Species' traits, en-  
1852 vironmental variation, and vertebrate diversification. *Proceedings of the National Academy of Sciences*,  
1853 113.29, 8041–8048. DOI: 10.1073/pnas.1602237113.
- 1854 Zhang, J., Qian, H., Girardello, M., Pellissier, V., Nielsen, S. E., and Svenning, J. C. (2018a). Trophic inter-  
1855 actions among vertebrate guilds and plants shape global patterns in species diversity. *Proceedings of the  
1856 Royal Society B: Biological Sciences*. DOI: 10.1098/rspb.2018.0949.
- 1857 Zhang, Y., Yang, K., Yang, P., Su, Y., Zheng, W., and Liu, J. (2018b). Food restriction decreases BMR, body  
1858 and organ mass, and cellular energetics, in the Chinese Bulbul (*Pycnonotus sinensis*). *Avian Research*,  
1859 9.1, 1–11. DOI: 10.1186/s40657-018-0131-8.
- 1860 Zuñiga-Palacios, J., Zuria, I., Castellanos, I., Lara, C., and Sánchez-Rojas, G. (2021). What do we know (and  
1861 need to know) about the role of urban habitats as ecological traps? Systematic review and meta-analysis.  
1862 *Science of the Total Environment*, 780. DOI: 10.1016/j.scitotenv.2021.146559.

# **Appendices**

## **Appendix 0: Supporting information for Chapter 1**

# Appendix 1: Supporting information for Chapter 2

## S2.1 Taxonomic corrections

Across the different sources, similar species could appear under different binomial names. This was a problem when matching datasets by species. Moreover, it is possible that within a source, a given species was appearing under two or more different, synonymous names. As such, taxonomic synonymy created duplicated rows for the same species, overall falsely increasing the total number of species and potentially inflating the number of missing trait values. Taxonomic synonymy was hence a major issue. Due to the large number of species across datasets, extensive manual checks could not be applied. The presence of typos in species names had the same effect as synonymy, erroneously duplicating species. We attempted to correct for taxonomy first by correcting for typos, and second by identifying species which were entered under non-accepted names and replacing these with the accepted name. To this end, we developed an automated procedure, complemented with a few manual entries where errors were opportunistically spotted. Such errors in taxonomy were notably spotted when attempting to retrieve trait data for subsets of species, for analyses not directly related to the work conducted here. Taxonomic synonymy was as such checked manually for 91 species (56 birds, 7 mammals and 28 reptiles); in that case, information was extracted from other diverse sources (such as the Reptile Database (<http://www.reptile-database.org/>); Avibase (<https://avibase.bsc-eoc.org/avibase.jsp?lang=EN&pg=home>); AmphibiaWeb (<https://amphibiaweb.org/>); and additional manual checks using the IUCN Red List for mammals). A column in the Synonym dataset mentions where manual checks were applied (in which case the Synonym dataset was manually corrected).

### Automated procedure and outputs.

#### Extracting names from the IUCN Red List and the Integrated Taxonomic Information System (ITIS).

The objectives of the automated procedure were to (1) extract species synonymous binomial names from the

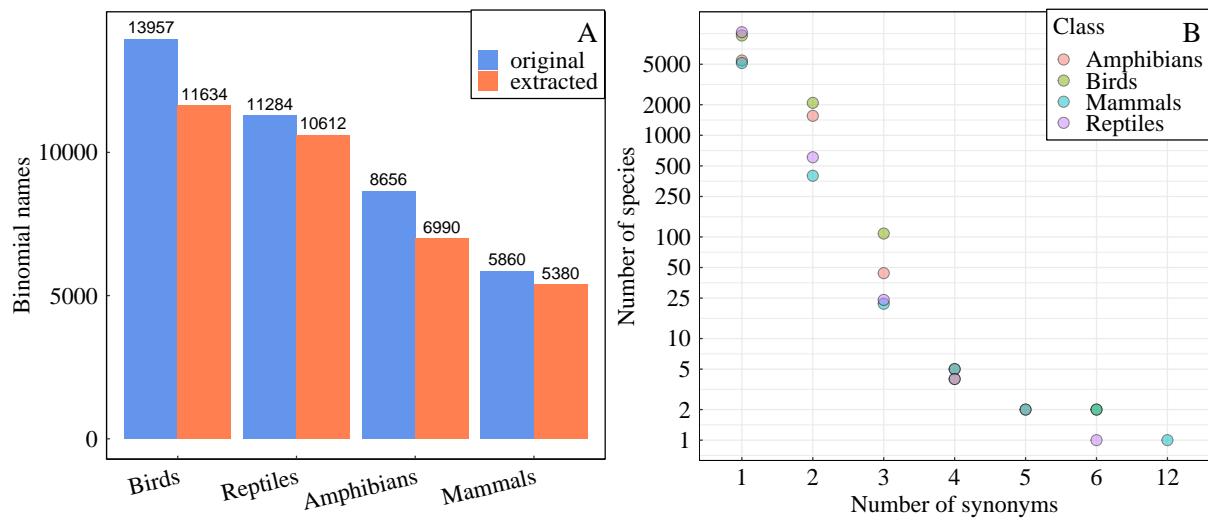
IUCN Red List or from ITIS, using the rredlist (Chamberlain, 2018) and taxize (Chamberlain and Szöcs, 2013) R packages; and (2) identify the status of each name (accepted or not accepted). We started by generating a list of all names featuring in any of the sources. These ‘original’ names were corrected for typos (using gnr\_resolve function in the taxize package). Then, the IUCN Red List was queried and any listed synonyms were stored, as well as the status of each synonym (accepted or not accepted). When species were not found in the IUCN Red List, synonyms were extracted from ITIS. When species were not found in ITIS either, corrected names (original names corrected for typos) were used. Family and order designations were extracted using the same procedure and some entries were retrieved from the Global Biodiversity Information Facility taxonomic backbone when not available in the Red List or in ITIS (GBIF, <https://www.gbif.org/tools/species-lookup>).

**NB:** for species entered with the forms *Genus cf.*, *Genus aff.* or *Genus spp.*, the accepted binomial name was left empty.

**Output.** We generated a ‘Synonym’ dataset containing records of binomial names (14,124 recorded names for birds; 8,743 for amphibians; 6,090 for mammals; and 11,678 for reptiles), and for each we recorded their status and their potential synonyms.

**Harmonising taxonomy in trait datasets.** Taxonomy across datasets was finally homogenised by replacing synonyms with a uniquely identified accepted name. As a consequence, the total number of identified unique species decreased (Figure S2.1). The species presenting the highest number of synonyms was the East African mole rat (*Tachyoryctes splendens*), for which we found 12 synonymous names (Figure S2.1B).

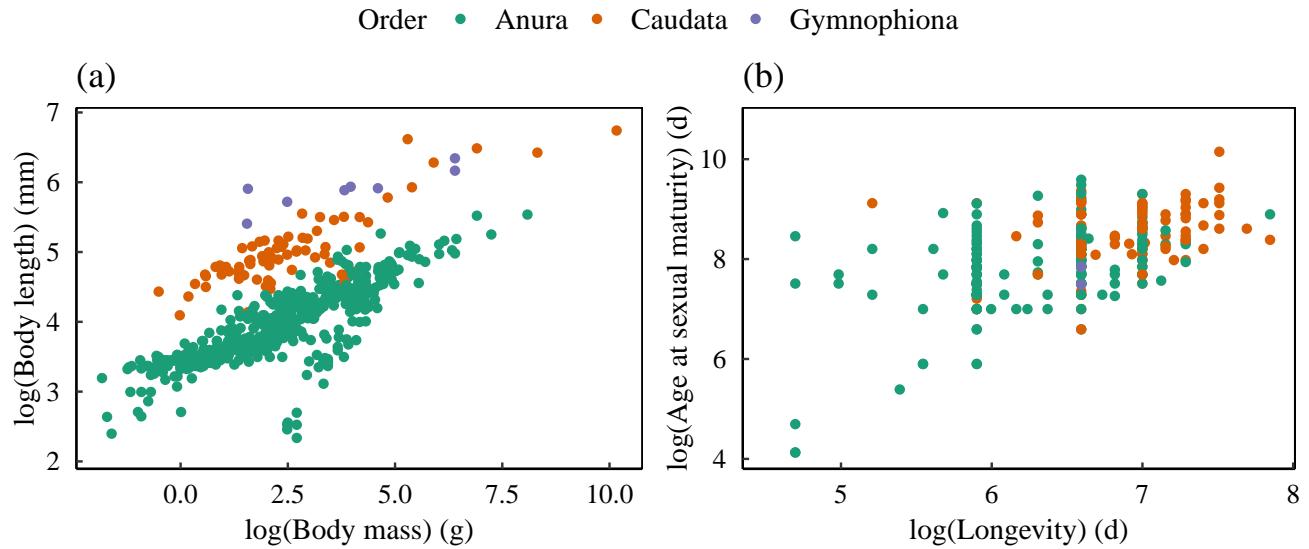
The automated procedure was not perfect, and taxonomic errors are likely to have persisted in the trait datasets. The Red List and the ITIS were not comprehensive taxonomic sources, and for clades with high degrees of synonymy in names, such as reptiles or amphibians, neither the Red List or the ITIS contained enough information. Taxonomy may be further improved by using class-specific sources in an automated procedure.



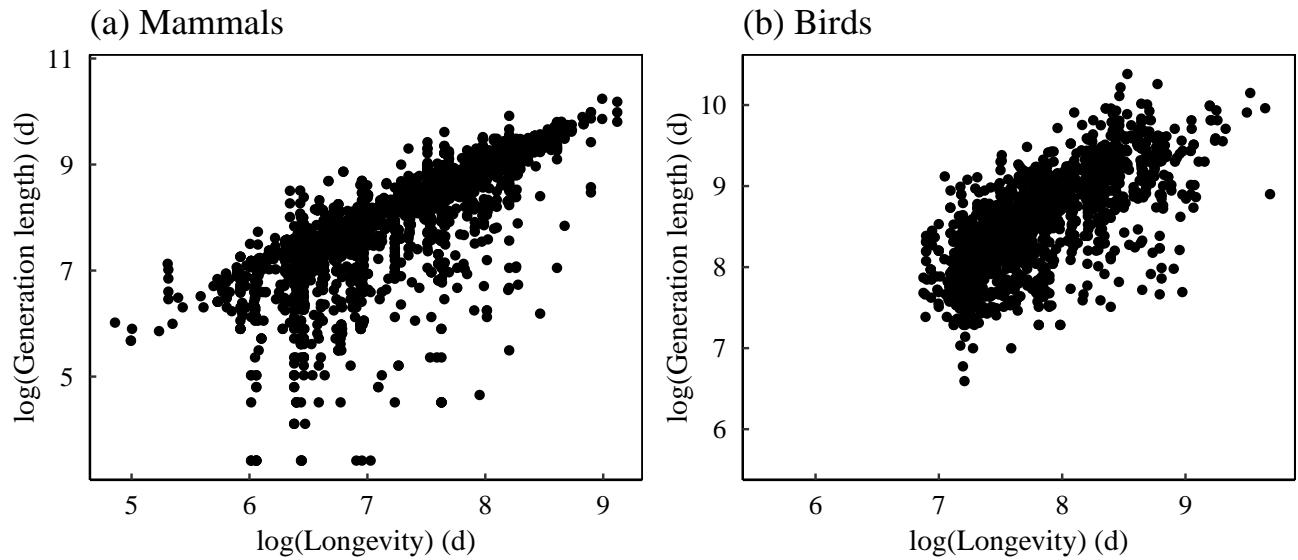
**Figure S2.1: Difference in species number before and after taxonomic correction (A) and distribution of number of synonyms across datasets (B).** (A) shows the number of species (binomial names) extracted from all sources (blue bars), and the number of uniquely identified accepted names (in red). Replacing non-accepted synonyms by one identified accepted name reduced the number of species in all classes, with the largest reduction for birds. (B) shows the distribution of the number of synonymous names. In all four classes, more than 5,000 species were known under one name only. Nevertheless, a large number of species had two identified synonyms (range: 400 species for mammals - 2086 for birds). The most potentially replicated species was the East African mole rat *Tachyoryctes splendens*, for which 12 synonyms were identified.

## S2.2 Additional information for trait compilation

### Correlations among closely related traits

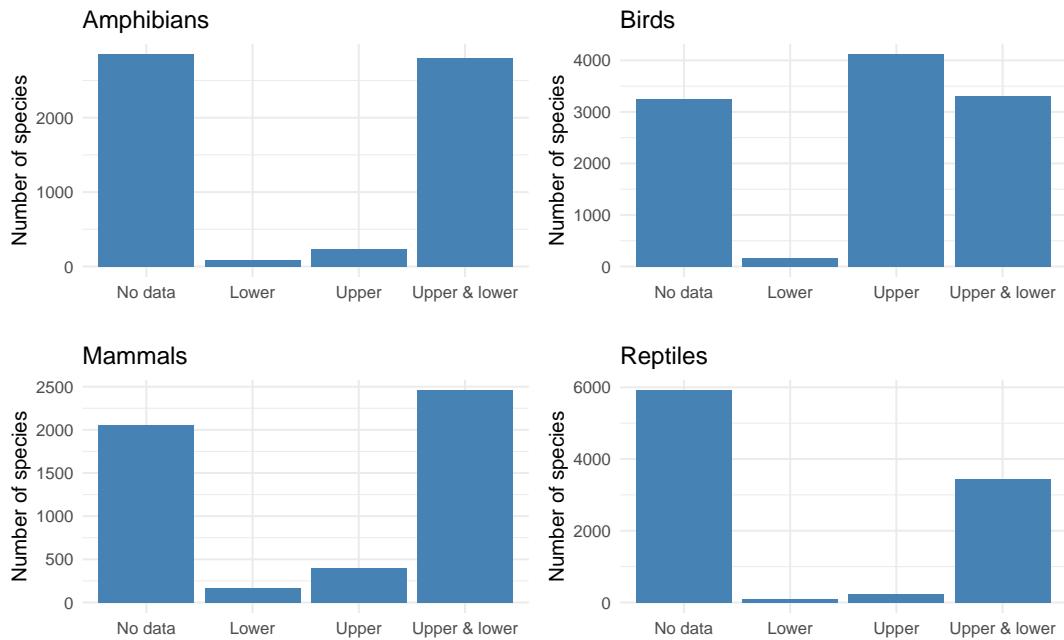


**Figure S2.2:** (a) Body mass versus body length and (b) longevity versus age at sexual maturity in amphibians. The Pearson's correlation coefficient was 0.71 in (a) and 0.55 in (b) (order was not included in these coefficients).

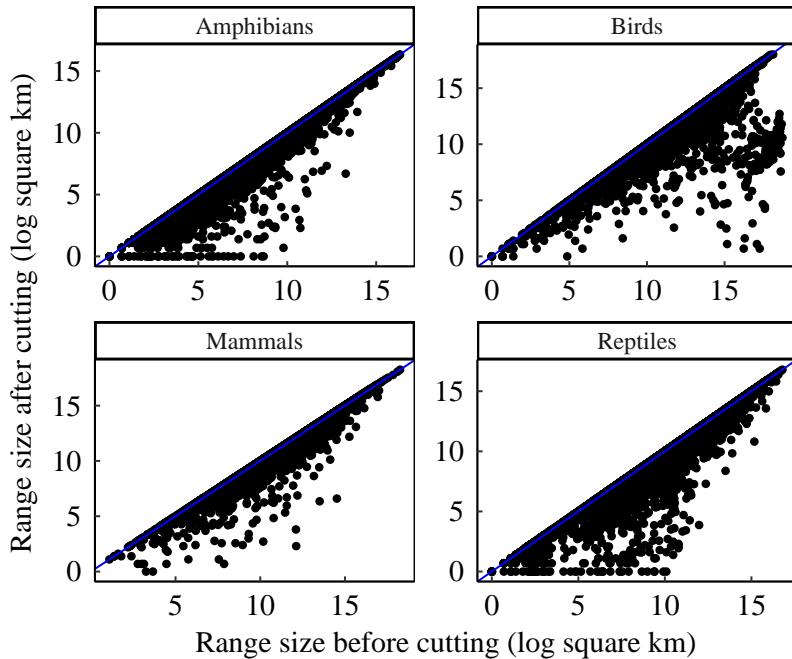


**Figure S2.3:** Generation length versus longevity data in mammals and birds. The Pearson's correlation coefficient was 0.74 in (a) and 0.70 in (b).

### S2.3 Cutting distribution maps by altitudinal limits

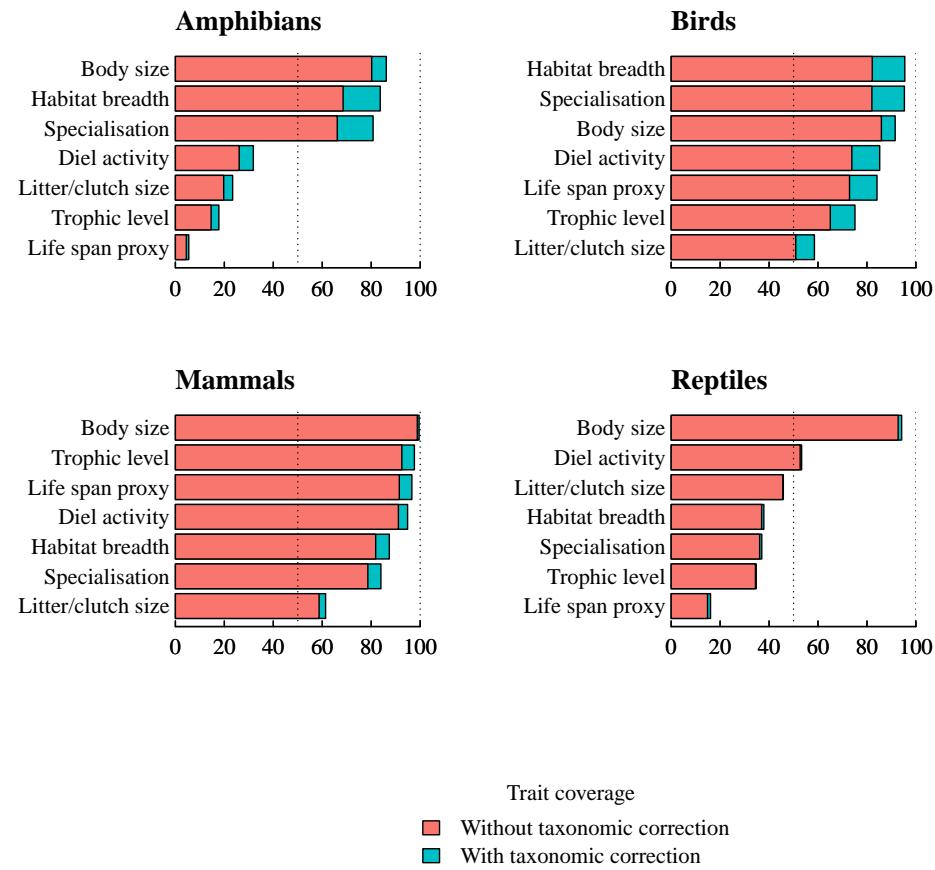


**Figure S2.4: Availability of altitudinal limits across species.** Upper and lower altitudinal limits were extracted from the IUCN Red List.



**Figure S2.5: Range sizes before versus after cutting by altitudinal limits.**

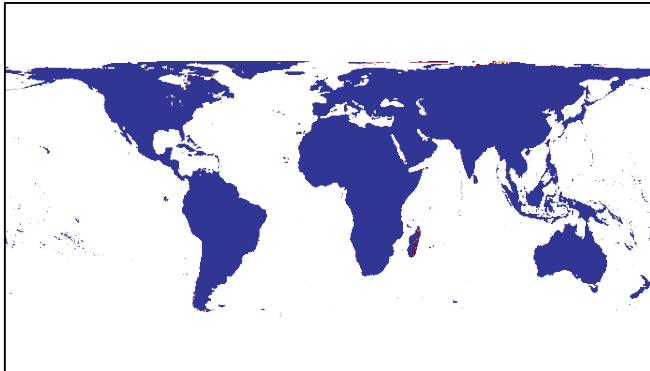
## S2.4 Impact of taxonomic corrections on trait coverage



**Figure S2.6:** Trait coverage when no taxonomic correction is applied before matching sources by binomial names, versus when we applied the described procedure. Identification of synonyms allowed to increase trait coverage in most cases.

## **S2.5 Assemblage-level median, mean and standard deviation of trait completeness (maps)**

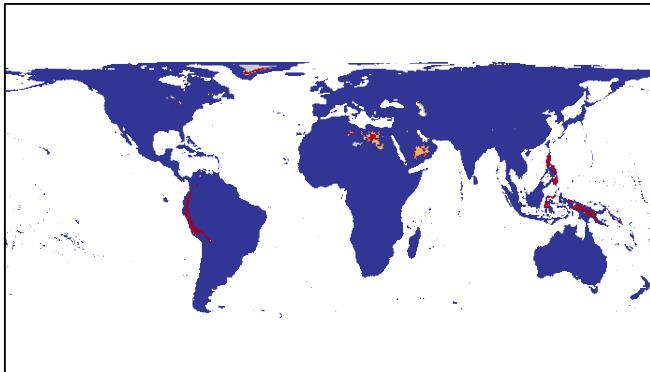
**(a) Mammals**



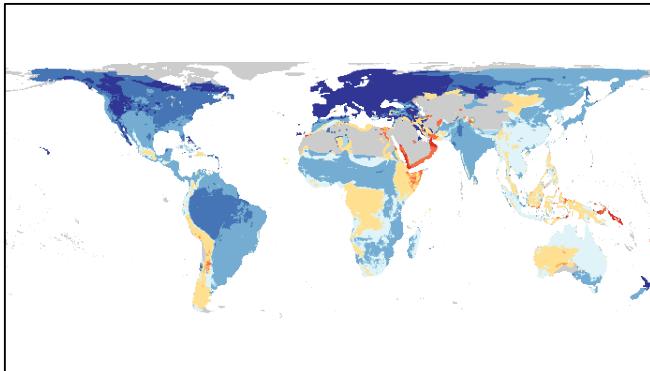
Median completeness  
(birds and mammals):

- NA
- 80 – 90%
- 90 – 91%
- 91 – 92%
- 92 – 93%
- 93 – 94%
- 94 – 95%
- 95 – 96%
- 96 – 97%
- 97 – 98%
- 98 – 99%
- 99 – 100%

**(b) Birds**



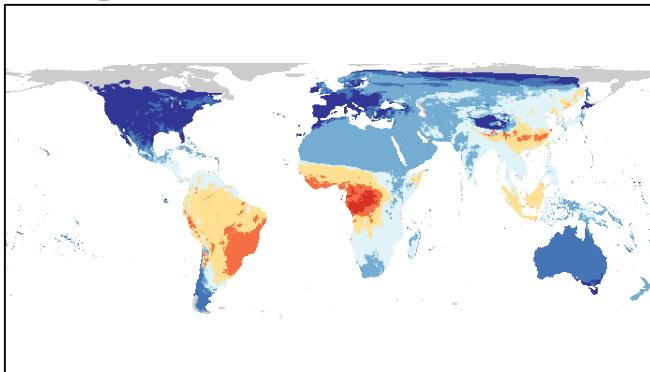
**(c) Amphibians**



Median completeness  
(herptiles):

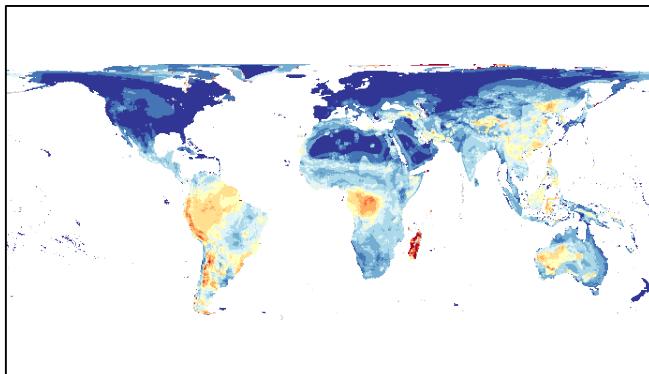
- NA
- 0 – 10%
- 10 – 20%
- 20 – 30%
- 30 – 40%
- 40 – 50%
- 50 – 60%
- 60 – 70%
- 70 – 80%
- 80 – 90%
- 90 – 100%

**(d) Reptiles**



**Figure S2.7: Spatial distribution of assemblage-level median trait completeness in each class.** Note that the color breaks differ for mammals and birds and for herptiles.

