



**The University of Edinburgh**

**School of Geosciences**

**GLOBAL ASSESSMENT OF LOCAL BIODIVERSITY  
CHANGES REVEALS COMPLEX RESPONSES TO  
LARGE-SCALE HUMAN ACTIVITY**

By

**DANIELA GARGYA**

in partial fulfilment of the requirement  
for the degree of BSc with Honours  
in Ecological and Environmental Sciences with Management

May 2020

## Summary

Globally, biodiversity is changing in complex ways in a time of accelerating human impact. We lack real-world data how cumulative large-scale anthropogenic activity, as a prominent contemporary global change driver, is influencing the re-shuffling of ecological communities over time. Despite calls for more encompassing tests, research has mostly focussed on individual types of environmental change. Here, I quantify the influence of multiple types of human impacts captured with the proxy accessibility to cities on ecological assemblages worldwide by analysing change in 5787 time-series across birds, mammals, plants and invertebrates. Surprisingly, I found that for every 10% increase in accessibility, species replacement (turnover) decreased by 1.6 %. I further highlight the need to account for both direct and indirect effects of human activities. Turnover was highest for invertebrates and mammals and lowest for birds and plants. My findings demonstrate the complexity of real-world biodiversity change and challenge the assumption that current conservation focus on protecting wilderness areas might not be sufficient to achieve biodiversity composition conservation goals.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Background . . . . .	1
1.2	Objectives and research questions . . . . .	3
1.3	Research hypotheses and predictions . . . . .	4
<b>2</b>	<b>Methods</b>	<b>6</b>
2.1	Databases . . . . .	6
2.2	Data processing . . . . .	8
2.3	Data analysis . . . . .	9
<b>3</b>	<b>Results</b>	<b>10</b>
3.1	Temporal turnover and accessibility (Research Question 1) . . . . .	10
3.2	Temporal turnover and human population density (Research Question 2 modified) . . . .	11
3.3	Temporal turnover and taxa (Research Question 3 modified) . . . . .	12
<b>4</b>	<b>Discussion</b>	<b>14</b>
4.1	Overview . . . . .	14
4.2	Temporal turnover and accessibility (Research Question 1) . . . . .	15
4.3	Temporal turnover and human population density (Research question 2 modified) . . . .	16
4.4	Temporal turnover and taxa (Research Question 3 modified) . . . . .	17
4.5	Study Limitations . . . . .	17
4.6	Future directions . . . . .	18
<b>5</b>	<b>Conclusions</b>	<b>19</b>
<b>6</b>	<b>References</b>	<b>20</b>

<b>7 Appendices</b>	<b>26</b>
7.1 Appendix A: Sensitivity analyses . . . . .	26
7.2 Appendix B: Additional figures . . . . .	27
7.3 Appendix C: Model outputs . . . . .	29
7.4 Appendix D: R Code . . . . .	31

## Acknowledgements

I would like to thank my supervisors, Dr Isla Myers-Smith and Gergana Daskalova for their constant support, shared knowledge and encouragement throughout the process of this dissertation and beyond. From finding a research question, to help with statistical analysis, coding, and writing tips, I always found open ears and patient explanations. Also, I am very grateful for the opportunity to have become part of Team Shrub, where I learned a lot during the weekly lab meetings while also having fun. The feedback Team Shrub provided me on one of my paragraphs was very useful.

I also want to acknowledge and thank my great support network. Starting with my wonderful flatmates, Anna, Larisa and Giada, for always keeping the spirits high, but also giving valuable feedback and proofreading drafts. Also, I want to thank my parents and Max, for constant support and proofreading the draft.