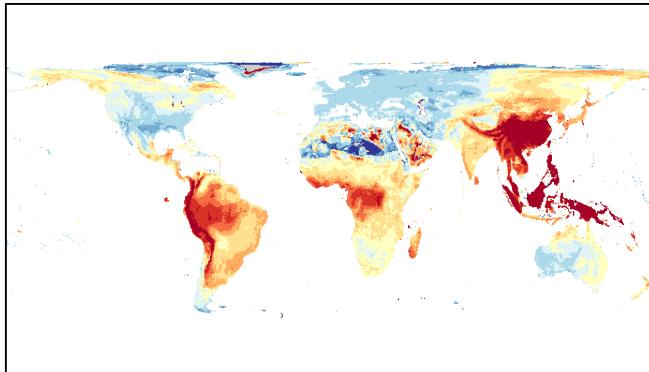
**(a) Mammals**



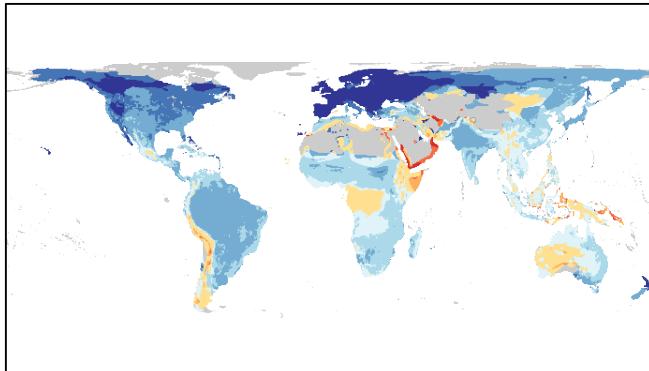
Mean completeness  
(birds and mammals):

- NA
- 80 – 90%
- 90 – 91%
- 91 – 92%
- 92 – 93%
- 93 – 94%
- 94 – 95%
- 95 – 96%
- 96 – 97%
- 97 – 98%
- 98 – 99%
- 99 – 100%

**(b) Birds**



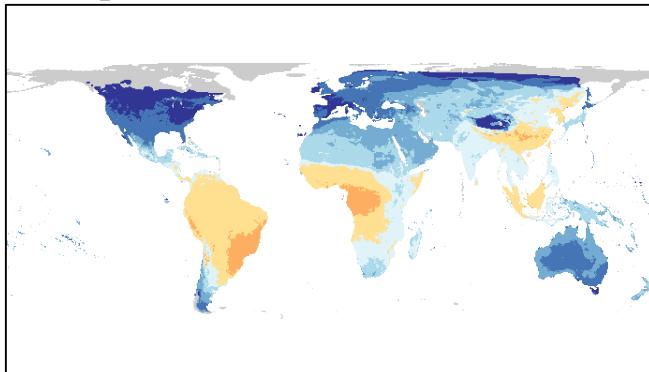
**(c) Amphibians**



Mean completeness  
(herptiles):

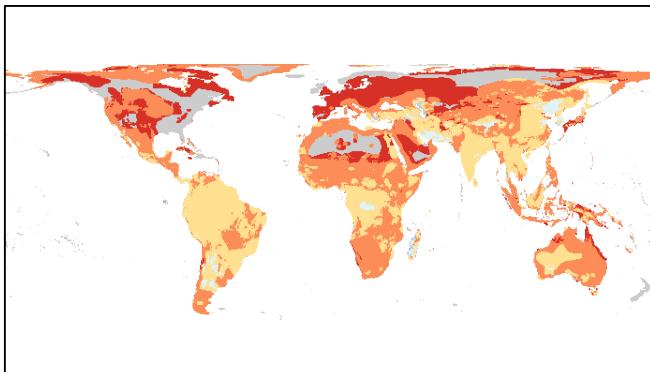
- NA
- 0 – 10%
- 10 – 20%
- 20 – 30%
- 30 – 40%
- 40 – 50%
- 50 – 60%
- 60 – 70%
- 70 – 80%
- 80 – 90%
- 90 – 100%

**(d) Reptiles**



**Figure S2.8:** Spatial distribution of assemblage-level mean trait completeness in each class. Note that the color breaks differ for mammals and birds and for herptiles.

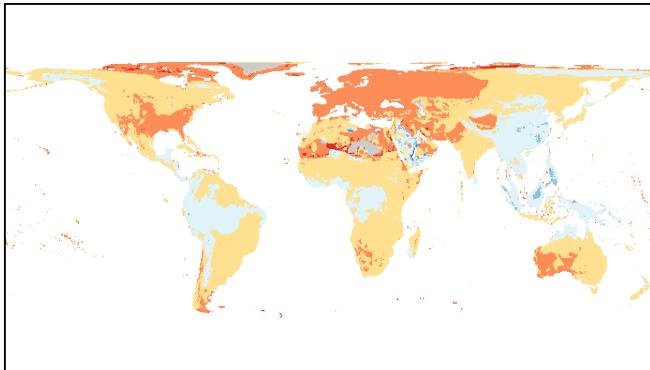
**(a) Mammals**



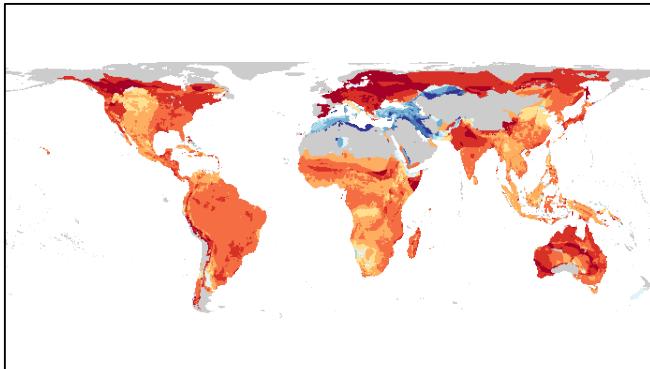
Standard deviation  
of completeness  
(birds and mammals):

- NA
- 0 – 5%
- 5 – 10%
- 10 – 15%
- 15 – 20%
- 20 – 25%
- 25 – 30%

**(b) Birds**



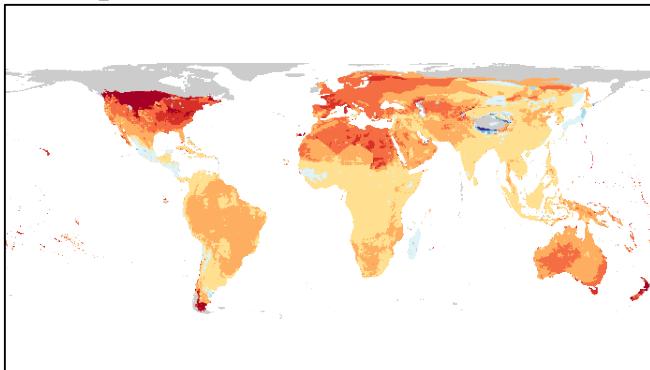
**(c) Amphibians**



Standard deviation  
of completeness  
(herptiles):

- NA
- 0 – 10%
- 10 – 15%
- 15 – 20%
- 20 – 25%
- 25 – 30%
- 30 – 35%
- 35 – 40%
- 40 – 50%
- 50 – 60%
- 60 – 75%

**(d) Reptiles**



**Figure S2.9: Spatial distribution of assemblage-level standard deviation of trait completeness in each class.** Note that the color breaks differ for mammals and birds and for herptiles.

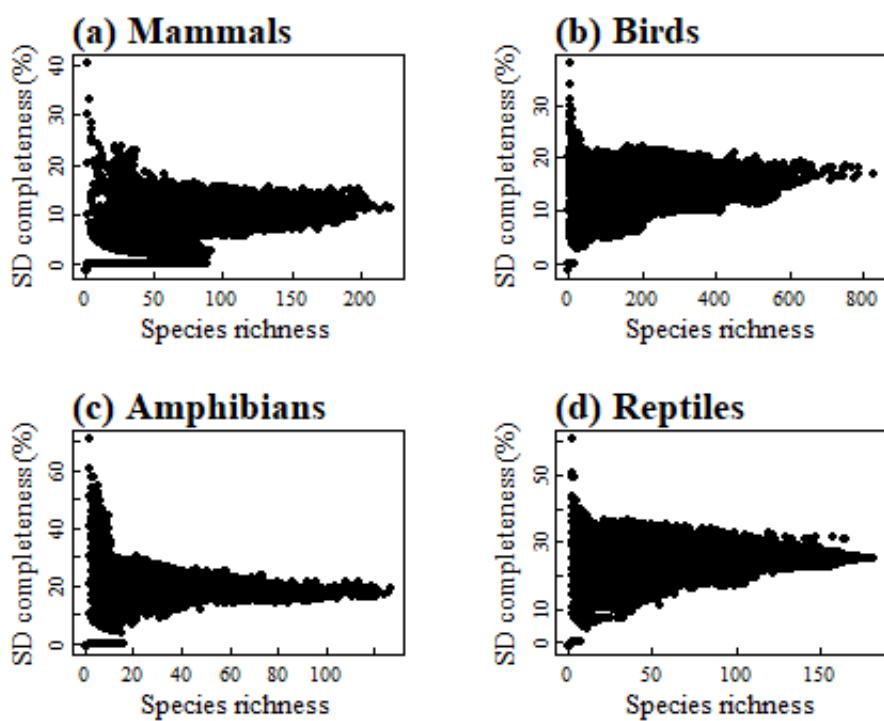
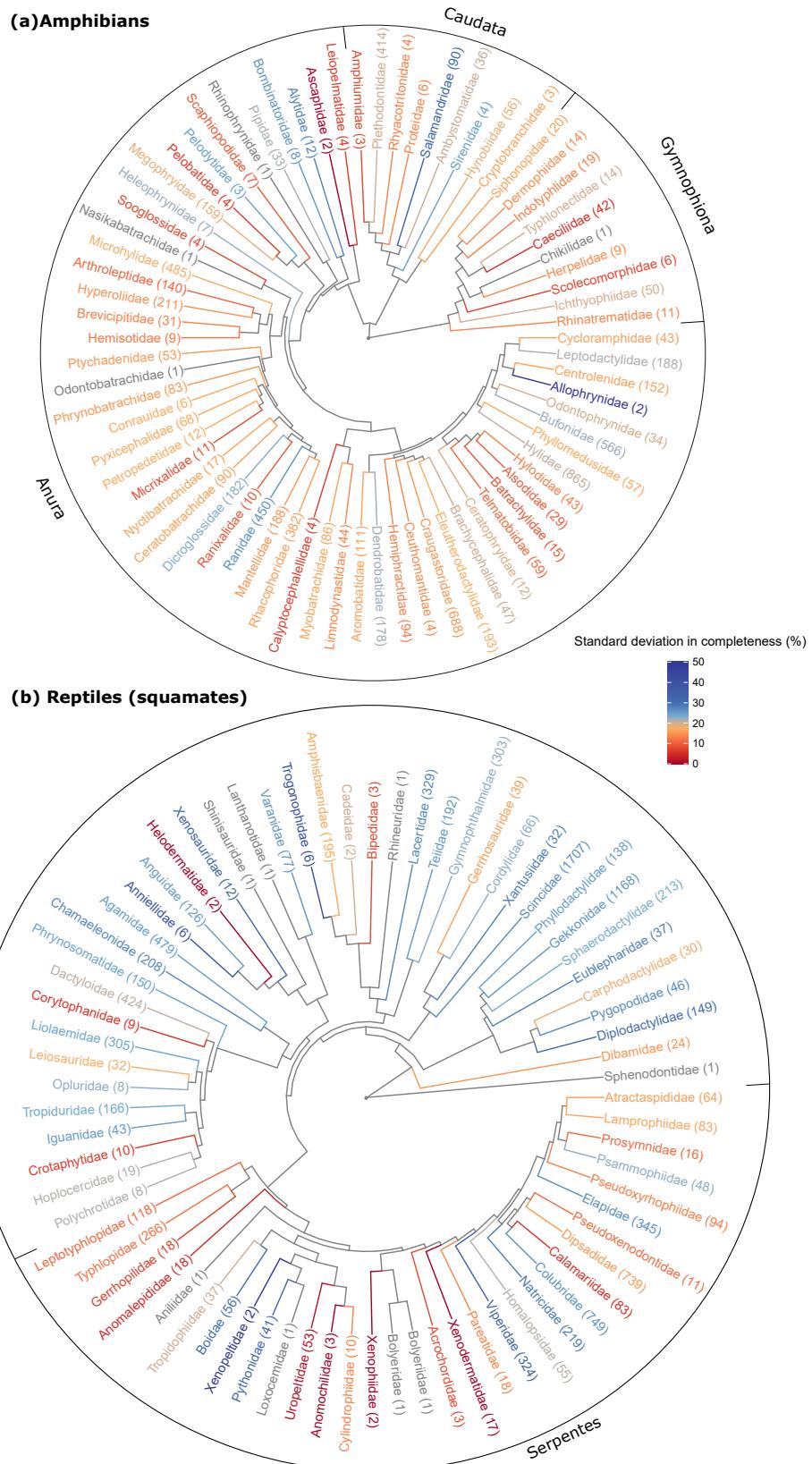


Figure S2.10: Assemblage-level species richness against standard deviation in completeness in each class.

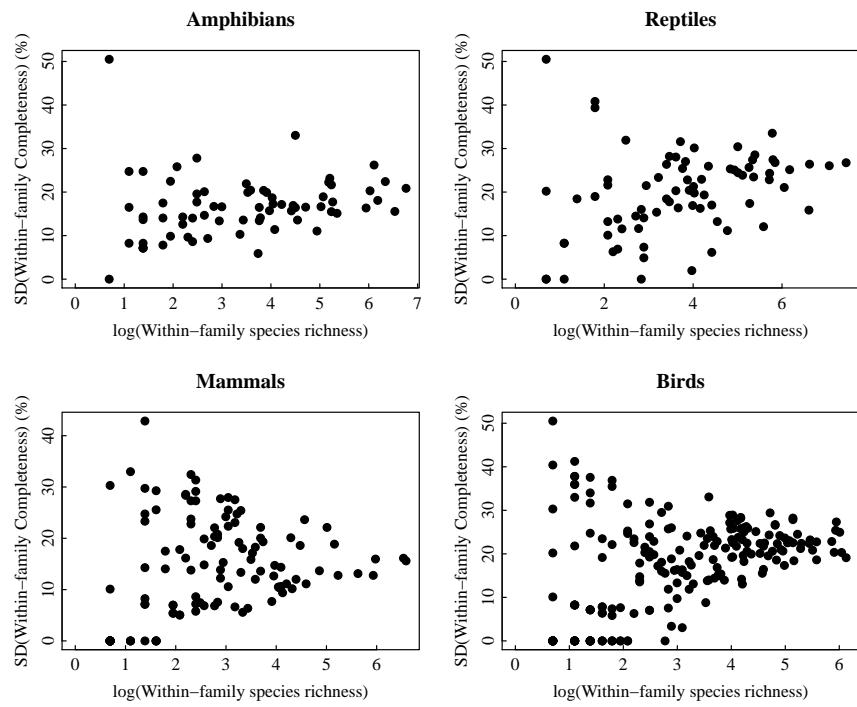
## S2.6 Phylogenetic patterns in trait completeness







**Figure S2.13: Within-family standard deviation in completeness (herptiles).**



**Figure S2.14: Within-family species richness against the within-family standard deviation of completeness.**

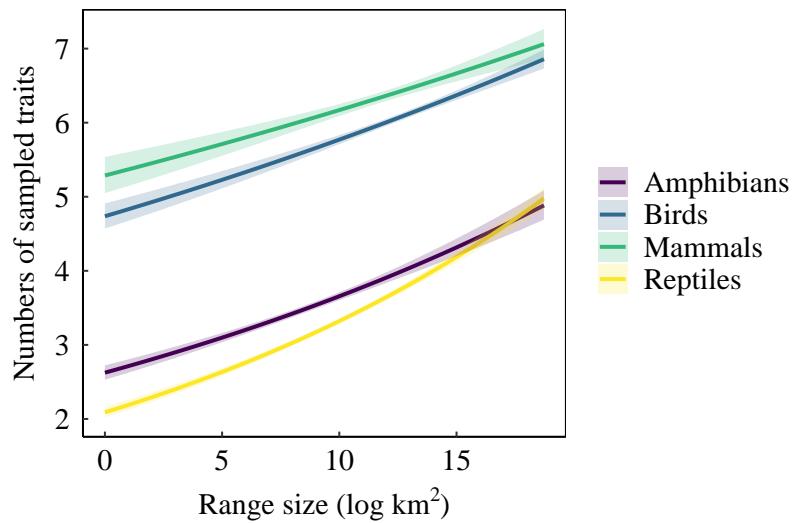
## S2.7 Model coefficients for Range size against Number of sampled traits (Poisson model).

**Table S2.1:** Coefficients of the model investigating whether species range size explained the number of sampled traits. Class was added as an interacting predictor. The reference level for class is mammals. The model was fitted using Poisson error distribution.

	Estimate	Std. Error	z value	Pr(> z )
Intercept	1.678	0.022	76.809	< 2e - 16
log Range Size	0.015	0.002	8.086	6.16e - 16
Class Birds	-0.092	0.028	-3.350	0.000809
Class Amphibians	-0.689	0.029	-24.099	< 2e - 16
Class Reptiles	-0.872	0.027	-31.856	< 2e - 16
log Range Size:Class Birds	0.003	0.002	1.415	0.157
log Range Size:Class Amphibians	0.017	0.003	6.427	1.30e - 10
log Range Size:Class Reptiles	0.026	0.002	11.159	< 2e - 16

**Table S2.2:** Coefficients of the model investigating whether species range size explained the number of sampled traits, *using range maps not cut by altitudinal limits*. Class was added as an interacting predictor. The reference level for class is mammals. The model was fitted using Poisson error distribution.

	Estimate	Std. Error	z value	Pr(> z )
Intercept	1.665	0.023	72.070	< 2e - 16
log Range Size	0.015	0.002	8.167	3.16e - 16
Class Birds	-0.110	0.029	-3.763	0.0002
Class Amphibians	-0.700	0.030	-23.721	< 2e - 16
Class Reptiles	-0.928	0.029	-32.403	< 2e - 16
log Range Size:Class Birds	0.004	0.002	1.840	0.066
log Range Size:Class Amphibians	0.018	0.003	6.564	5.24e - 11
log Range Size:Class Reptiles	0.031	0.002	12.630	< 2e - 16



**Figure S2.15: Relationship between number of sampled traits and geographical range size using distribution maps not cut by altitudinal limits.** Models were fitted using a Poisson error distribution. Class was added as a predictor interacting with range size. Rates of increase were not significantly different for mammals and birds, but differed for reptiles and amphibians, with steeper rates of increase for reptiles overall. Cutting range maps by altitudinal limits had little effects on the results (see Figure 2.5 in Main text).

## S2.8 Spatial models summaries

**Table S2.3: Spatial model summary for amphibians.** The spatial model was fitted to explain assemblage-level median completeness with species richness. Biogeographic realm was added as an interacting explanatory variable.

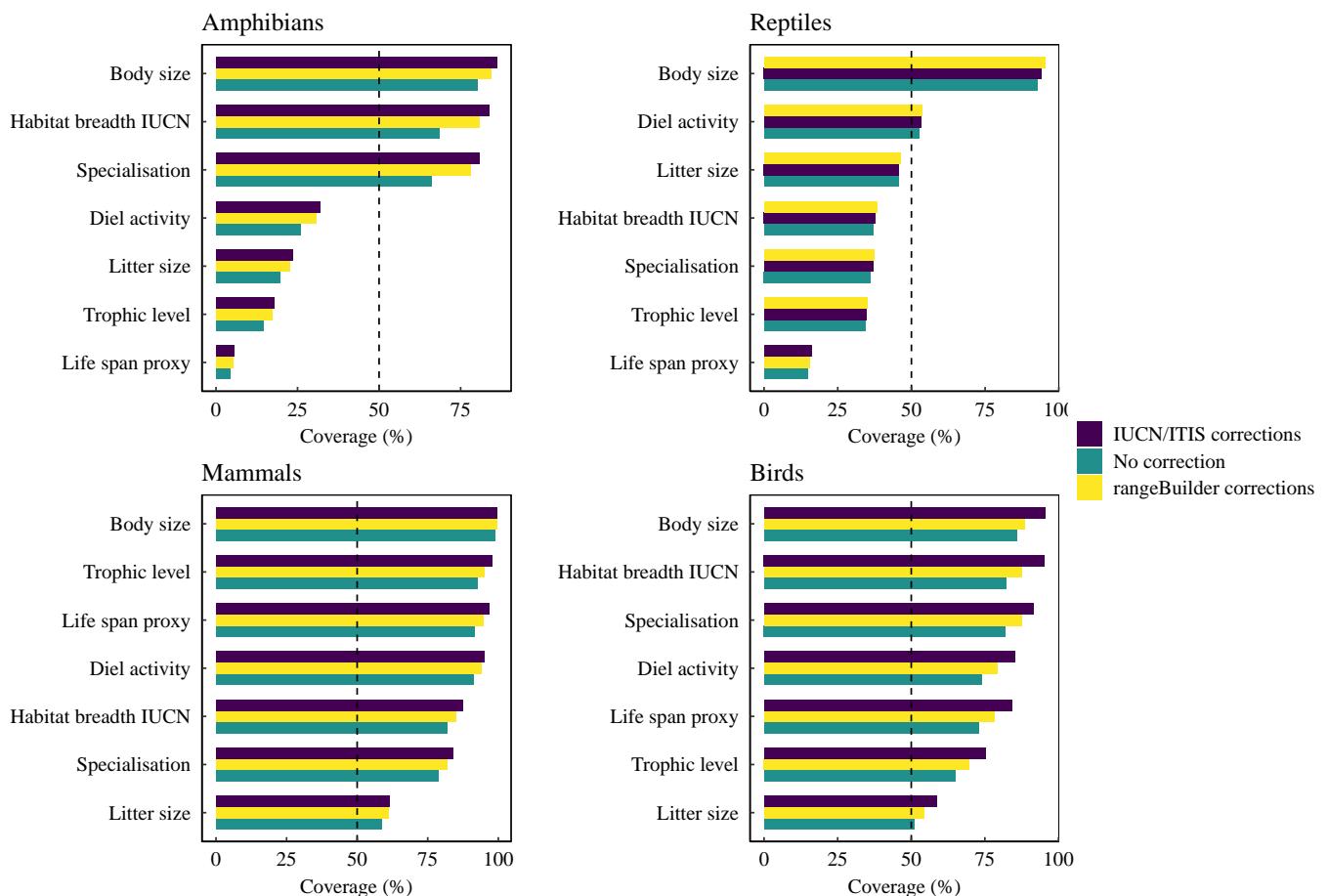
	Estimate	Std. Error	z value	Pr(> z )
Intercept - Realm: Afrotropic	0.0738	0.0064	11.4908	0
log(Species richness)	-0.0025	0.0017	-1.4261	0.1538
Realm: Australasia	-0.0109	0.0095	-1.1453	0.2521
Realm: Indo-Malay	0.0455	0.0119	3.8294	0.0001
Realm: Nearctic	0.0441	0.0082	5.3905	0.000000
Realm: Neotropic	-0.0377	0.0083	-4.5538	0.00001
Realm: Palearctic	0.0047	0.0067	0.6992	0.4844
log(Species richness):Australasia	0.0018	0.0038	0.4789	0.6320
log(Species richness):Indo-Malay	-0.0147	0.0039	-3.7294	0.0002
log(Species richness):Nearctic	-0.0097	0.0030	-3.2003	0.0014
log(Species richness):Neotropic	0.0144	0.0026	5.6454	0.000000
log(Species richness):Palearctic	0.0109	0.0029	3.7358	0.0002

**Table S2.4: Spatial model summary for reptiles.** The spatial model was fitted to explain assemblage-level median completeness with species richness. Biogeographic realm was added as an interacting explanatory variable.

	Estimate	Std. Error	z value	Pr(> z )
Intercept - Realm: Afrotropic	0.2001	0.0144	13.9349	0
log(Species richness)	-0.0316	0.0031	-10.0547	0
Realm: Australasia	-0.1284	0.0189	-6.7851	0
Realm: Indo-Malay	-0.0453	0.0263	-1.7215	0.0852
Realm: Nearctic	-0.0788	0.0140	-5.6366	0.000000
Realm: Neotropic	-0.0932	0.0145	-6.4425	0
Realm: Palearctic	-0.1030	0.0131	-7.8787	0
log(Species richness):Australasia	0.0386	0.0046	8.4019	0
log(Species richness):Indo-Malay	0.0124	0.0061	2.0397	0.0414
log(Species richness):Nearctic	0.0346	0.0038	9.1601	0
log(Species richness):Neotropic	0.0220	0.0034	6.4231	0
log(Species richness):Palearctic	0.0286	0.0033	8.6153	0

## S2.9 Trait coverage and taxonomic matching

Here, we briefly explore the robustness of our work to taxonomic uncertainty by comparing trait coverage obtained with our procedure for taxonomic matching against trait coverage obtained when extracting synonyms from class-specific sources, which could contain more information, notably for herptiles. We aligned taxonomy again using the rangeBuilder R package (Rabosky et al., 2016), which allows the extraction of accepted names from class-specific sources. Overall, our results are robust to the use of different taxonomic backbones; the main conclusions are likely to be unaffected by taxonomic uncertainty.



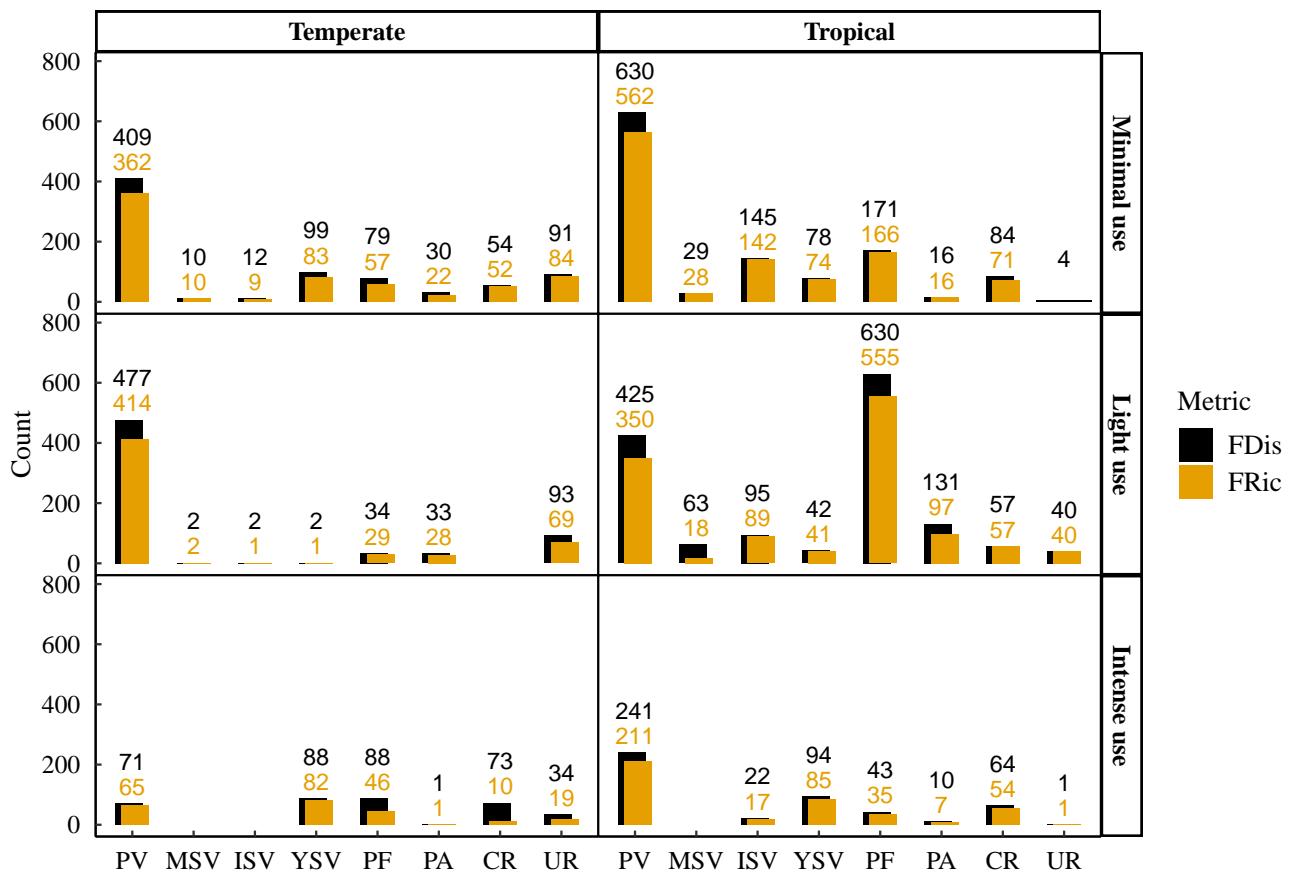
**Figure S2.16: Comparison of trait coverage among datasets corrected for taxonomy in different ways:** using the described procedure (purple bars), datasets corrected using the rangeBuilder package (extraction of synonyms from class-specific sources, yellow bars; Rabosky et al. (2016)) and datasets where no taxonomic correction was applied when matching sources (green bars).

# Appendix 2: Supporting information for Chapter 3

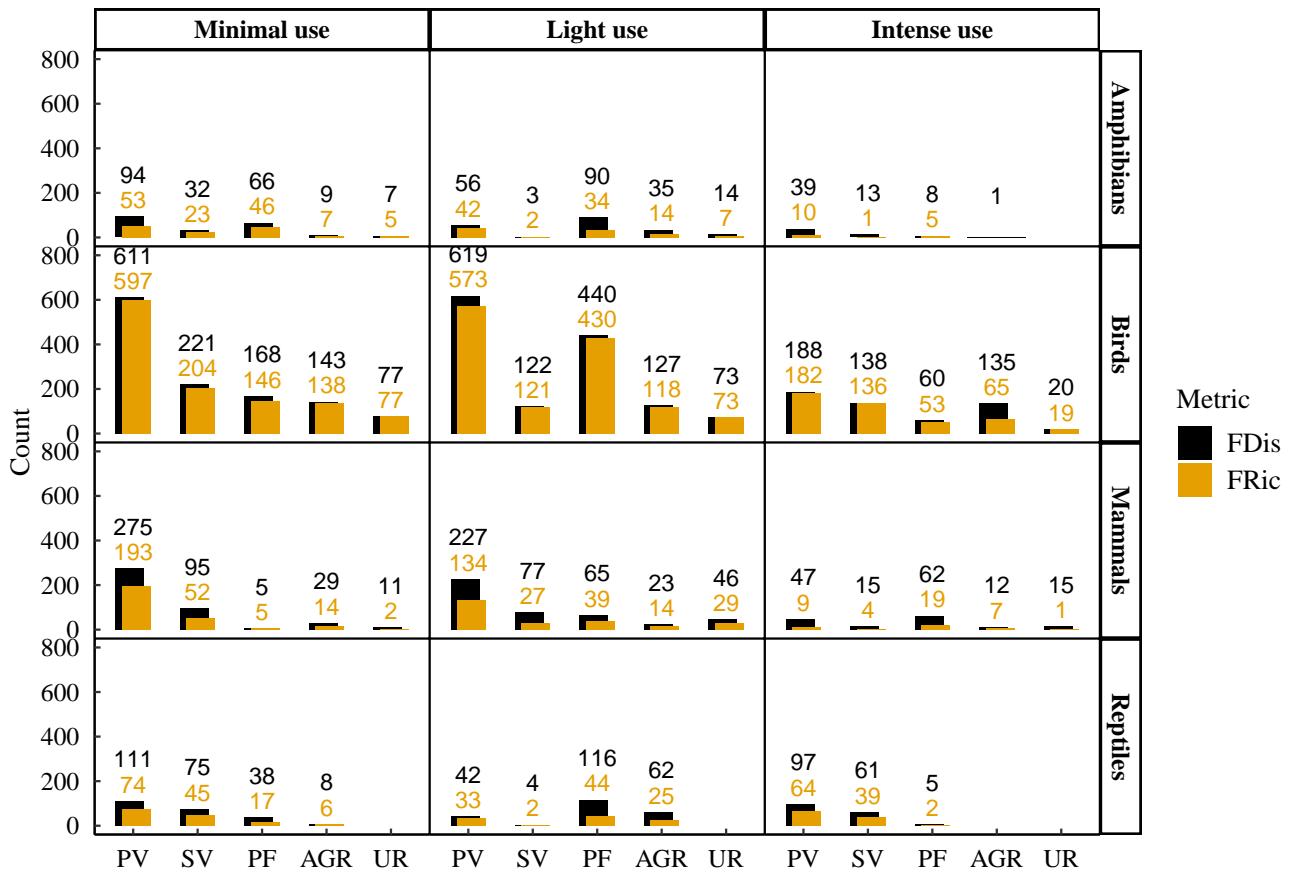
## S3.1 Land-use categories in PREDICTS and sample sizes

**Table S3.1: Land-use categories in the PREDICTS database.** See Hudson et al. (2014, 2017) for more details.

Land-use category	Definition
Primary vegetation	Native vegetation, undisturbed since its establishment under current climatic conditions. No known alterations due to human activities or to extreme natural events.
Mature secondary vegetation	Vegetation recovering after complete destruction of primary vegetation & where succession is near complete – the structure approaches that of primary vegetation.
Intermediate secondary vegetation	Vegetation recovering after complete destruction of primary vegetation at a mid-successional stage.
Young secondary vegetation	Vegetation recovering after complete destruction of primary vegetation at an early successional stage.
Plantation forest	Previously cleared areas planted with crop trees or shrubs grown and harvested for human consumption or for commercial purposes (includes wood, fruit, oil, biofuel, rubber, etc).
Pasture	Areas grazed by livestock, permanently or regularly. Can be improved through cultivation techniques.
Cropland	Previously cleared areas planted with herbaceous crops and harvested for human or animal consumption (including animal feed and crops used in the food industry), or for commercial purposes (e.g., crops grown for the textile industry).
Urban	Previously cleared areas built up by humans. Vegetation is managed for civic or personal purposes.



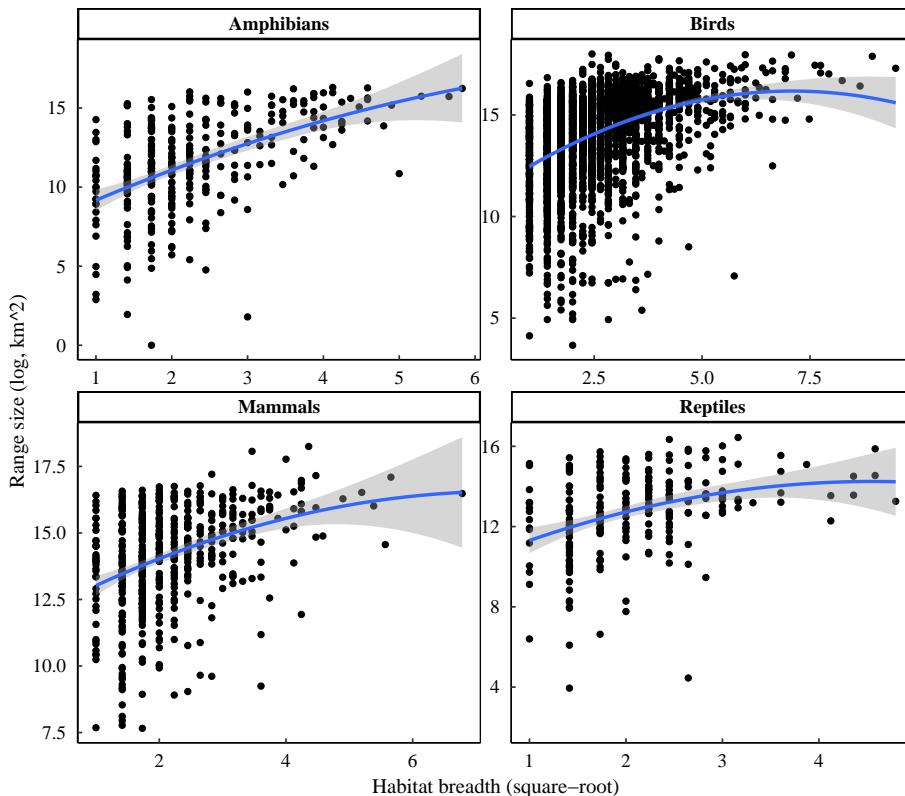
**Figure S3.1: Number of sites in each land use and land-use intensity for which FRic and FDis were calculated, across all species.** The number of sites for FRic can be smaller than the number of sites for FDis because FRic could not necessarily be computed for all assemblages in which FDis was estimated (in assemblages where species richness was 2 or 3, FRic could not be computed).



**Figure S3.2: Number of sites in each land use, land-use intensity and for which FRic and FDis were calculated, within each class.** The number of sites for FRic can be smaller than the number of sites for FDis because FRic could not necessarily be computed for all assemblages in which FDis was estimated (in assemblages where species richness was 2 or 3, FRic could not be computed).

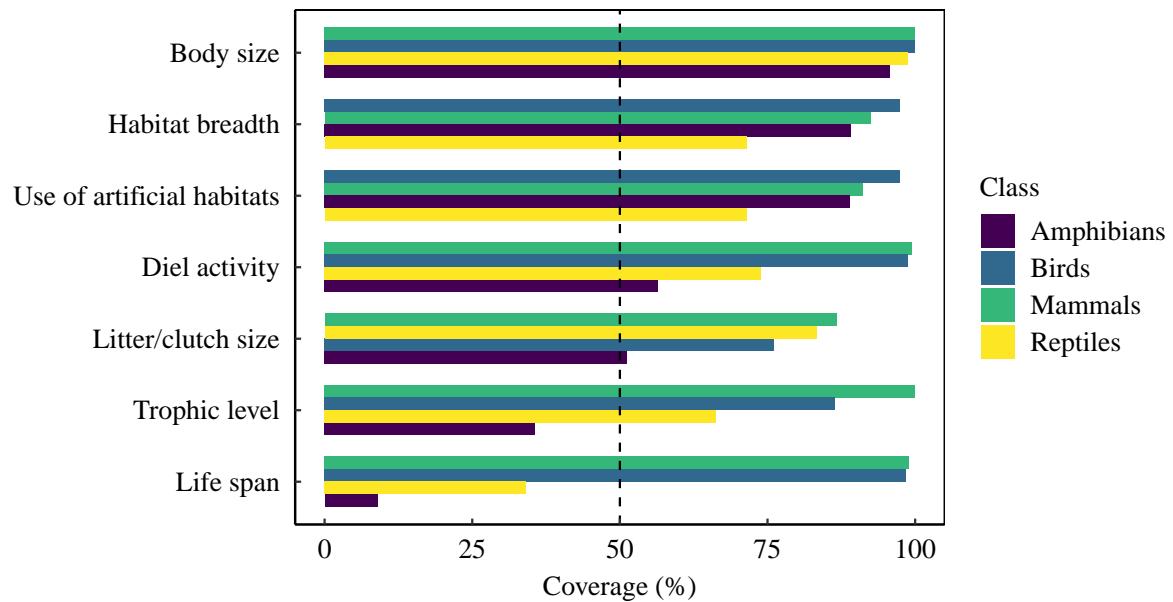
## S3.2 Trait data & imputation of missing trait values

I used the compilation of trait data across terrestrial vertebrates from Chapter 2 (and published in Etard et al. (2020)). I selected seven traits that were available at least for a subset of the species in each class: body mass, trophic level, lifespan, litter/clutch size, diel activity, habitat breadth and use of artificial habitats. These traits related to life-history, habitat specialisation and use of geographical space (for instance, habitat breadth is a significant predictor of geographical range size in all classes, Figure S3.3).

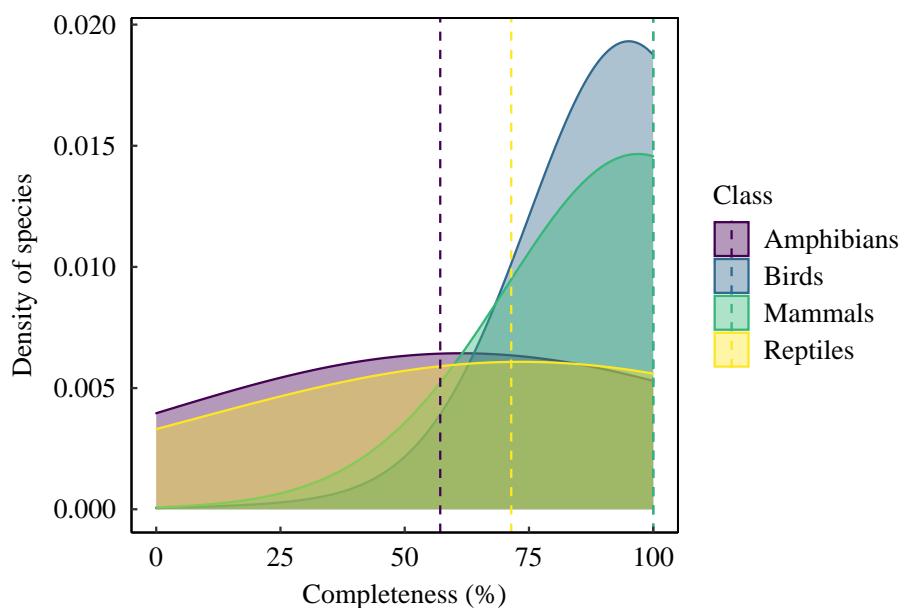


**Figure S3.3: Relationship between habitat breadth and geographical range size across species in each class.** The derivation of geographical range sizes is described in Chapter 2.

Trait-data coverage was highly variable among classes and traits, with important geographical and phylogenetic biases in trait data for reptiles and amphibians (Chapter 2; Etard et al., 2020; Figure S3.4, Figure S3.5). To obtain complete species-trait datasets, I imputed missing trait values. Further, in order to assess the sensitivity of our models to variation in imputed values, I imputed the missing trait values eight independent times. This allowed me to assess the congruence of the model predictions when using the different imputed trait datasets in the analyses.



**Figure S3.4: Trait coverage for the vertebrate species sampled in the PREDICTS database.** For a given trait, coverage is calculated as the percentage of species for which an estimate was available.



**Figure S3.5: Distribution of trait completeness across the vertebrate species sampled in the PREDICTS database.** For a given species, trait completeness is calculated as the proportion of traits for which an estimate was available. Dashed lines represent the median trait completeness.

### S3.2.1 Choice of imputation technique

There exist several imputation techniques (Debastiani et al., 2021; Etard et al., 2020; Johnson et al., 2021; Penone et al., 2014), such as K-nearest neighbour (Troyanskaya et al., 2001), multivariate imputation by chained equations (Buuren and Oudshoorn, 2007), random forest algorithms (implementable in R with missForest; Stekhoven, 2016; Stekhoven and Bühlmann, 2012), and phylogenetic imputations (implementable in R with PhyloPars, Bruggeman et al., 2009). Penone et al. (2014) assessed the performance of these four techniques and showed that missForest and PhyloPars performed better when traits were phylogenetically conserved, and when the species phylogenetic position was included as a predictor of missing trait values. PhyloPars can only handle continuous data, while missForest is compatible with mixed-type (including categorical) data. When no phylogenetic information was included, mice was found to be the best method, with fast imputations of mixed-type data (Penone et al., 2014). Therefore, to assess whether missForest or mice was more appropriate here, I measured the phylogenetic signal in trait data. For continuous traits, I used Pagel's  $\lambda$  (Pagel, 1999), and for categorical traits I used Borges'  $\delta$  (Borges et al., 2018). Strong phylogenetic signal would indicate that traits are phylogenetically conserved, and hence missForest would be the most suited approach for imputing missing trait values, with the inclusion of species' phylogenetic positions as a predictor.