## List of abbreviations

BACI Before after control impact

CI Confidence Interval

RQ Research question

# 1 Introduction

## 1.1 Background

Humans have modified 75% of terrestrial land cover, leading to major pressures on the biosphere and its inhabitants (IPBES, 2019). Growth of human population and the expansion of cities and road networks, which predicate UN's equity agenda by advancing accessibility (UN, 2015), are expected to intensify pressures on the natural environment (IPBES, 2019). Large-scale human activity has altered abundance, richness and composition of ecological assemblages in complex ways (Dornelas et al., 2014; Vellend, Baeten, et al., 2017; Hillebrand et al., 2018; Magurran et al., 2018). Even when consequences on ecosystem processes and resilience are not fully understood (Vellend, Baeten, et al., 2017), biodiversity is increasingly unstable (Eriksson and Hillebrand, 2019). Disentangling drivers of complex biodiversity change and establishing effective conservation practices without compromising human development is an urgent issue. Currently, we only have a limited quantitative understanding of how global change drivers, such as large-scale human activity, influence local patterns of biodiversity over time. Despite calls for more encompassing tests of the effects of global change drivers on biodiversity changes (Sirami et al., 2017; Mazor et al., 2018), research has mostly focussed on individual types of environmental change such as forest loss (Daskalova et al., 2019) and land-use change (Newbold et al., 2015). Recent global databases of large-scale human activity such as the accessibility to cities maps (Weiss et al., 2018), together with long-term biodiversity time-series (Dornelas et al., 2018) allow us to quantitatively test more encompassing global change drivers' effects on biodiversity change. Biodiversity change linked to large-scale human activity can provide the information and predictions needed for better international policy making in the light of our rapidly changing Anthropocene.

Biodiversity changes depend on the scale and metric of observation (McGill et al., 2015; Chase et al., 2019). On a global scale, biodiversity richness is declining as is reflected in increasing global extinction rates (Barnosky et al., 2011). Contrary, on a local scale, biodiversity richness seems to be stable or even increasing (Butchart et al., 2010; Dornelas et al., 2014; Vellend, Dornelas, et al., 2017). A potential explanation can be found when looking at species turnover (changes in the composition of ecological communities), which can reveal changes in biodiversity that are often uncoupled from changes in species richness (Dornelas et al., 2014; Magurran et al., 2018). Communities seem to undergo replacement of some species by others with no overall change in total number. Although turnover is part of natural community persistence processes to some extent, current levels exceed baseline trends of existing ecological models (Dornelas et al., 2014; Magurran et al., 2018). This

highlights that the biodiversity crisis is not about decline but rather about large-scale reorganization of communities leading to homogenization across space (Blowes et al., 2019). Homogeneity has serious implications on the ability of species communities to adapt to future environmental changes (Eriksson and Hillebrand, 2019). Thus, to comprehend biodiversity change across scales and make predictions for the future, it is important to understand changes in communities' compositions at the local scale.

Biodiversity change and its attribution to a specific driver can be assessed through different methods, each with its own advantages and disadvantages (De Palma et al., 2018). Based on differences in methods used, there is ongoing controversy of local biodiversity changes driven by human activity. Research that used space-for-time (Newbold et al., 2015) and modelling approaches (Newbold et al., 2014) generally report negative trends of biodiversity change following human impact, such as land-use change. Similarly, absence of human activity is reported as beneficial in reducing extinction rates (Di Marco et al., 2019). Research that uses temporal trends and real-world data reports a more complex picture of biodiversity change with both negative and positive trends co-occurring (Dunic et al., 2017; Daskalova et al., 2019). Space-for-time models can utilise much data but have been criticised for the uncertainty in the assumption that time can be substituted for space (De Palma et al., 2018). Especially when researching biodiversity change, this approach is lacking important temporal aspects such as ecological lags (Daskalova et al., 2019) and community self-regulation (Gotelli et al., 2017), with the consequent risk of overestimating the effect of global change drivers. Temporal assessments based on real world data can infer driver-response relationships by observing fluctuations through time. These attributions heavily depend on long-term availability and matching of data of both driver and response variables under observation (De Palma et al., 2018). Assessing biodiversity changes and their relationship to global change drivers benefits from looking at real-world temporal data.

Biodiversity change can be driven by multiple human impacts simultaneously. Humans have impacted the natural environment through a variety of activities such as habitat change, exploitation, pollution, climate change and introduction of invasive species (Pereira, Navarro and Martins, 2012; IPBES, 2019). Previous studies have focussed on individual drivers of biodiversity change, such as forest loss (Daskalova et al., 2019) and land-use transitions (Newbold et al., 2014, 2015), ignoring the fact that several drivers can be present at the same time. Therefore, these studies are limited in their extent to capture potential interactions of simultaneous human activities. To implement effective conservation measures it is not sufficient to look at one driver in isolation, but we seek to gain a more general understanding of biodiversity changes in reaction to large-scale human activities acting together (Mazor et al., 2018). The accessibility to cities metric, a measure of the closest travel time to the next urban centre, can be seen as a proxy for multiple human influences. An increase in accessibility is tightly



linked with increases of road network and urbanisation (Weiss et al., 2018). This metric can capture changes to the environment and its inhabitants such as habitat fragmentation, land-use change, alien species and habitat loss as a cumulative measure. It allows to observe outcomes of potential additive, synergistic or antagonistic interactions between such environmental changes, as they are likely to co-occur (Bowler et al., 2020) and interact (Chazal and Rounsevell, 2009; Sirami et al., 2017). Not only multiple drivers, but also the magnitude of exposure can influence biodiversity changes (Bowler et al., 2020). It can be hypothesized that biodiversity changes might be amplified with higher exposure to human activities, driven by human population density (Geary et al., 2019). Biodiversity changes are influenced by co-occurring drivers of human activities with potential interaction effects and are thus captured better with a metric that accounts for a variety of human activities.

Biodiversity change is dependent on both the exposure to various drivers and species responses to environmental change (Foden et al., 2019). Species' vulnerability varies with functional traits such as life history, range-size and niche breadth (Purvis et al., 2000; Isaac and Cowlshaw, 2004). Examining them separately for taxa can improve conservation strategy targeting (Longton and Hedderson, 2000). Species with small population sizes, high habitat specificity and restricted geographic extent are more likely to go extinct locally (Kareiva et al., 1990; Lande, 1993; Newbold et al., 2014), thus increase the likelihood of biodiversity change. Less mobile species might struggle to disperse and find a suitable refuge, as the exposure to environmental change increases (Concepción et al., 2015; Ehrlén and Morris, 2015). Small population sizes are more prone to undergo stochastic fluctuation, as well as higher inbreeding rates with resulting reduced fitness (Kilpatrick and Ives, 2003). High habitat specificity means that colonization of new areas is restricted by limited habitat preferences (Gaston and Fuller, 2008). Species traits associated with higher vulnerability shared across a taxon make it more prone to declines and extinctions. Studies looking at different species responses to a range of human impacts have found both positive and negative responses among taxa, related to functional traits (Newbold et al., 2014; Daskalova, Myers-Smith and Godlee, 2019). Understanding biodiversity changes on a taxa-specific level allows for more specific targeting of conservation efforts.

## **1.2 Objectives and research questions**

My aim is to quantify how global change drivers influence ecological communities over time. Specifically, I will focus on large-scale human activity captured in the accessibility to cities, and human population density and their relationship with temporal changes in community composition overall and across taxa. To achieve this, I investigate the influence of accessibility to cities on ecological assem-

blages worldwide and across taxa (birds, mammals, plants and invertebrates) by analysing change in 5787 time series, addressing the following research questions (RQ):

1. Do sites with higher accessibility to cities experience more changes in assemblage composition over time (temporal turnover) than locations with lower accessibility?
  - How does the duration of ecological monitoring influence the magnitude of detected temporal turnover trends?
2. How is temporal turnover influenced by an interaction between human population density and accessibility?
3. How does temporal turnover of assemblage composition respond to levels of accessibility across taxa (birds, mammals, terrestrial invertebrates, terrestrial plants)?

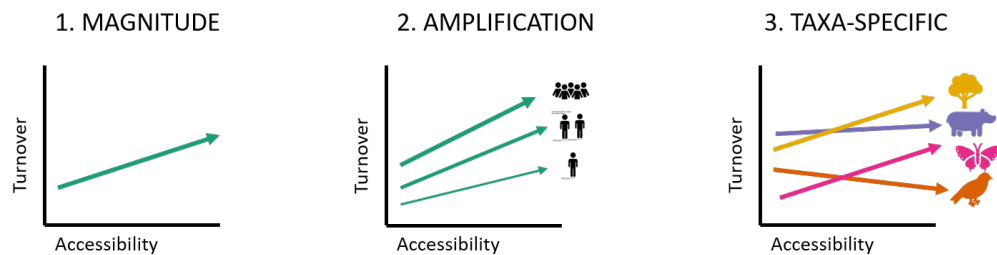
### **1.3 Research hypotheses and predictions**

I hypothesize that there will be a positive relationship between higher accessibility to cities and greater temporal turnover. The magnitude of temporal turnover increases for sites which have been monitored over longer durations. The relationship between accessibility to cities and temporal turnover is steeper when human population density is higher. When looking at individual taxa, I hypothesize that taxa responses will vary for temporal turnover across levels of accessibility, with both positive and negative relationships present. Those hypotheses will be tested against the null hypotheses of no relationship between levels of accessibility and temporal turnover, no amplification of effect when human population density is higher and no differences in relationships between temporal turnover and accessibility across taxa.