### S3.2.2 Phylogenetic signal in traits

Across all classes, similar traits were used for calculating functional diversity metrics: body mass, litter/clutch size, lifespan (using different proxies in different vertebrate classes: generation length for birds and mammals, longevity for reptiles, and age at sexual maturity for amphibians), trophic level, diel activity, habitat breadth and use of artificial habitats. In addition, I included some class-specific traits for the imputations, as certain class-specific traits could be useful predictors of other traits (such as body length for instance in amphibians (Santini et al., 2018)). Table S3.2 details the traits that were included for the imputations in each class and the phylogenetic signal for each of these traits. Continuous traits were log-10 transformed before assessing Pagel's  $\lambda$  to improve normality. Pagel's  $\lambda$  was estimated using the phylosig function of the phytools package (Revell, 2012), and Borges'  $\delta$  was assessed using code provided by Borges et al. (2018), available at : [https://github.com/mrborges23/delta\\_statistic](https://github.com/mrborges23/delta_statistic). To test for the significance of  $\delta$ , I generated null distributions of  $\delta$  for each categorical trait by randomising trait vectors 50 times, and calculating  $\delta$  for each randomised vector – following the guidelines proposed by Borges et al. (2018). I then tested whether the observed medians were greater than the null distributions using one-sided Wilcoxon rank sum tests.

I used class-specific phylogenies to estimate phylogenetic signal, all downloaded on 13th April 2020. Trees from Faurby et al. (2018, 2020) were used for mammals (downloaded from <https://zenodo.org/record/3690867#.Xyc5wyhKhPZ>). For amphibians, birds and reptiles (squamates only), I downloaded trees from <https://data.vertlife.org/>. Trees were from Jetz et al. (2012) for birds, from Jetz and Pyron (2018) for amphibians and from Tonini et al. (2016) for squamates. For each class, I downloaded a distribution of 1,000 trees, from which I obtained consensus trees to estimate phylogenetic signal (to that end, I used the TreeAnnotator programme of the BEAST software (Bouckaert et al., 2014)).

**Table S3.2: Phylogenetic signal in continuous and categorical traits.** BM: body mass; BL: body length; GL: generation length; MA: age at sexual maturity; ML: maximum longevity; L: longevity; LCS: litter/clutch size; HB: habitat breadth; TL: trophic level; DA: diel activity; UA: use of artificial habitats. Continuous traits were log-10 transformed to improve normality before estimating Pagel's  $\lambda$  – except for habitat breadth which was square-rooted. A star indicates a significant signal (p-value<0.05 for the log-likelihood ratio test in the case of  $\lambda$ ; and a significant difference from the simulated null distribution of  $\delta$  for categorical traits). 'NA' indicates traits that were not considered for a given class. All traits showed significant phylogenetic signal, with signals for BM, BL, L, GL, MA and LCS being particularly strong (above 0.8) across the four classes.

Class	Pagel's $\lambda$								Borges' $\delta$		
	BM	BL	GL	MA	ML	L	LCS	HB	TL	DA	UA
<b>Amphibians</b>	0.98*	0.94*	NA	0.85*	0.82*	NA	0.93*	0.99*	18*	3.4*	4.5*
<b>Birds</b>	0.99*	NA	0.97*	NA	NA	NA	0.95*	0.60*	13*	32e3*	1.8*
<b>Mammals</b>	0.99*	NA	0.97	NA	NA	NA	0.99	0.71	26*	17*	1.3*
<b>Reptiles</b>	1.0*	NA	NA	NA	0.94*	0.98*	1.0*	0.52*	6.3*	6.4*	1.4*

### S3.2.3 Implementation of missForest imputations

As phylogenetic signals were strong in many categorical and continuous traits (Table S3.2), I imputed missing trait values using random forest algorithms, as implemented in R with missForest (Stekhoven, 2016; Stekhoven and Bühlmann, 2012). Another advantage of missForest was that, being a non-parametric approach, no prior assumption about data distribution was required. The data were therefore not transformed prior to imputations. In addition, Penone et al. (2014) showed that including phylogenetic information did not decrease the accuracy of imputations for traits that were less phylogenetically conserved, such as habitat breadth in this work.

Phylogenetic relationships were included as additional predictors in the form of phylogenetic eigenvectors (Diniz-Filho et al., 2012), extracted from the phylogenies using the PVR package (Santos, 2018). Following Penone et al. (2014), I included the first ten phylogenetic eigenvectors as additional predictors of missing trait values in each class, enough to minimise imputation error. As not all species were represented in the phylogenies, I also added taxonomic order as a predictor for all species. All traits in Table S3.2 were included in the imputations. Tuning parameters of missForest were set to ten maximum iterations and to one

hundred trees grown in each forest.

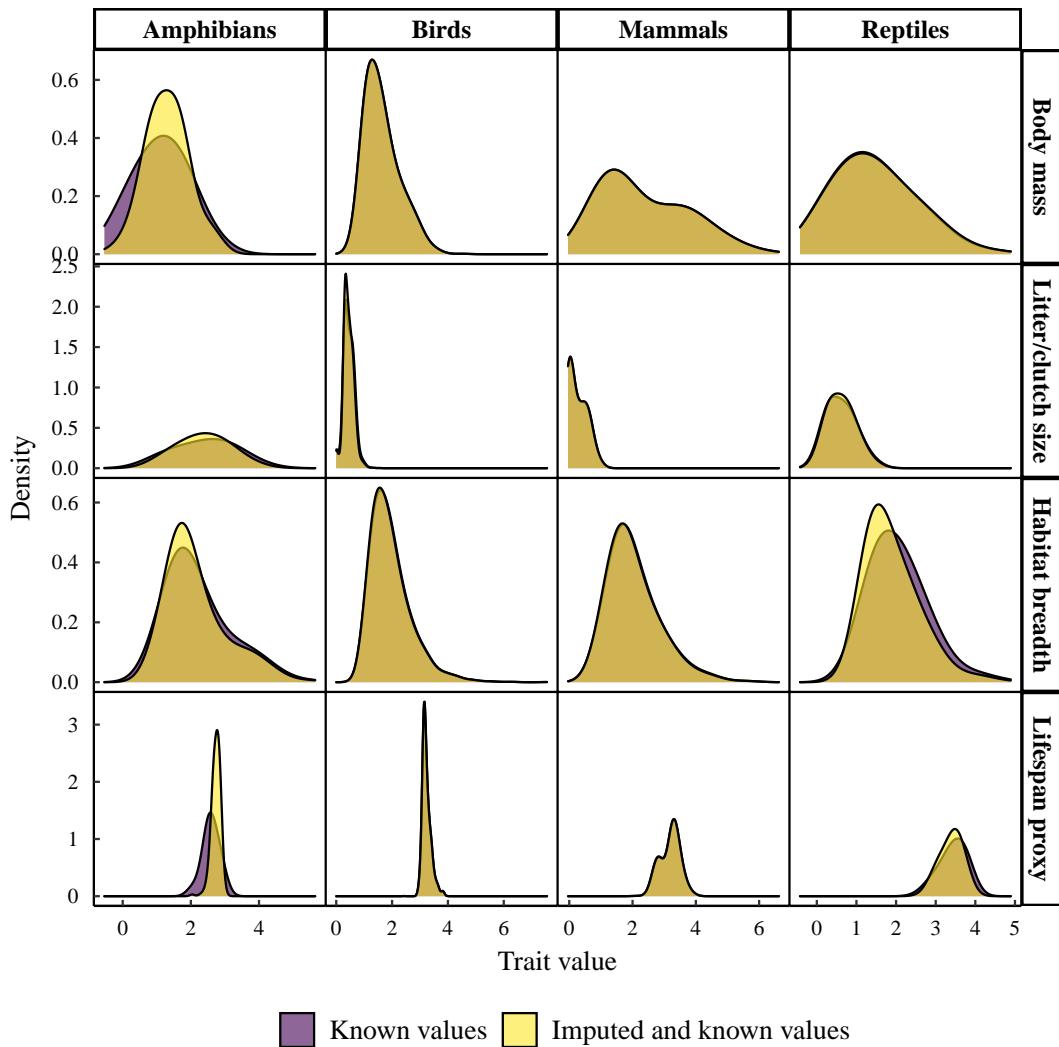
### S3.3 Degree of multicollinearity among traits

Multicollinearity among traits can be problematic when calculating functional diversity indices (Cadotte et al., 2011). After imputing missing trait values and before estimating functional metrics, I hence assessed whether the degree of multicollinearity among categorical and continuous traits was not problematically high. To that end, I used generalised variance inflation factors (Fox and Monette, 1992). Given a regression model, variance inflation factors quantify the overestimation in the variance of estimated regression coefficients due to multicollinearity among the predictors. A GVIF value of 5 or 10 is commonly used as a threshold to select out collinear predictors (Dormann et al., 2013). I used the stepwise.vif function of the Rnalytica package (<https://github.com/awsm-research/Rnalytica>), with a threshold of 5, to determine the GVIF of each trait. I used the imputed traits from the 8th imputation iteration to assess whether multicollinearity was problematically high. Continuous traits were log-10 transformed (except for habitat breadth which was square-rooted). Multicollinearity across traits was not detected to be problematically high, as all traits had a GVIF value below 2 (Table S3.3). As such, all seven traits were included in the calculation of functional diversity indices.

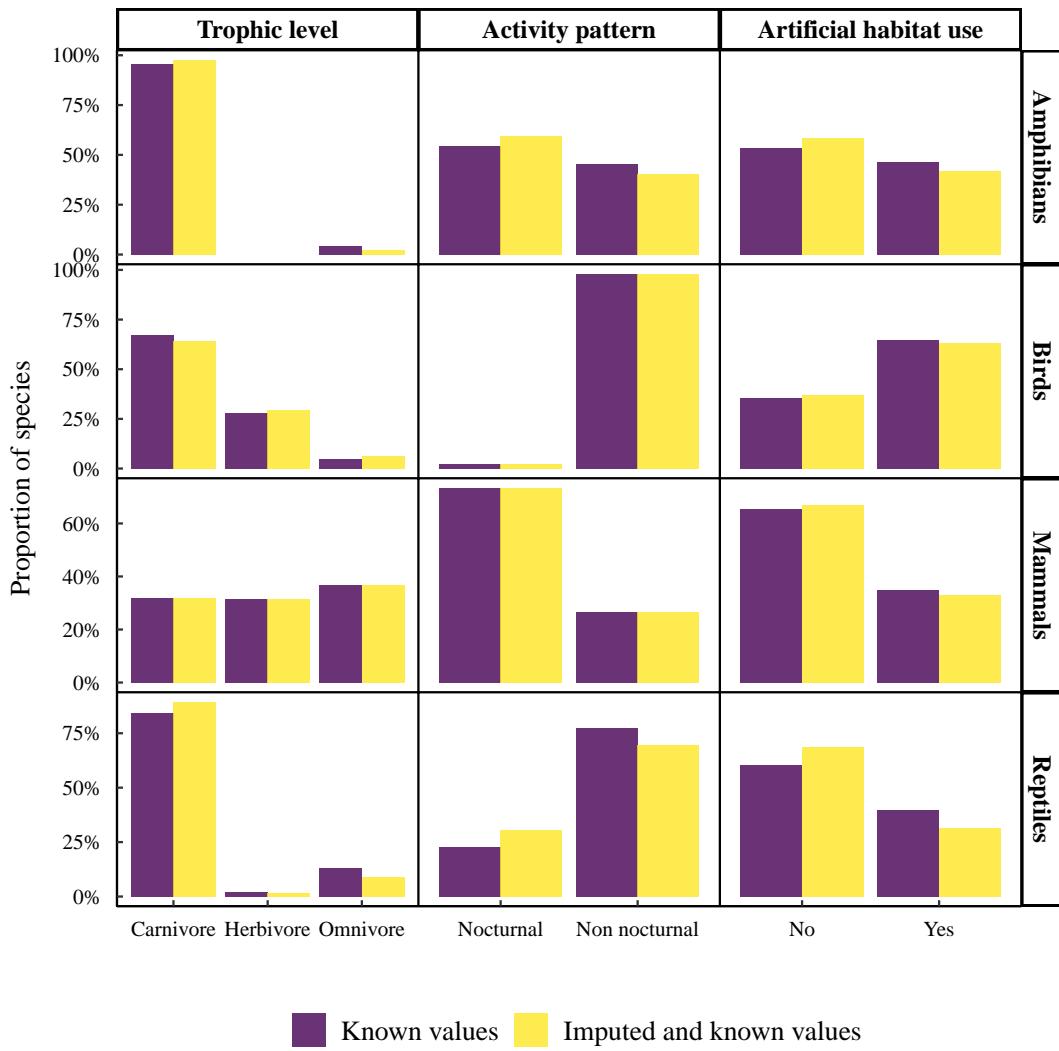
**Table S3.3: Variance inflation factors across considered (imputed) traits.**

Trait	GVIF
Diel activity	1.1
Trophic level	1.3
Use of artificial habitats	1.4
Body mass (log10)	1.5
Habitat breadth (square-root)	1.5
Litter/clutch size (log10)	1.6
Lifespan proxy (log10)	1.7

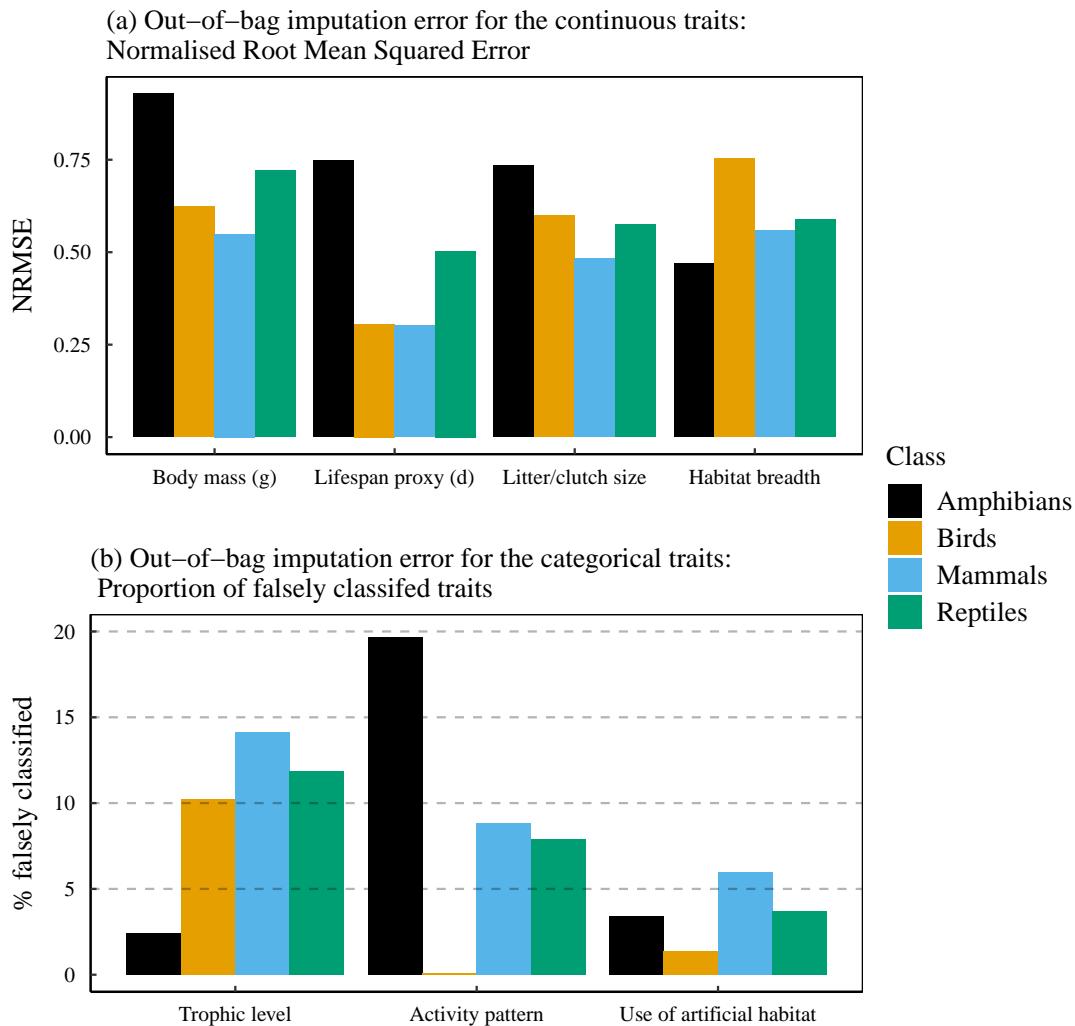
## S3.4 Imputation performance



**Figure S3.6:** Distribution of continuous traits considered in the calculation of the functional diversity metrics (shown as density plots), before and after missing value imputations, in each class and for the species occurring in the PREDICTS database. All traits were log<sub>10</sub>-transformed except Habitat breadth, which was square-rooted.



**Figure S3.7:** Frequency distribution for categorical traits considered in the calculation of the functional diversity metrics (shown as % of total species in each category) before and after missing value imputations, in each class, for the species occurring in the PREDICTS database.



**Figure S3.8: Out-of-bag imputation errors for the continuous traits (a) and categorical traits (b).** For continuous traits, the Normalised Root Mean Squared Error (NRMSE) was obtained by dividing the Mean Squared Error (MSE) by the variance of the known trait distribution, then square-rooting the result. The MSE was returned for each trait by the missForest function (missForest package, Stekhoven, 2016; Stekhoven and Bühlmann, 2012) and corresponds to an out-of-bag error. For categorical traits, the error was estimated as the out-of-bag proportion of falsely classified traits.

### S3.5 Functional loss and functional gain

Across all vertebrates, I estimated functional loss and gain using 84 studies for the tropical subset and 39 studies for the temperate subset (51,514 and 30,470 pairwise comparisons between sites respectively, Table S3.4). Because of this large number of pairwise comparisons, I did not develop a null modelling approach (if I used 100 randomisations per pair of sites, I would need to compute functional loss and gain for more than 8 million pairs, which would be very computationally demanding). I grouped mature, intermediate and young secondary vegetation together in this analysis. I could not estimate the effects in all land uses (for instance, sample sizes for tropical urban sites were too small).

Within classes, I used 18 tropical studies and 1 temperate study for amphibians; 38 and 21 for birds (respectively); 28 and 9 for mammals; and 11 and 7 for reptiles. As sample sizes differed among pairs of land uses and use I was not able to estimate all effects, notably for the intensely-used land uses.

To calculate functional loss and functional gain, the Gower distance matrix was first subsetted to the species occurring in a given pair of sites (see Chapter 3, Methods, ‘Functional traits and diversity indices’). Cailliez corrections were applied when the distance matrix was not Euclidian (Cailliez corrections consist of applying the smallest positive constant to the distances so as to make them Euclidian (Cailliez, 1983); ade4 R package (Dray et al., 2007)). I then performed a principal coordinates analysis on the (corrected) Gower distance matrix, retaining the first two axes to reduce the computational load in the calculation of convex hulls. Sites that contained fewer than three functionally different species were excluded (the computation of a convex hull requiring more species in the assemblage than PCoA axes). Then I estimated the volume of trait space occupied by each assemblage of a given pair, as well as the volume of the shared trait space (intersection), from which I derived functional loss and functional gain.

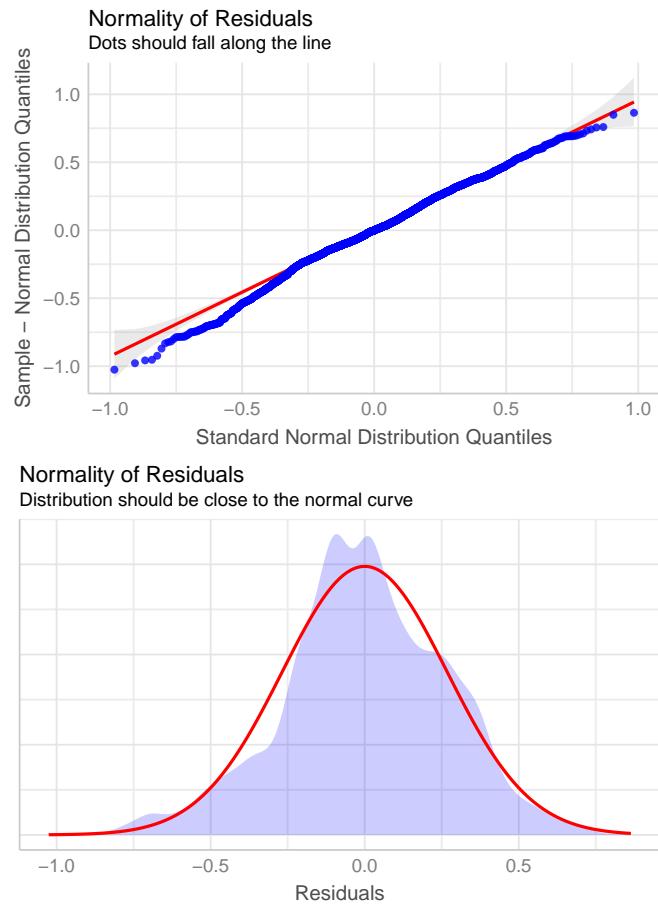
**Table S3.4: Sample sizes (number of pairs of sites) for the calculation of functional loss and functional gain across all species.**

Region	Pairs	Minimal use	Light use	Intense use
Temperate	PV-PV	7626	22546	492
	PV-SV	511	72	–
	PV-PF	9	166	–
	PV-PA	8	40	–
	PV-CR	150	–	–
	PV-UR	6306	1197	7
Tropical	PV-PV	8547	4016	16722
	PV-SV	6584	1124	9713
	PV-PF	580	1378	–
	PV-PA	36	20	22
	PV-CR	1700	1088	–
	PV-UR	–	–	–

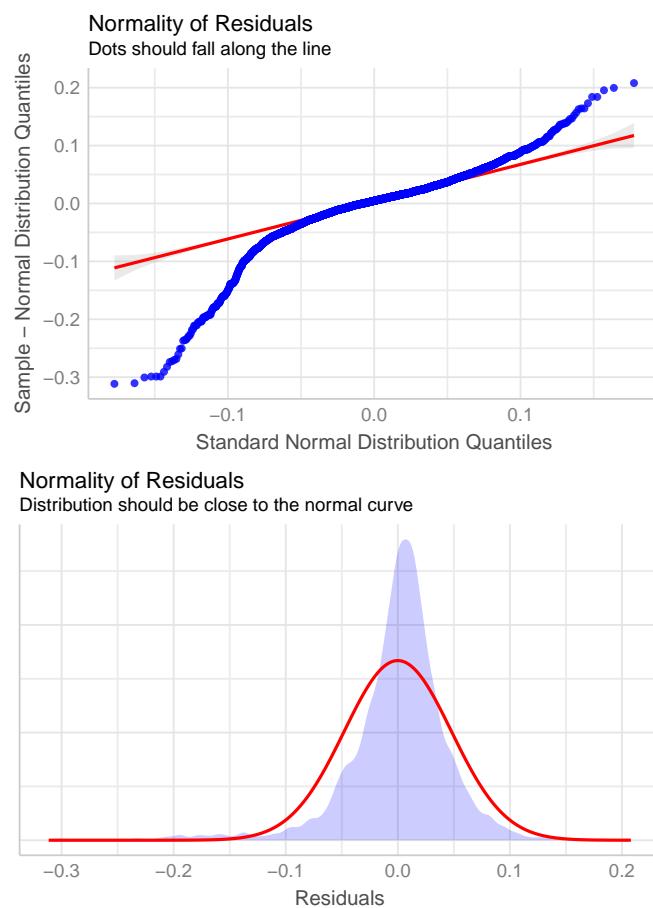
**Table S3.5: Sample sizes (number of pairs of sites) for the calculation of functional loss and functional gain within each class.**

Class	Region	Pair of land uses	Minimal use	Light use	Intense use
Amphibians	Temperate	PV/PV	–	45	–
Amphibians	Temperate	PV/SV	8	70	–
Amphibians	Temperate	PV/AGR	3	–	–
Amphibians	Temperate	PV/UR	6	100	–
Amphibians	Tropical	PV/PV	501	241	307
Amphibians	Tropical	PV/SV	838	–	90
Amphibians	Tropical	PV/PF	422	91	–
Amphibians	Tropical	PV/AGR	1	3	1
Birds	Temperate	PV/PV	7,382	19,300	491
Birds	Temperate	PV/SV	150	1	–
Birds	Temperate	PV/PF	9	166	–
Birds	Temperate	PV/AGR	145	40	–
Birds	Temperate	PV/UR	6,300	992	–
Birds	Tropical	PV/PV	5,059	3,117	9,014
Birds	Tropical	PV/SV	3,491	1,058	5,225
Birds	Tropical	PV/PF	156	994	–
Birds	Tropical	PV/AGR	1,626	1,085	–
Mammals	Temperate	PV/PV	110	3,030	–
Mammals	Temperate	PV/SV	25	–	–
Mammals	Temperate	PV/AGR	5	–	–
Mammals	Temperate	PV/UR	–	105	7
Mammals	Tropical	PV/PV	1,989	637	64
Mammals	Tropical	PV/SV	230	65	8
Mammals	Tropical	PV/PF	2	–	–
Mammals	Tropical	PV/AGR	109	20	21
Reptiles	Temperate	PV/PV	132	2	1
Reptiles	Temperate	PV/SV	250	1	–
Reptiles	Temperate	PV/AGR	5	–	–
Reptiles	Tropical	PV/PV	989	137	5,140
Reptiles	Tropical	PV/SV	1,760	1	3,456
Reptiles	Tropical	PV/PF	–	190	–

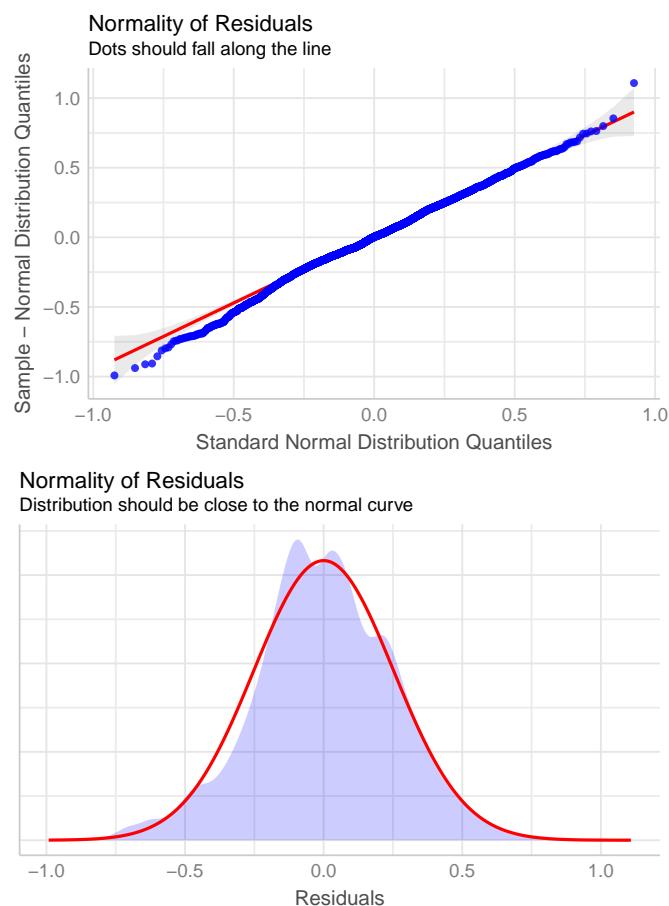
## S3.6 Diagnostic plots



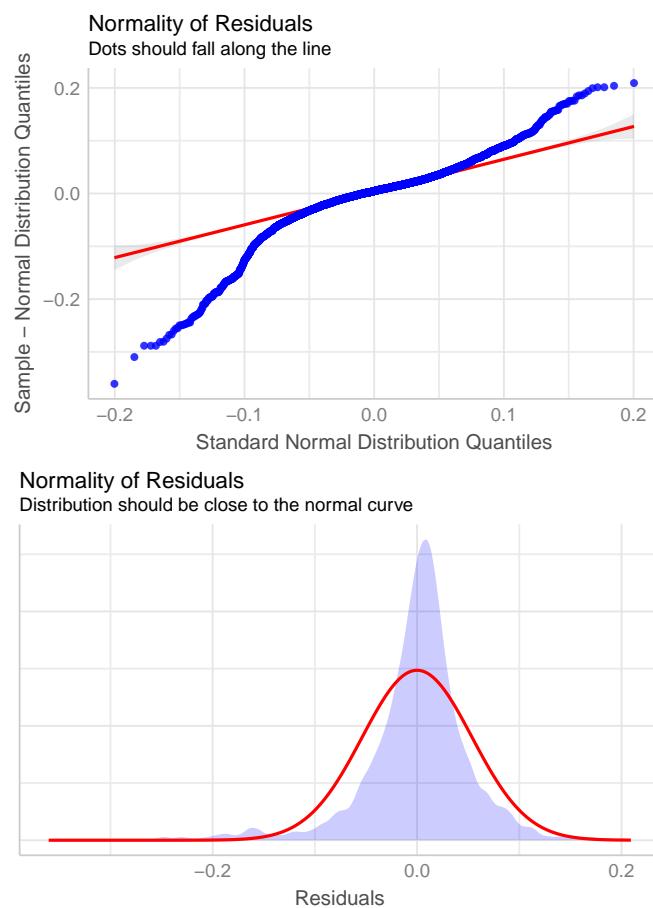
**Figure S3.9:** Diagnostic plots for Model 1a, obtained using the 'performance' R package (Lüdecke et al., 2021).



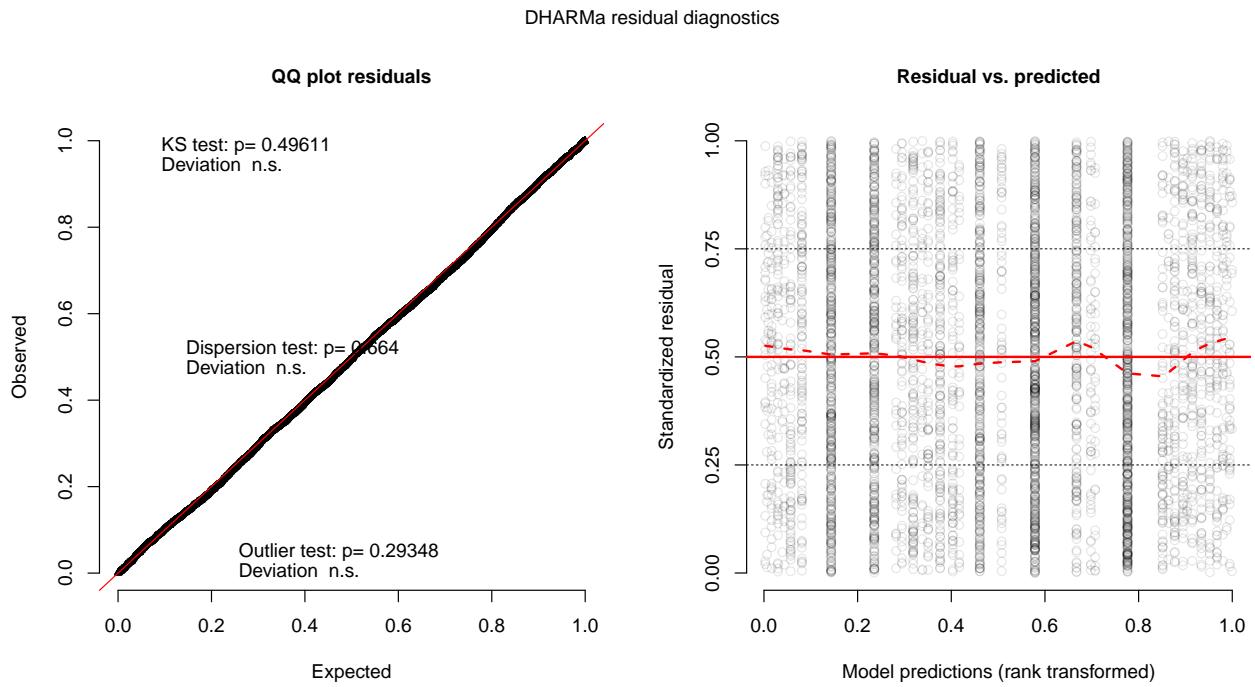
**Figure S3.10:** Diagnostic plots for **Model 1b**, obtained using the ‘performance’ R package (Lüdecke et al., 2021).



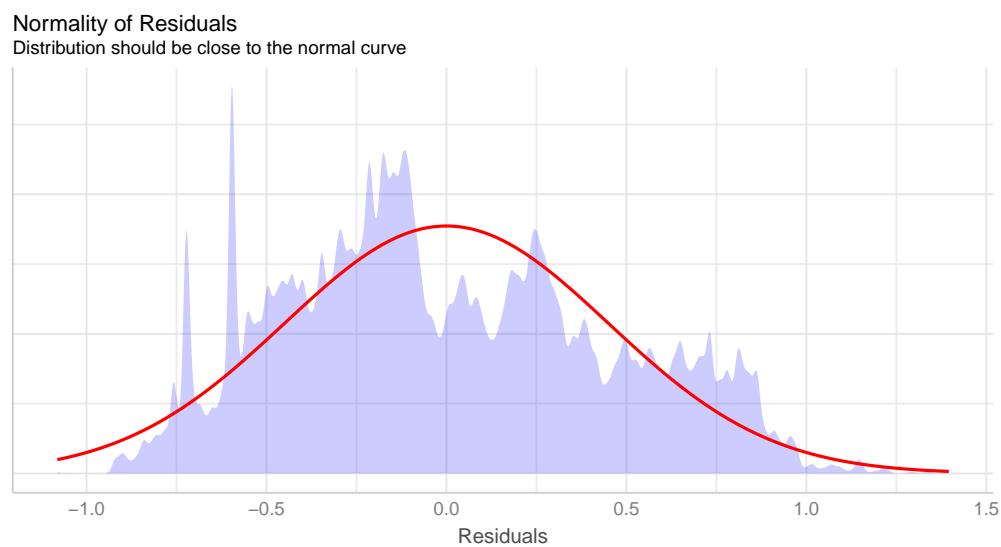
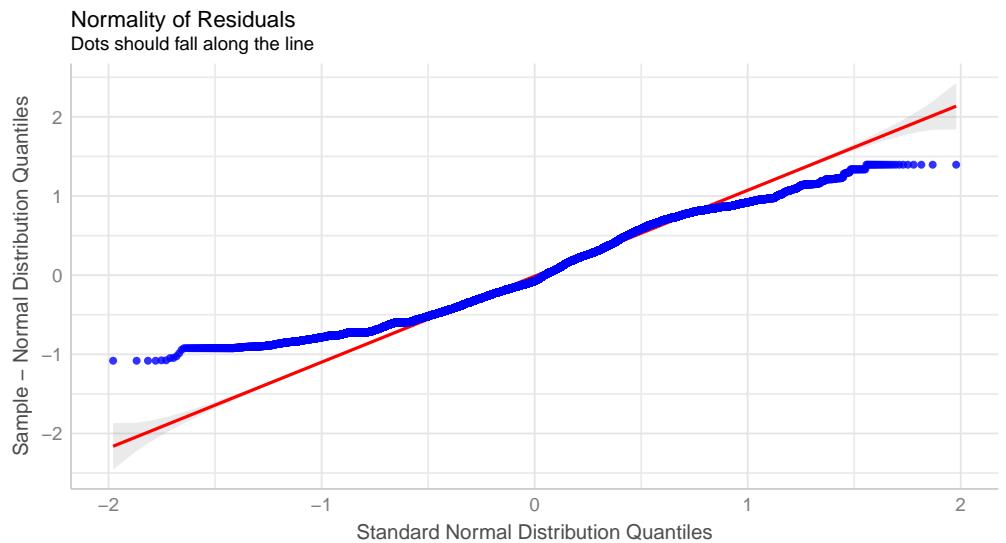
**Figure S3.11:** Diagnostic plots for Model 2a, obtained using the ‘performance’ R package (Lüdecke et al., 2021).



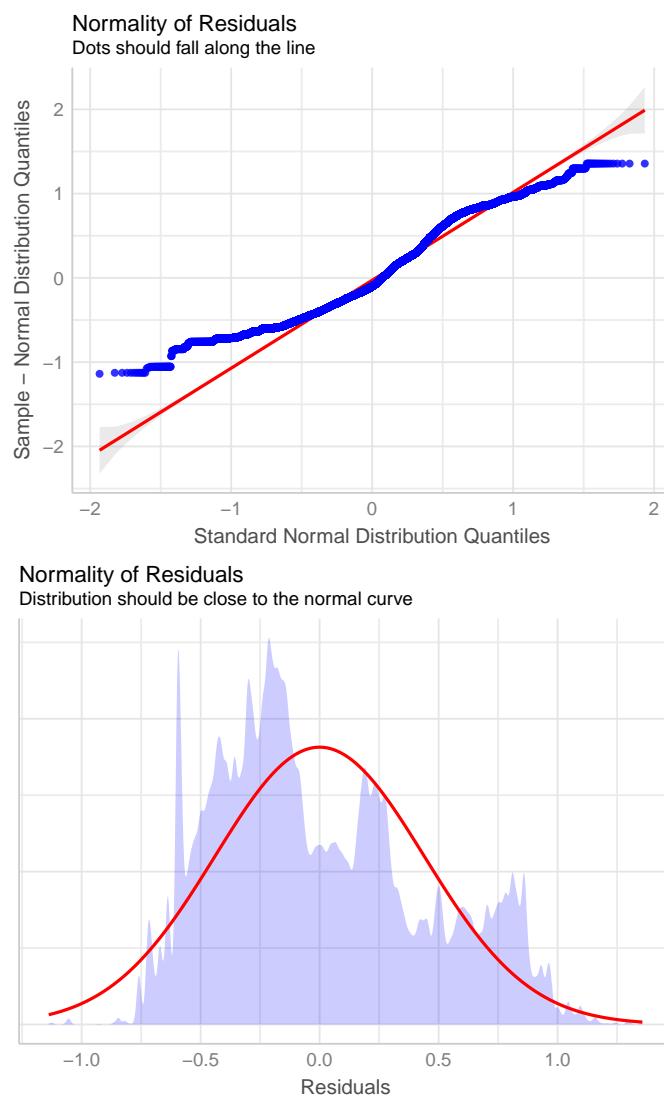
**Figure S3.12:** Diagnostic plots for Model 2b, obtained using the ‘performance’ R package (Lüdecke et al., 2021).



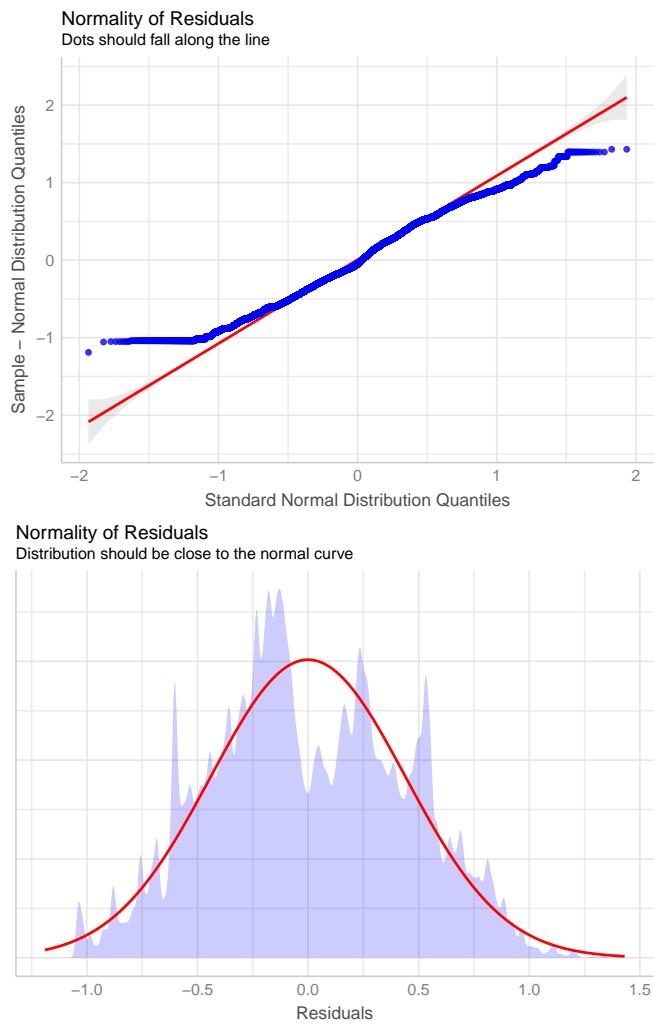
**Figure S3.13: Diagnostic plots for Model 3**, obtained using the ‘DHARMA’ R package (Hartig, 2021).



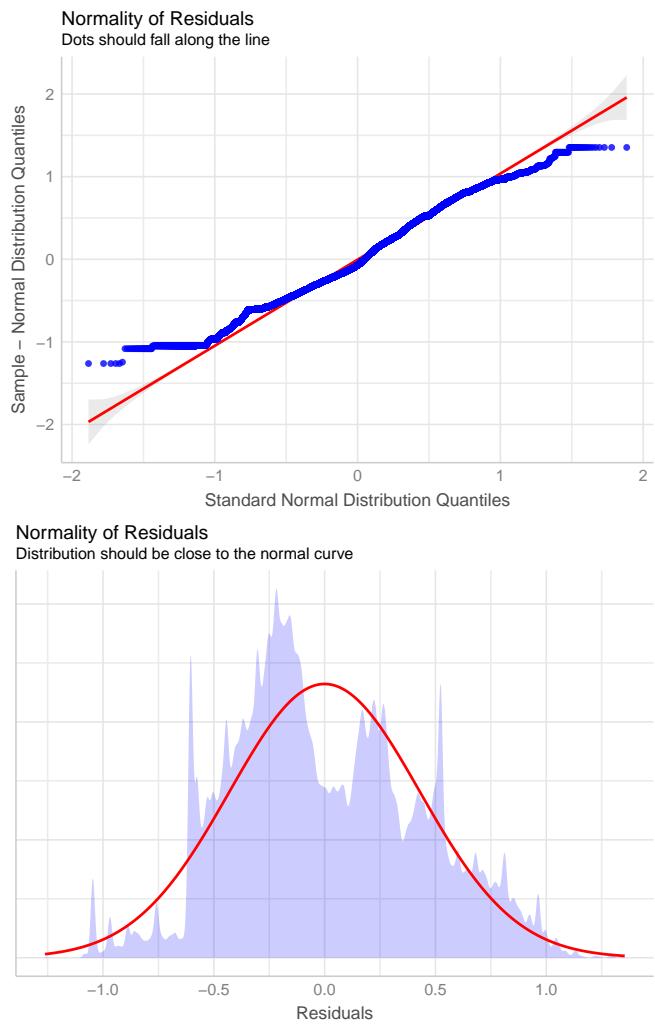
**Figure S3.14:** Diagnostic plots for Model 4a, obtained using the ‘performance’ R package (Lüdecke et al., 2021).



**Figure S3.15:** Diagnostic plots for Model 4b, obtained using the ‘performance’ R package (Lüdecke et al., 2021).

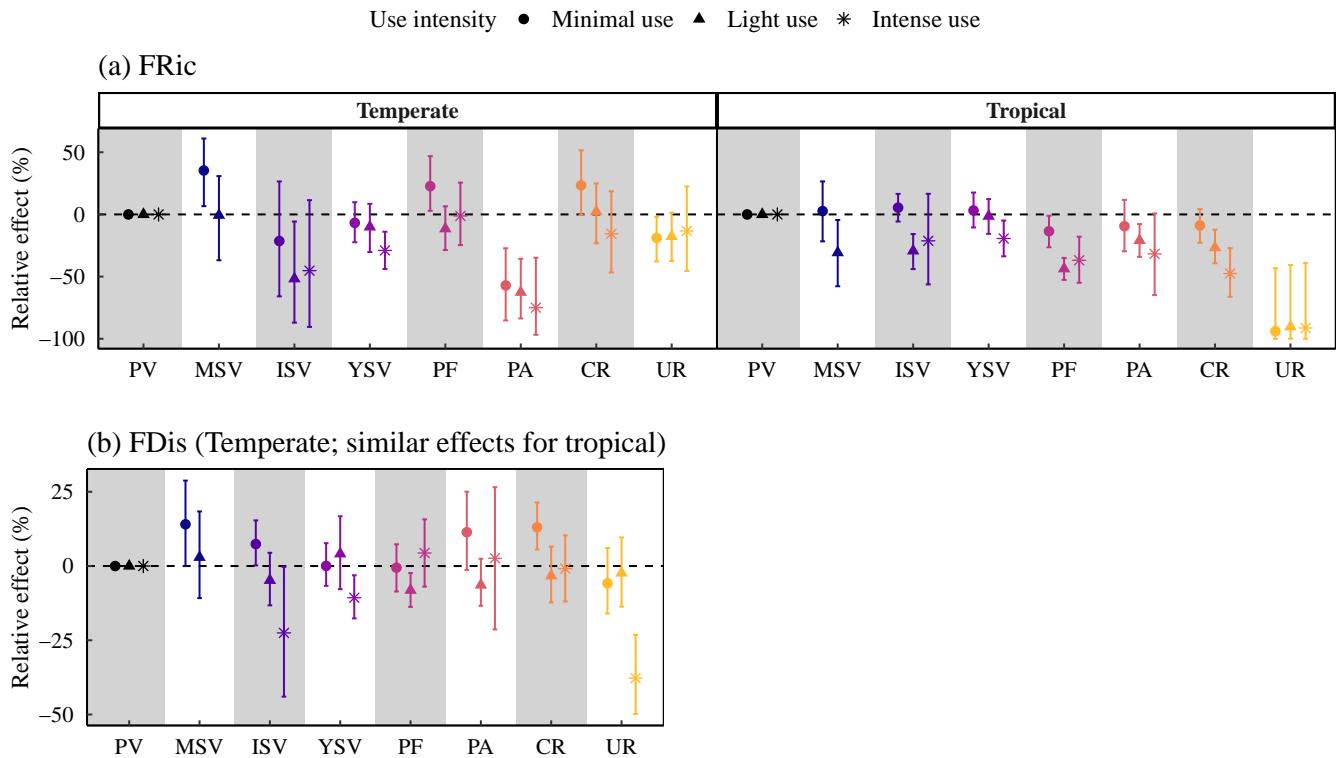


**Figure S3.16:** Diagnostic plots for Model 5a, obtained using the ‘performance’ R package (Lüdecke et al., 2021).

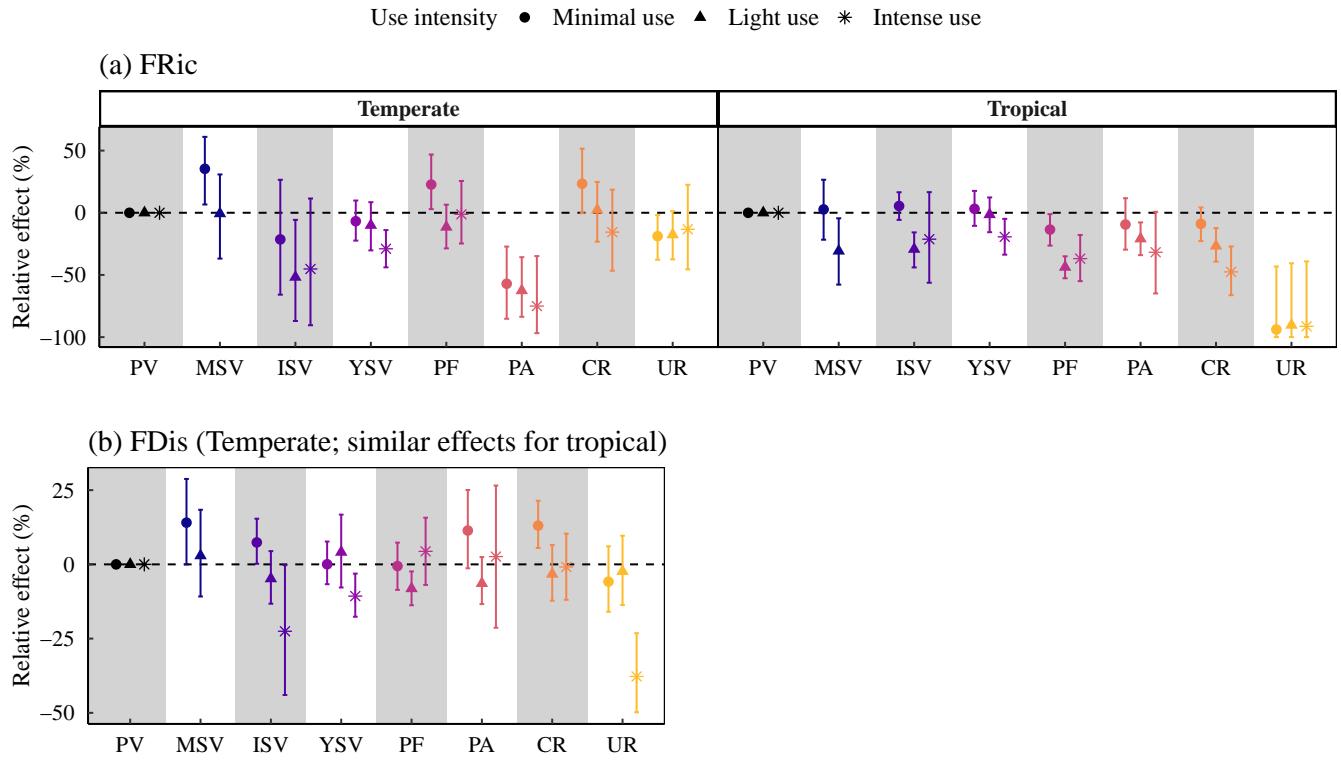


**Figure S3.17:** Diagnostic plots for Model 5b, obtained using the ‘performance’ R package (Lüdecke et al., 2021).

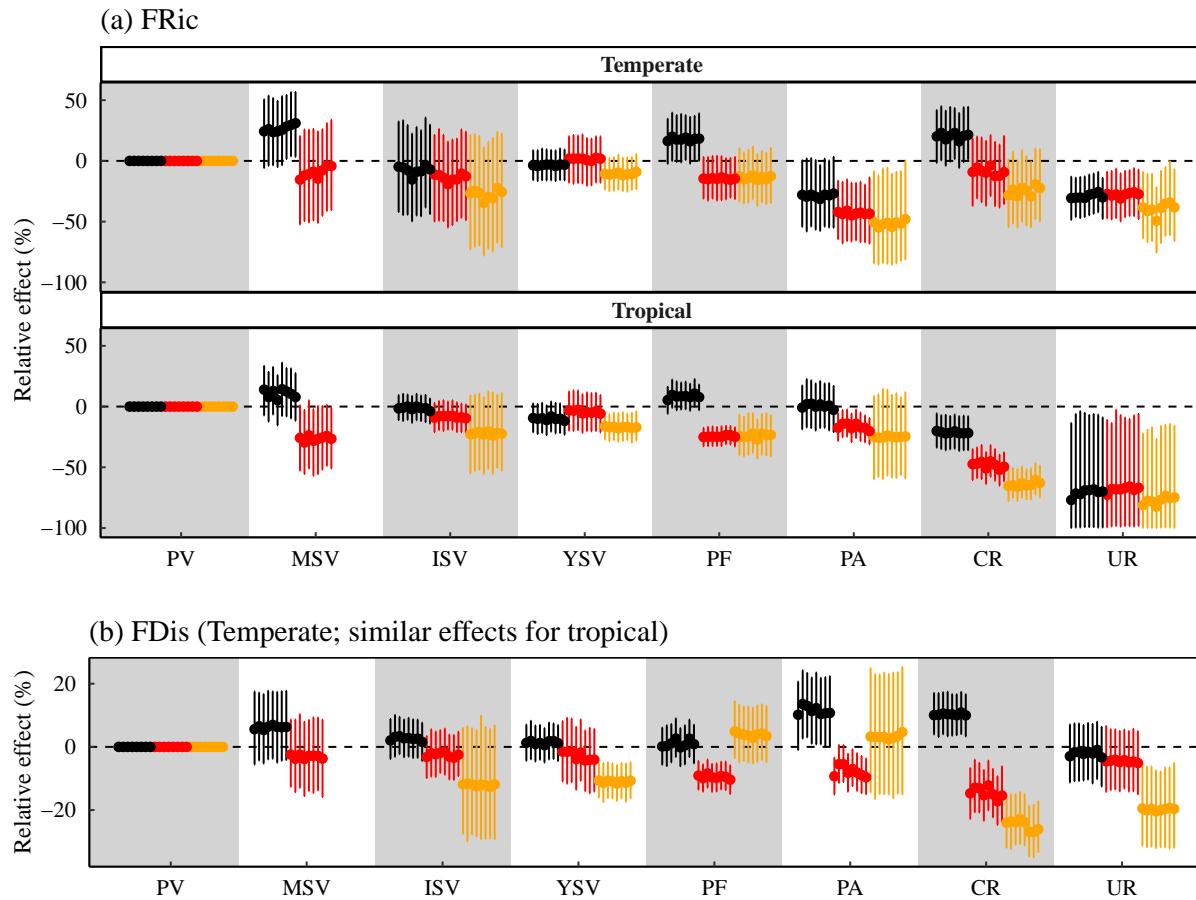
### S3.7 Model robustness



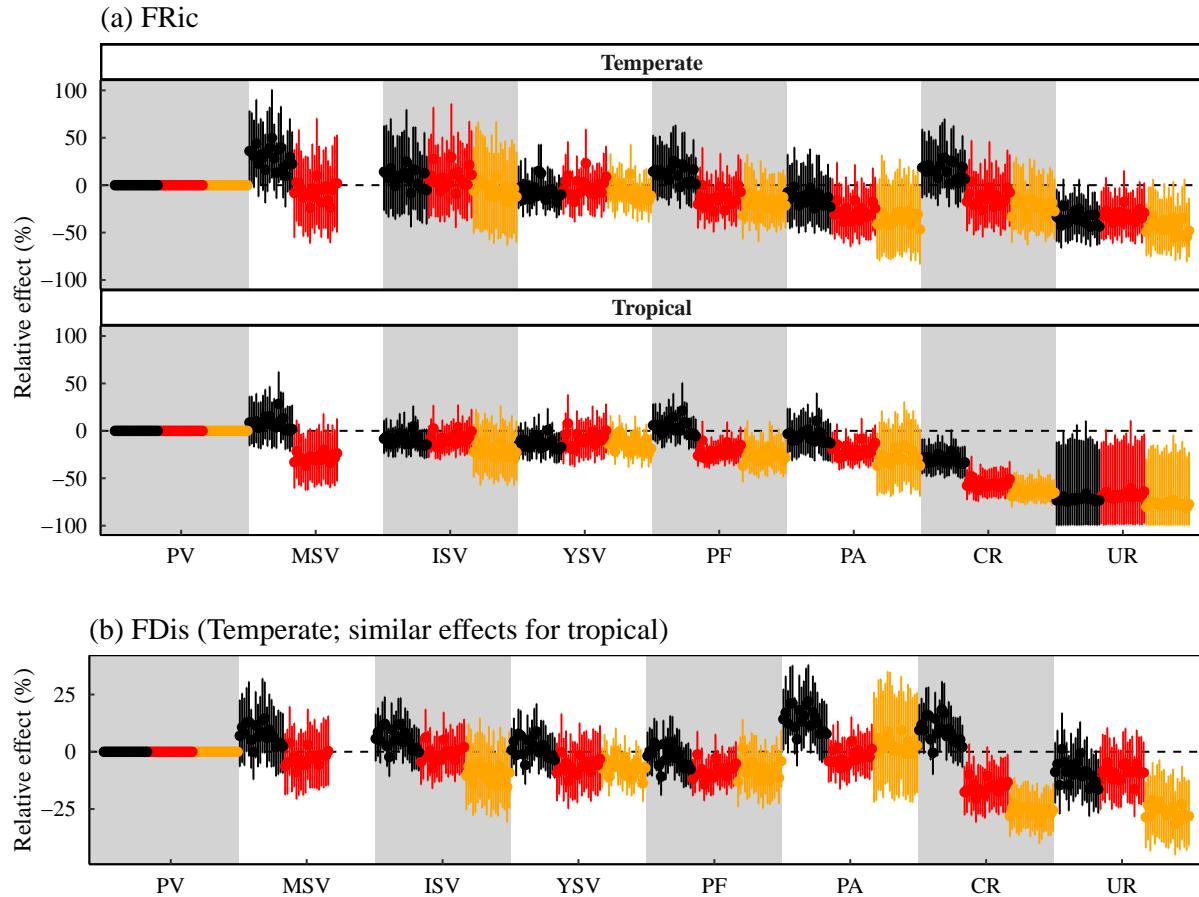
**Figure S3.18: Effects of land use, land-use intensity and region on FRic (a) and FDis (b) across vertebrates, for the subset of species with complete trait data (i.e., excluding species with less than 100% trait completeness).** Effects are plotted as a % difference relative to the reference land use (primary vegetation, PV). For FRic, I fitted Model 1a (see main text, Chapter 3), which included the effects of land use, land-use intensity and region, as well as interactions between land use and land-use intensity and between land use and region. For FDis, I fitted Model 1b, which included effects of land use, land-use intensity and region, and interactions between land use and land-use intensity. Error bars represent 95% confidence intervals. MSV: mature secondary vegetation; ISV: intermediate secondary vegetation; YSV: young secondary vegetation; PF: plantation forest; PA: pasture; CR: cropland; UR: urban. Effects for Intense use in MSV could not be estimated as there weren't enough sampled sites.



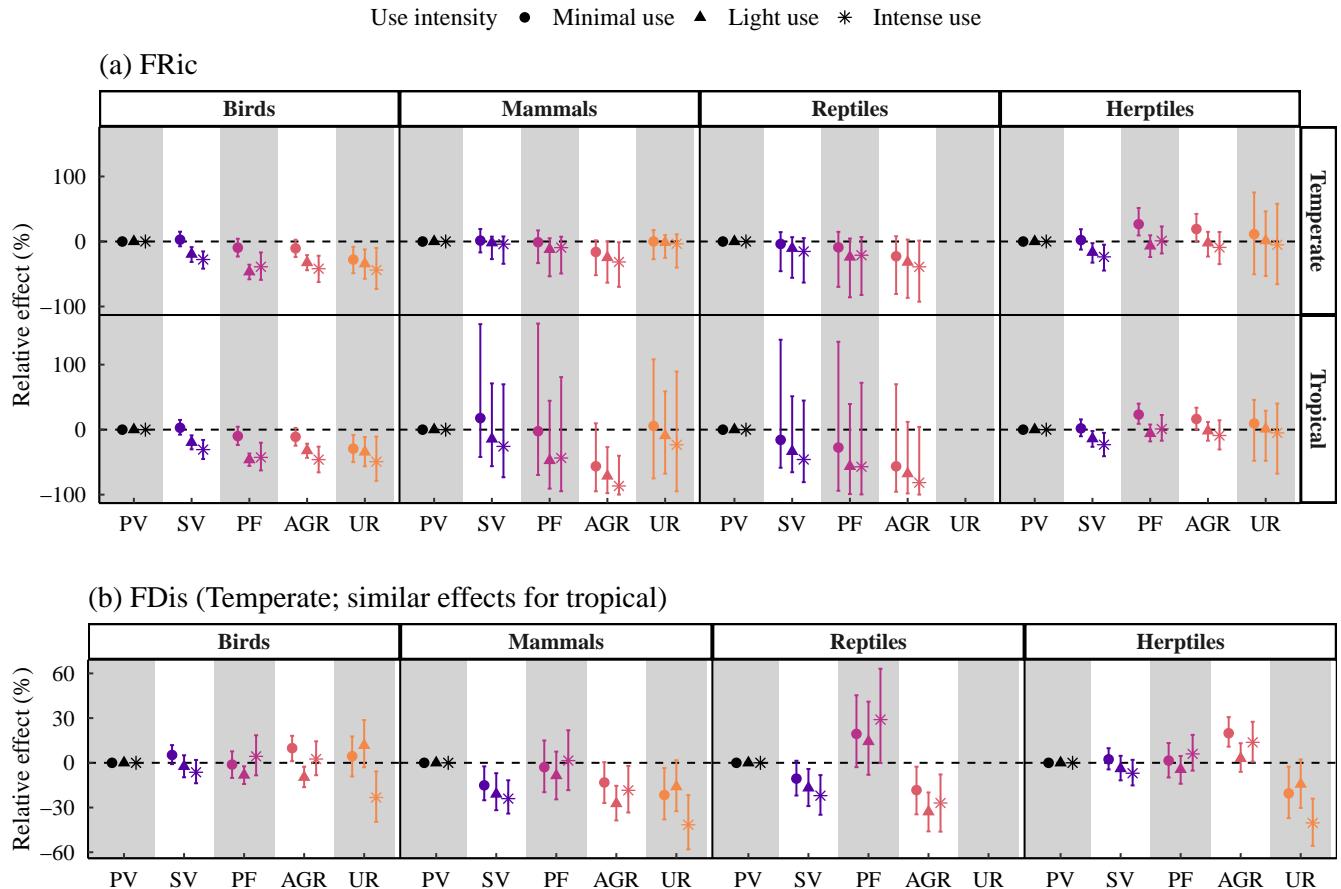
**Figure S3.19: Effects of land use, land-use intensity and region on FRic (a) and FDis (b), for the subset of species with complete trait data (i.e., excluding species with less than 100% trait completeness), with geographical range size as an additional trait considered in the calculation of functional diversity metrics.** Effects are plotted as a % difference relative to the reference land use (primary vegetation, PV). For FRic, I fitted Model 1a (see main text, Chapter 3), which included the effects of land use, land-use intensity and region, as well as interactions between land use and land-use intensity and between land use and region. For FDis, I fitted Model 1b , which included effects of land use, land-use intensity and region, and interactions between land use and land-use intensity. Error bars represent 95% confidence intervals. MSV: mature secondary vegetation; ISV: intermediate secondary vegetation; YSV: young secondary vegetation; PF: plantation forest; PA: pasture; CR: cropland; UR: urban. Effects for Intense use in MSV could not be estimated as there weren't enough sampled sites.



**Figure S3.20: Effects of land use and land-use intensity on FRic (a) and FDis (b), obtained when calculating FRic and FDis with each set of imputed traits (eight in total). Effects are plotted as a % difference relative to the reference land use (primary vegetation, PV). For FRic, I fitted Model 1a (see main text, Chapter 3), which included the effects of land use, land-use intensity and region, as well as interactions between land use and land-use intensity and between land use and region. For FDis, I fitted Model 1b, which included effects of land use, land-use intensity and region, and interactions between land use and land-use intensity. Error bars represent 95% confidence intervals. MSV: mature secondary vegetation; ISV: intermediate secondary vegetation; YSV: young secondary vegetation; PF: plantation forest; PA: pasture; CR: cropland; UR: urban. Effects for Intense use in MSV could not be estimated as there weren't enough sampled sites.**



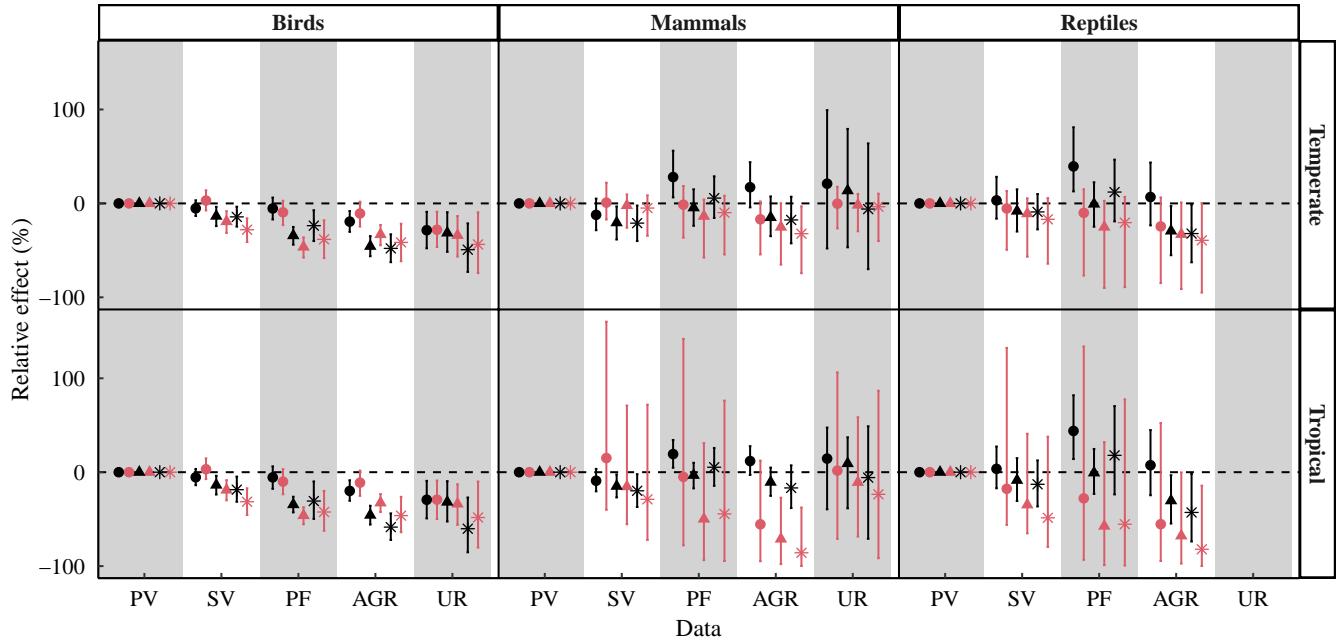
**Figure S3.21: Effects of land use and land-use intensity on FRic (a) and FDis (b), obtained when re-sampling primary vegetation sites twenty independent times.** I fixed the sample size for primary vegetation sites at 50. For FRic, we fitted Model 1a (see main text, Chapter 3), which included the effects of land use, land-use intensity and region, as well as interactions between land use and land-use intensity and between land use and region. For FDis, I fitted Model 1b, which included effects of land use, land-use intensity and region, and interactions between land use and land-use intensity. Error bars represent 95% confidence intervals. MSV: mature secondary vegetation; ISV: intermediate secondary vegetation; YSV: young secondary vegetation; PF: plantation forest; PA: pasture; CR: cropland; UR: urban. Effects for Intense use in MSV could not be estimated as there weren't enough sampled sites.



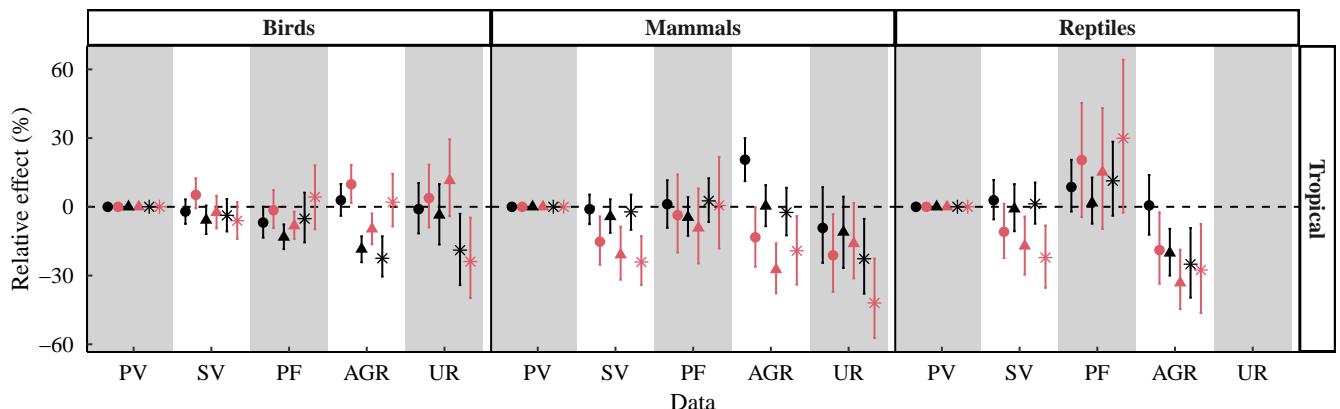
**Figure S3.22: Effects of land use, land-use intensity and taxonomic class on FRic (a) and FDis (c), for the subset of species with complete trait data (i.e., excluding species with less than 100% trait completeness).** Effects are plotted as a % difference relative to the reference land use (primary vegetation, PV). I did not include the effects of region here as sample sizes were not large enough for some classes. For FRic, the model included the effects of land use, land-use intensity and class, and interactions between land use and land-use intensity as well as land use and class. For FDis, the model included an additional interaction between land-use intensity and class. Error bars represent 95% confidence intervals. SV: secondary vegetation; PF: plantation forest; PA: pasture; CR: cropland; UR: urban. Effects for reptiles in urban land uses could not be estimated as there weren't enough sampled sites.

—●— complete and imputed trait data —●— complete trait data subset      Use intensity    ● Minimal use    ▲ Light use    \* Intense use

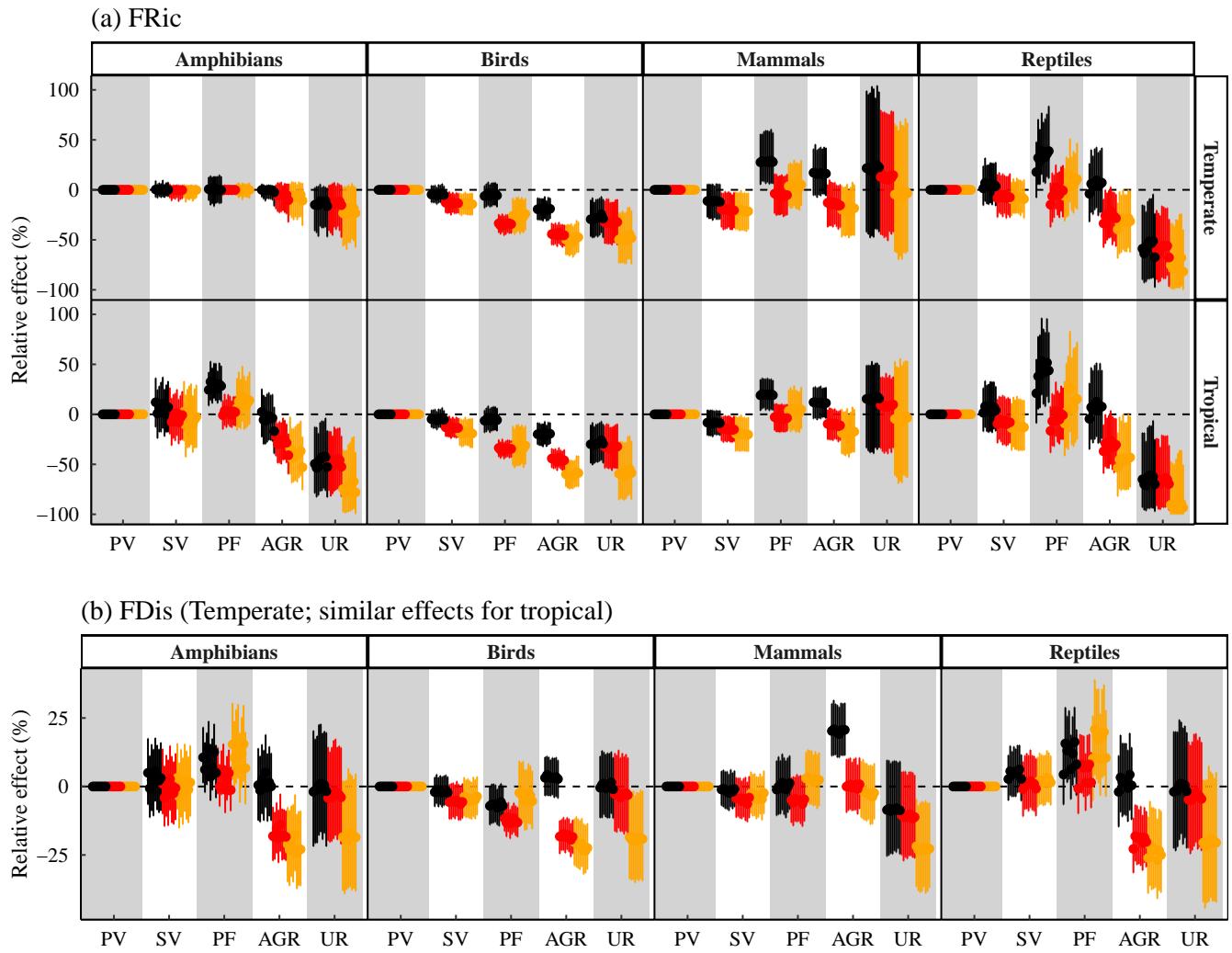
(a) FRic



(a) FDis



**Figure S3.23: Effects of land use, region, land-use intensity and taxonomic class on FRic and FDis obtained with the imputed trait data (black points) or with the complete data subsets (red points).** Effects are plotted as a % difference relative to the reference land use (primary vegetation, PV). For FRic, I fitted Model 2a (see main text, Chapter 3), and I fitted Model 2b for FDis. Error bars represent 95% confidence intervals. SV: secondary vegetation; PF: plantation forest; AGR: agricultural (cropland and pasture); UR: urban. Effects for reptiles in urban land uses could not be estimated as there weren't enough sampled sites.



**Figure S3.24: Effects of land use, region, land-use intensity and taxonomic class on FRic and FDis, obtained when calculating FRic and FDis with each set of imputed traits (eight in total).** Effects are plotted as a % difference relative to the reference land use (primary vegetation, PV). For FRic, I fitted Model 2a (see main text, Chapter 3), and Model 2b for FDis. Error bars represent 95% confidence intervals. SV: secondary vegetation; PF: plantation forest; AGR: agricultural (cropland and pasture); UR: urban. Effects for reptiles in urban land uses could not be estimated as there weren't enough sampled sites.

## S3.8 Model robustness – time since land-use conversion

Time since land-use conversion could have important impacts on assemblage composition and thus, on local functional diversity. I did not investigate these effects because PREDICTS contained data on time since land-use conversion only for about 22% of the sites, considerably reducing samples sizes. Here, I investigated whether the results are likely robust to the inclusion of time since land-use conversion using the subset of sites for which time since land-use conversion was provided. To this end, I found the best-fitting models explaining FRic and FDis, using backwards stepwise selection, starting with complete models that included the effects of land use, time since land-use conversion, region, land-use intensity (for FRic only) and all two-way interactions among these predictors.

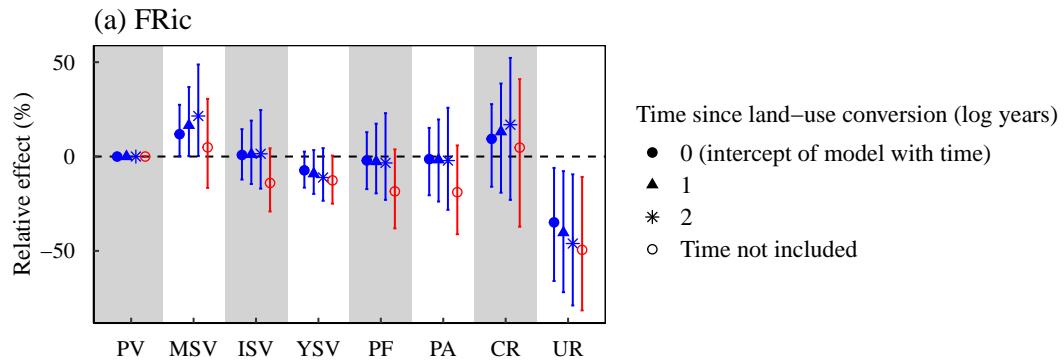
- For FRic, the best-fitting model includes the main effects of land use and time since land-use conversion, but no interaction between these predictors. The model's summary (Table S3.6) show that time since conversion has a significant negative effect on FRic, but the relationship between FRic and time since land-use conversion is similar in different land uses (as there are no interactions between land use and time since conversion, such that the slopes are similar in different land uses, and so the rate at which FRic decreases with time is similar in different land uses). The intercept is only different for urban land uses (significantly lower). Thus, based on this data subset, I expect time since land-use conversion to have a similar effect in different land uses.

**Table S3.6:** Summary of the model explaining FRic by land use and time since land-use conversion, fitted on the subset of data for which there are information on time since land-use conversion.

	Estimate	Std. Error	t value
Intercept: Primary vegetation	1.156	0.073	15.921
Mature secondary vegetation	0.178	0.093	1.907
Intermediate secondary vegetation	0.018	0.072	0.249
Young secondary vegetation	-0.078	0.051	-1.532
Plantation forest	-0.018	0.082	-0.224
Pasture	-0.005	0.093	-0.054
Cropland	0.133	0.152	0.875
Urban	-0.316	0.133	-2.368
log_Years	-0.094	0.021	-4.566

I then compare this model's predictions with a simpler model that does not account for time since land-use conversion ( $\text{FRic} \sim \text{Land use}$ ). The predictions (Fig. S3.25) show that including time since land-use conversion does not bias the results, as I find a similar significant effect with both models in

urban land uses, and elsewhere the effects are congruent. Thus, given this data subset, I argue that the results are robust to the inclusion of time since land-use conversion.



**Figure S3.25: Effects of land use on FRic for the model that includes time since land-use conversion (blue points) versus the model that does not take time since land-use conversion into account (red points).**

- For FDis, the best-fitting model includes the main effects of land use, time since land-use conversion as well as interactions between land use and time since land-use conversion (I didn't consider land-use intensity in the starting model because of sample size issues). Nevertheless, the main effect of time since land-use conversion is not significant (Table S3.7), and the relationship between time since land-use conversion and FDis is not significant in most land uses (except for plantation forest). Thus, I argue the available data do not allow us to properly investigate the relationship between time since land-use conversion and FDis.

**Table S3.7:** Summary of the model explaining FDis by land use and time since land-use conversion, fitted on the subset of data for which I have information on time since land-use conversion.

	Estimate	Std. Error	t value
Intercept: Primary vegetation	0.366	0.011	32.219
Mature secondary vegetation	0.032	0.055	0.577
Intermediate secondary vegetation	-0.015	0.050	-0.298
Young secondary vegetation	0.020	0.015	1.386
Plantation forest	0.074	0.023	3.213
Pasture	-0.017	0.048	-0.346
Cropland	-0.013	0.042	-0.317
Urban	0.031	0.054	0.573
log_Years	-0.004	0.004	-1.186
Mature secondary vegetation:log_Years	-0.005	0.015	-0.335
Intermediate secondary vegetation:log_Years	0.011	0.016	0.650
Young secondary vegetation:log_Years	-0.008	0.007	-1.170
Plantation forest:log_Years	-0.023	0.007	-3.077
Pasture:log_Years	0.010	0.015	0.688
Cropland:log_Years	0.007	0.012	0.620
Urban:log_Years	-0.016	0.022	-0.714

## **Appendix 3: Supporting information for Chapter 4**

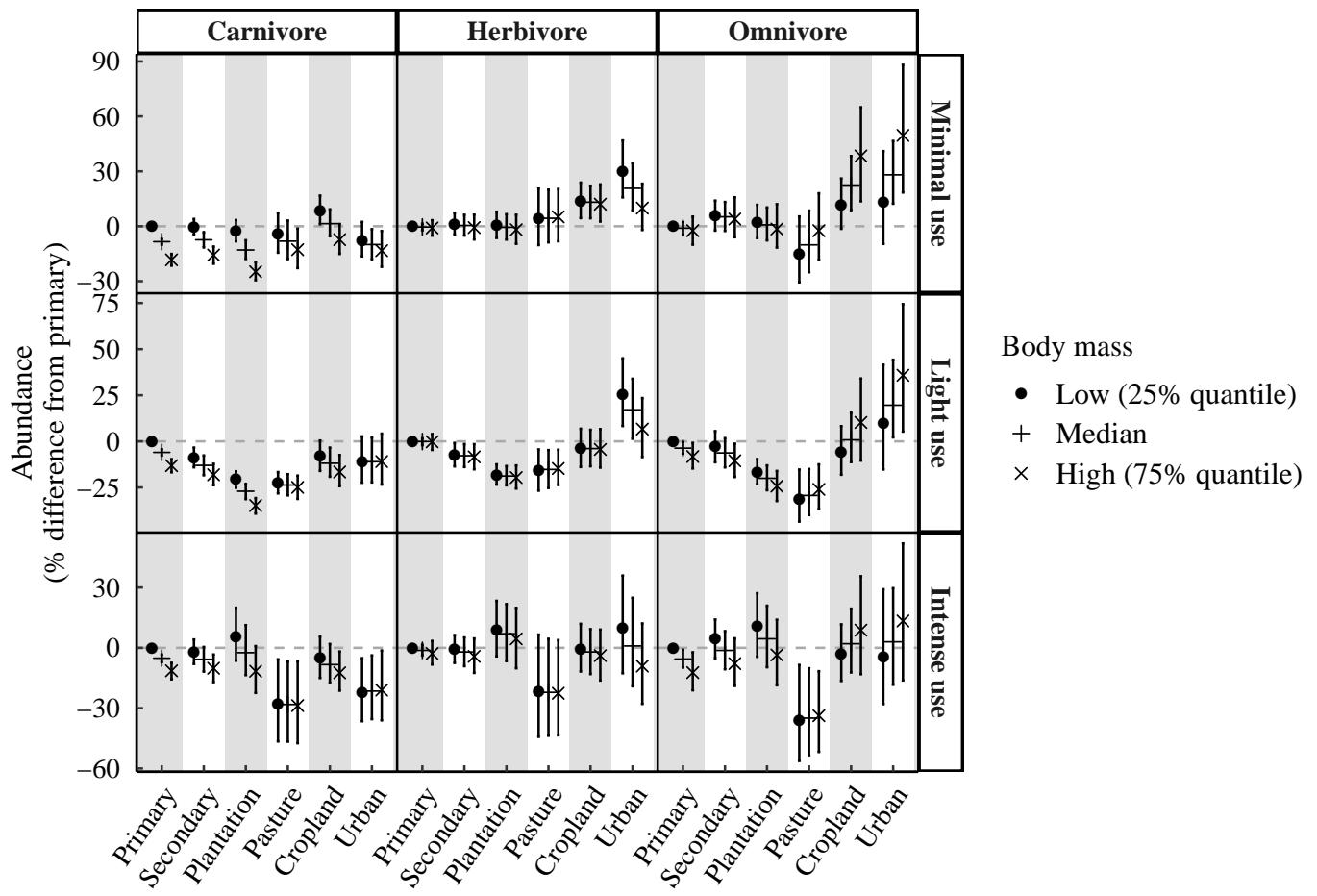
# Appendix 4: Supporting information for Chapter 5

**Table S5.1:** Number of species for which RMR data was collected in each class & data sources, initial RMR coverage for PREDICTS species and phylogenetic signal in RMR.

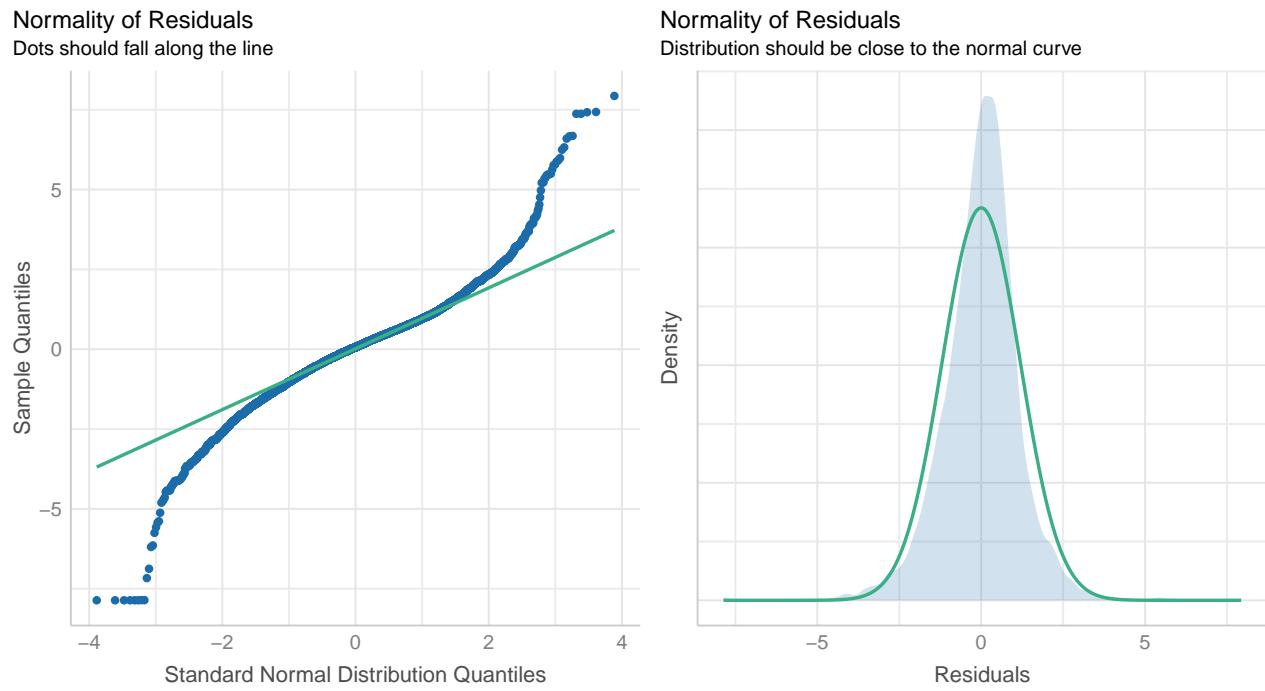
Class	RMR data	Coverage for PREDICTS species	Phylogenetic signal (Pagel's $\lambda$ , $\pm 95\%$ CI)
<b>Amphibians</b>	126 species from Stark et al. (2020)	16/379 species (4%)	0.89 (0.86-0.91)
<b>Birds</b>	719 species from McNab (2009) Fristoe et al. (2015) Londoño et al. (2015) Stark et al. (2020)	317/3129 species (10%)	0.97 (0.95-0.98)
<b>Mammals</b>	685 species from PanTHERIA (Jones et al., 2009) Fristoe et al. (2015) Stark et al. (2020)	148/556 species (27%)	0.99 (0.98-0.99)
<b>Reptiles</b>	173 species from Stark et al. (2020)	24/329 species (7.3%)	0.90 (0.86-0.92)

$$\begin{aligned} \log(\text{Abundance}) = & \text{LU} + \text{LUI} + \text{TG} + \log(\text{BM}) + \\ & \text{LU:LUI} + \text{LU:TG} + \text{LU:log(BM)} + \text{LUI:TG} + \text{LUI:log(BM)} + \text{TG:log(BM)} + \\ & \text{LU:TG:log(BM)} + \text{LUI:TG:log(BM)} \end{aligned}$$

**Figure S5.1:** Formula for the abundance model, used to understand the role of shifts in the body mass of species on observed changes in tRMR (see main text, ‘Disentangling the effects of body mass and abundance on tRMR’). I fitted a model to explain changes in species abundance (given presence) by land use, land-use intensity, trophic group, body mass and their interactions. The model included all two-way interactions among these predictors. To account for potential differences in the slope of the relationship between abundance and body mass among the different trophic groups, I also included two three-way interactions in the model (among land use, trophic group and body mass; and among land-use intensity, trophic group and body mass). Random effects included study, site and species identity. LU: land use; LUI: land-use intensity; TG: trophic group; BM: body mass.

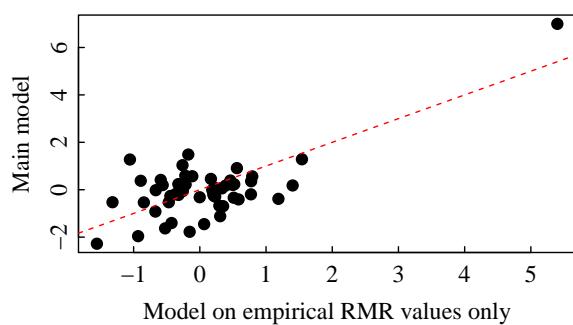


**Figure S5.2: Effects of land use, land-use intensity, trophic group, body mass and their interactions on assemblage-level total abundance**, estimated from the model specified in Fig. S1. The predictions are rescaled with reference to minimally used primary vegetation, considered to be the undisturbed baseline. Primary: primary vegetation; secondary: secondary vegetation; plantation: plantation forest. For visualisation purposes, I plotted the predictions for three body mass levels (but body mass was considered as a continuous variable in the model).

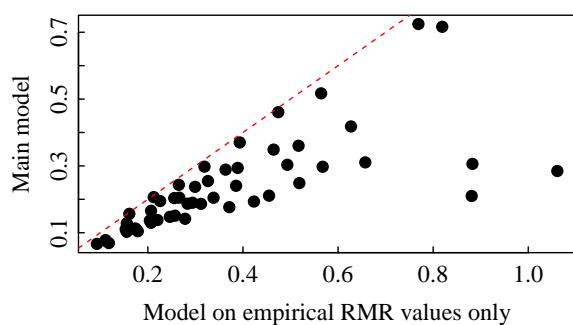


**Figure S5.3:** Diagnostic plots (qq-plot and residual distribution) for the linear mixed-effects model looking at the effects of land use, land-use intensity, trophic group and their interactions on assemblage-level total RMR. The diagnostic plots were obtained with the ‘performance’ R package (Lüdecke et al., 2021).

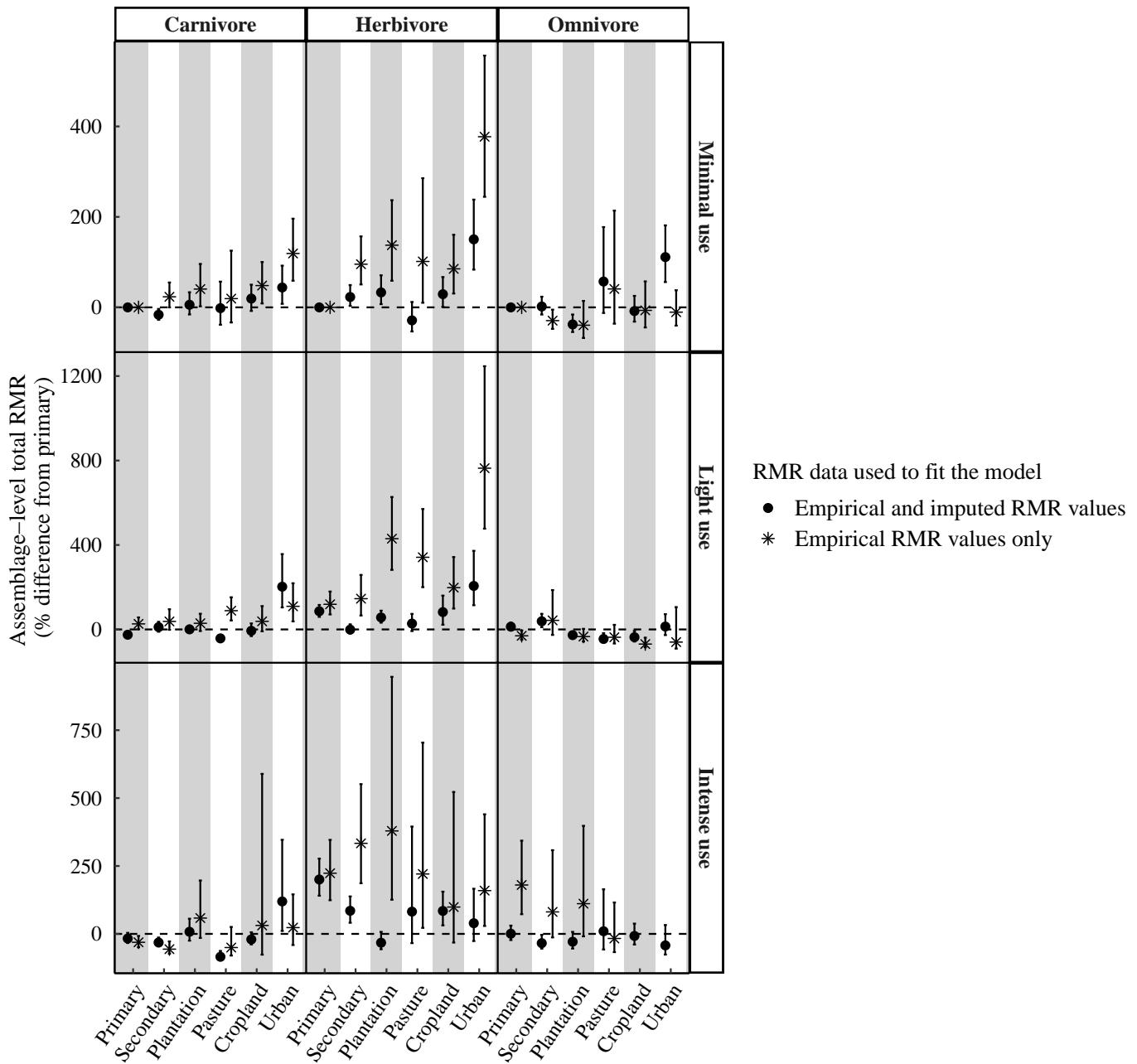
**(a) Coefficient estimates**



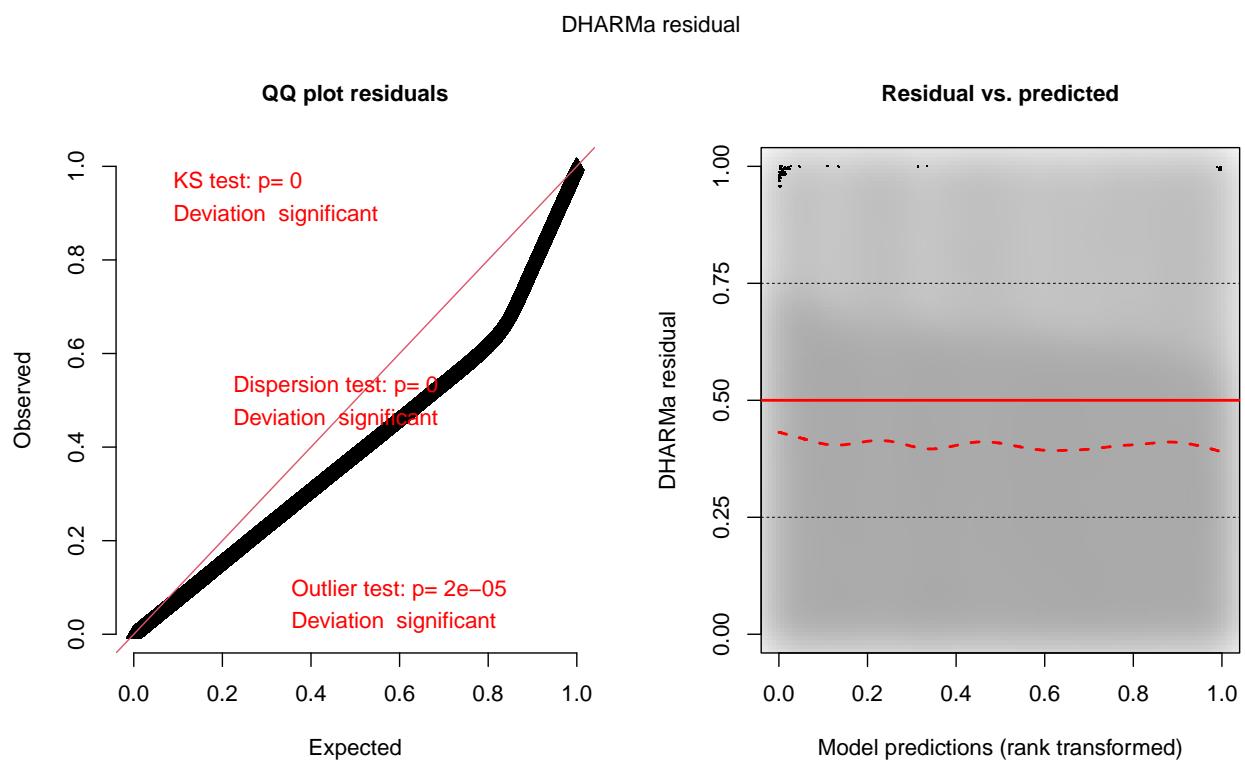
**(b) Standard error estimates**



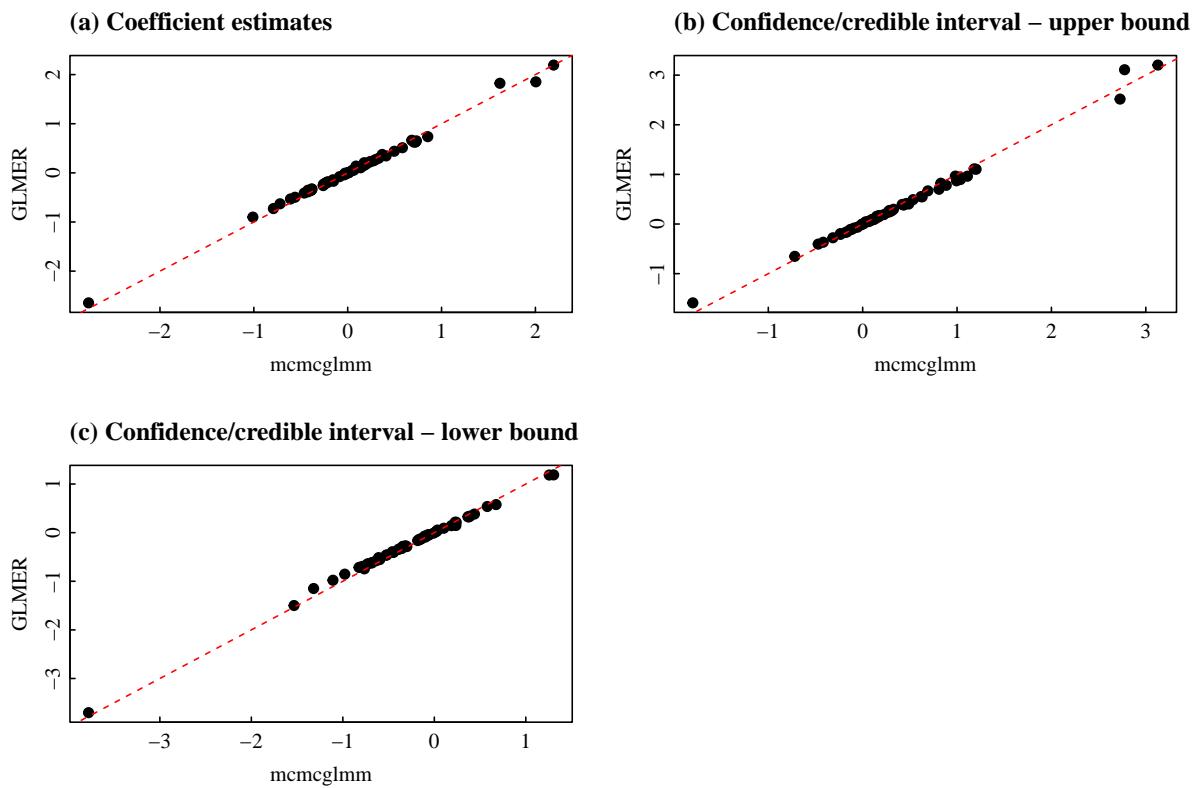
**Figure S5.4: Estimates for the models looking at the effects of land use, land-use intensity, trophic group and their interactions on assemblage-level total RMR.** I plotted the estimates from the model fitted on the empirical and imputed RMR values (presented in the main text) on the y-axis, and the estimates from the model fitted on the empirical RMR values only on the x-axis.



**Figure S5.5: Effects of land use, land-use intensity, trophic level and their interactions on assemblage-level total RMR,** estimated from the model fitted on the empirical and imputed RMR values (presented in the main text) and from the model fitted on the empirical values only. The predictions are rescaled with reference to minimally used primary vegetation, considered to be the undisturbed baseline. Primary: primary vegetation; secondary: secondary vegetation; plantation: plantation forest.

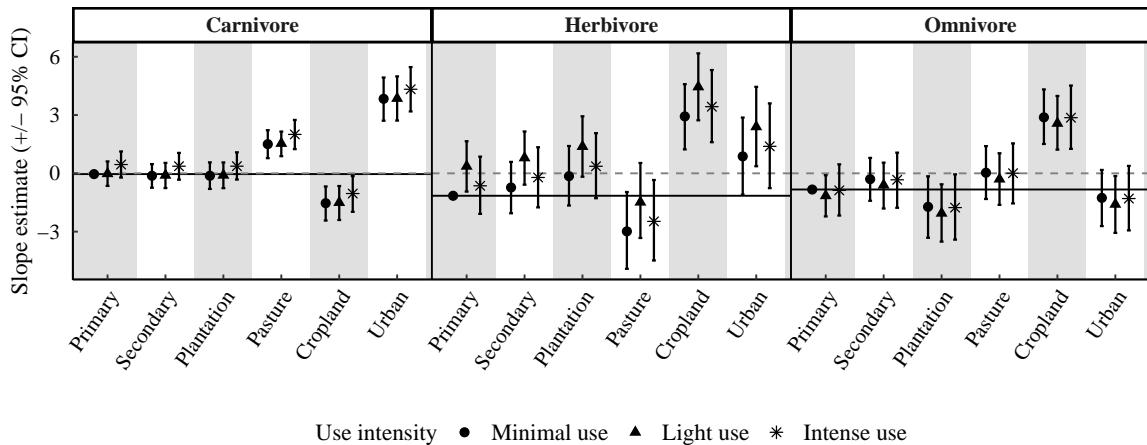


**Figure S5.6:** Diagnostic plots for the generalised mixed-effects model looking at the effects of land use, land-use intensity, trophic group, residual RMR and their interactions on species' probability of occurrence. The diagnostic plots were obtained with the 'DHARMA' R package (Hartig, 2021).

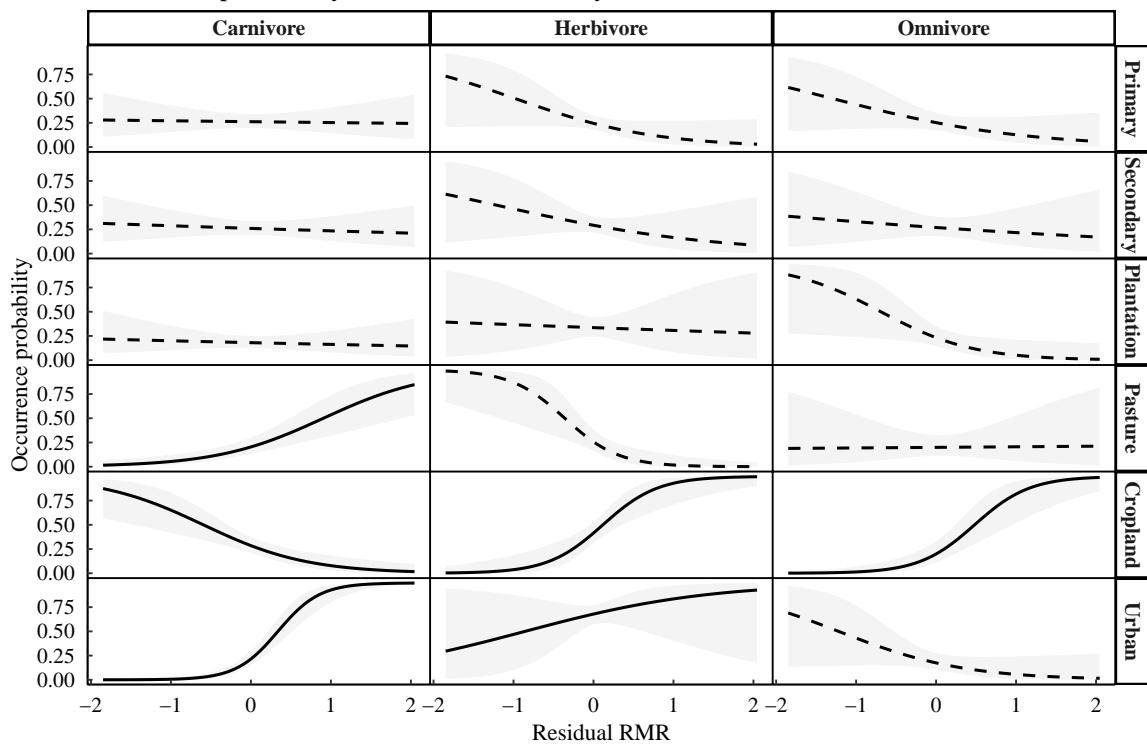


**Figure S5.7:** Model's coefficients from the occurrence model fitted using the **lme4** package (Bates et al., 2015) against coefficients from the model fitted using a Bayesian framework with the **MCMCglmm** package (Hadfield, 2010). The models were fitted to investigate the effects of land use, land-use intensity, trophic group and residual RMR on species occurrence probability.

(a) Slope of the relationships between occurrence probability (log-odds) and residual RMR



(b) Predicted effects of residual RMR on occurrence probability (minimal use intensity)



**Figure S5.8:** (a) Slope estimates for the relationship between residual RMR and occurrence probability in each land-use type and for the three levels of land-use intensity, from the model fitted using the empirical RMR values (i.e., excluding imputed RMR values). The black horizontal line indicates the slope for the reference level (primary vegetation) for minimal land-use intensity. The grey dashed line marks 0. Error bars are 95% confidence intervals. (b) Effect of residual RMR on species probability of occurrence within each trophic level and for each land use-type. I plotted the predictions for minimal land-use intensity only. Solid lines represent significant relationships. Primary: primary vegetation; secondary: secondary vegetation; plantation: plantation forest.

# Bibliography for the Appendices

- Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. DOI: 10.18637/jss.v067.i01.
- Borges, R., Machado, J. P., Gomes, C., Rocha, A. P., and Antunes, A. (2018). Measuring phylogenetic signal between categorical traits and phylogenies. *Bioinformatics*. DOI: 10.1093/bioinformatics/bty800.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A., and Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*. DOI: 10.1371/journal.pcbi.1003537.
- Bruggeman, J., Heringa, J., and Brandt, B. W. (2009). PhyloPars: Estimation of missing parameter values using phylogeny. *Nucleic Acids Research*. DOI: 10.1093/nar/gkp370.
- Buuren, S van and Oudshoorn, C. G. M. (2007). MICE: multivariate imputation by chained equations. *R package version*.
- Cadotte, M. W., Carscadden, K., and Mirochnick, N. (2011). Beyond species: Functional diversity and the maintenance of ecological processes and services. *Journal of Applied Ecology*. DOI: 10.1111/j.1365-2664.2011.02048.x.
- Cailliez, F. (1983). The analytical solution of the additive constant problem. *Psychometrika*. DOI: 10.1007/BF02294026.
- Chamberlain, S. (2018). *rredlist: 'IUCN' Red List Client*. R package version 0.5.0.
- Chamberlain, S. A. and Szöcs, E. (2013). taxize : taxonomic search and retrieval in R [version 2; referees: 3 approved]. *F1000Research*. DOI: 10.12688/f1000research.2-191.v2.
- Debastiani, V. J., Bastazini, V. A. G., and Pillar, D (2021). Ecological Informatics Using phylogenetic information to impute missing functional trait values in ecological databases. 63.April. DOI: 10.1016/j.ecoinf.2021.101315.
- Diniz-Filho, J. A. F., Bini, L. M., Rangel, T. F., Morales-Castilla, I., Olalla-Tárraga, M. Á., Rodríguez, M. Á., and Hawkins, B. A. (2012). On the selection of phylogenetic eigenvectors for ecological analyses. *Ecography*. DOI: 10.1111/j.1600-0587.2011.06949.x.
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., and Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*. DOI: 10.1111/j.1600-0587.2012.07348.x.

- Dray, S., Dufour, A.-B., and Chessel, D. (2007). The ade4 Package – II: Two-Table and K-Table Methods. *R News*, 7.2, 47–52.
- Etard, A., Morrill, S., and Newbold, T. (2020). Global gaps in trait data for terrestrial vertebrates. *Global Ecology and Biogeography*, November 2019, 1–16. DOI: 10.1111/geb.13184.
- Faurby, S., Davis, M., Pedersen, R., Schowanek, S. D., Antonelli1, A., and Svenning, J.-C. (2018). PHYLACINE 1.2: The Phylogenetic Atlas of Mammal Macroecology. *Ecology*, 99.11, 2626–2626. DOI: 10.1002/ecy.2443.
- Faurby, S., Pedersen, R., Davis, M., Schowanek, S. D., Jarvie, S., Antonelli, A., and Svenning, J.-C. (2020). PHYLACINE 1.2.1: An update to the Phylogenetic Atlas of Mammal Macroecology. DOI: 10.5281/zenodo.3690867.
- Fox, J. and Monette, G. (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association*. DOI: 10.1080/01621459.1992.10475190.
- Fristoe, T. S., Burger, J. R., Balk, M. A., Khaliq, I., Hof, C., and Brown, J. H. (2015). Metabolic heat production and thermal conductance are mass-independent adaptations to thermal environment in birds and mammals. *Proceedings of the National Academy of Sciences of the United States of America*, 112.52, 15934–15939. DOI: 10.1073/pnas.1521662112.
- Hadfield, J. D. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software*, 33.2, 1–22.
- Hartig, F. (2021). *DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models*. R package version 0.4.1.
- Hudson, L. N. et al. (2014). The PREDICTS database: A global database of how local terrestrial biodiversity responds to human impacts. *Ecology and Evolution*. DOI: 10.1002/ece3.1303.
- (2017). The database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) project. *Ecology and Evolution*. DOI: 10.1002/ece3.2579.
- Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., and Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature*. DOI: 10.1038/nature11631.
- Jetz, W. and Pyron, R. A. (2018). The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. *Nature Ecology and Evolution*. DOI: 10.1038/s41559-018-0515-5.
- Johnson, T. F., Isaac, N. J., Paviolo, A., and González-Suárez, M. (2021). Handling missing values in trait data. *Global Ecology and Biogeography*. DOI: 10.1111/geb.13185.
- Jones, K. E. et al. (2009). PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology*. DOI: 10.1890/08-1494.1.
- Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., and Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, 6.60, 3139. DOI: 10.21105/joss.03139.
- Londoño, G. A., Chappell, M. A., Castañeda, M. d. R., Jankowski, J. E., and Robinson, S. K. (2015). Basal metabolism in tropical birds: Latitude, altitude, and the 'pace of life'. *Functional Ecology*, 29.3, 338–346. DOI: 10.1111/1365-2435.12348.

- McNab, B. K. (2009). Ecological factors affect the level and scaling of avian BMR. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, 152.1, 22–45. DOI: 10.1016/j.cbpa.2008.08.021.
- Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*. DOI: 10.1038/44766.
- Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M., Young, B. E., Graham, C. H., and Costa, G. C. (2014). Imputation of missing data in life-history trait datasets: Which approach performs the best? *Methods in Ecology and Evolution*. DOI: 10.1111/2041-210X.12232.
- Rabosky, A. D., Cox, C., Rabosky, D., Title, P., Holmes, I., Feldman, A., and McGuire, J. (2016). Coral snakes predict the evolution of mimicry across New World snakes. *Nature Communications*, 7, 11484.
- Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3, 217–223.
- Santini, L., Benítez-López, A., Ficetola, G. F., and Huijbregts, M. A. J. (2018). Length–mass allometries in amphibians. *Integrative Zoology*, 13.1, 36–45. DOI: 10.1111/1749-4877.12268.
- Santos, T. (2018). Package ‘PVR’. Phylogenetic Eigenvectors Regression and Phylogenetic Signal-Representation Curve.
- Stark, G., Pincheira-Donoso, D., and Meiri, S. (2020). No evidence for the ‘rate-of-living’ theory across the tetrapod tree of life. *Global Ecology and Biogeography*, 29.5, 857–884. DOI: 10.1111/geb.13069.
- Stekhoven, D. J. (2016). Nonparametric Missing Value Imputation using Random Forest. *R Package version 1.4*. DOI: 10.1093/bioinformatics/btr597.
- Stekhoven, D. J. and Bühlmann, P. (2012). Missforest-Non-parametric missing value imputation for mixed-type data. *Bioinformatics*. DOI: 10.1093/bioinformatics/btr597.
- Tonini, J. F. R., Beard, K. H., Ferreira, R. B., Jetz, W., and Pyron, R. A. (2016). Fully-sampled phylogenies of squamates reveal evolutionary patterns in threat status. *Biological Conservation*. DOI: 10.1016/j.biocon.2016.03.039.
- Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., and Altman, R. B. (2001). Missing value estimation methods for DNA microarrays. *Bioinformatics*. DOI: 10.1093/bioinformatics/17.6.520.