I predict greater temporal turnover with greater magnitude of accessibility, as alterations of the natural environment caused by human activities facilitate colonisations of some species, while also leading to local extinctions. I predict greater temporal turnover in sites which have been monitored over longer durations, as the effect of turnover becomes more apparent over time. I predict the relationship between accessibility and temporal turnover to be steeper, when human population density is higher, as high human population density accelerates pressures on the natural systems, leading to higher turnover. I predict both positive and negative trends to be present among taxa, as they differ in functional traits and thus in their vulnerability to environmental changes (Figure 1).

If I find support for my null hypotheses, this will indicate that the metric accessibility cannot sufficiently capture direct and indirect impacts affecting biodiversity change, but also questions whether the current conservation emphasis to preserve wilderness areas is enough. If the results are in line with my alternative hypotheses, this will demonstrate the importance of humanly undisturbed places for the conservation of biodiversity.

#### A) Predictions



#### B) Workflow

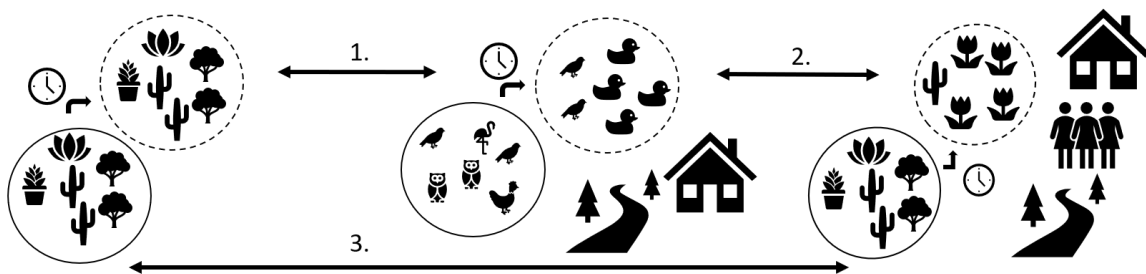


Figure 1: My study looks at three different aspects related to biodiversity change. **A** Conceptual diagrams of my predictions. **B** Workflow (figure inspired by Daskalova et al., 2019).

## 2 Methods

In this study, I assess how 5787 ecological assemblages across four taxa (birds, mammals, invertebrates and plants) respond to levels of accessibility and human population density over time. I used three global databases to quantitatively test the relationship of ecological communities and large-scale anthropogenic activity.

### 2.1 Databases

#### **BioTIME Database – biodiversity time series data**

I analysed 5787 time series from 181 different studies from 1023 unique terrestrial locations around the world covering a range of taxa, including birds, mammals, invertebrates and plants (Dornelas et al., 2018) that make up a part of the BioTIME database as of 12th of March 2020. I did not predetermine sample size but used all available data that met my inclusion criteria: part of the terrestrial realm, minimum study duration of 5 years, at least 15 studies per taxa, no more than 5000 time-series per study, time-series fixed to one location and at least 2 survey points per time-series. Time-series that had no unique coordinates for their location, were located with the coordinates given by the study they belonged to. The resulting sample sizes across space, time and taxa can be found in Figure 2.

The time-series used represent repeated studies of species abundance and identity of all species found within an ecological community. The data collection of BioTIME was consistent within studies but not between studies.

Other research, that used BioTIME to research beta diversity, partitioned studies with large extent and consequently used sample-based rarefaction to standardise sampling within each time-series (Blowes et al., 2019; Daskalova et al., 2019). My research did not conform with previous methods as 64% of my studies' areas were equal to or smaller than 1km<sup>2</sup>. 37% of studies were located in protected areas. The duration varied across time-series, ranging from 1923 – 2016, with a mean duration of 19 years (standard deviation of +/- 13 years). Mean time-series per study were 64 (+/- standard deviation of 145).

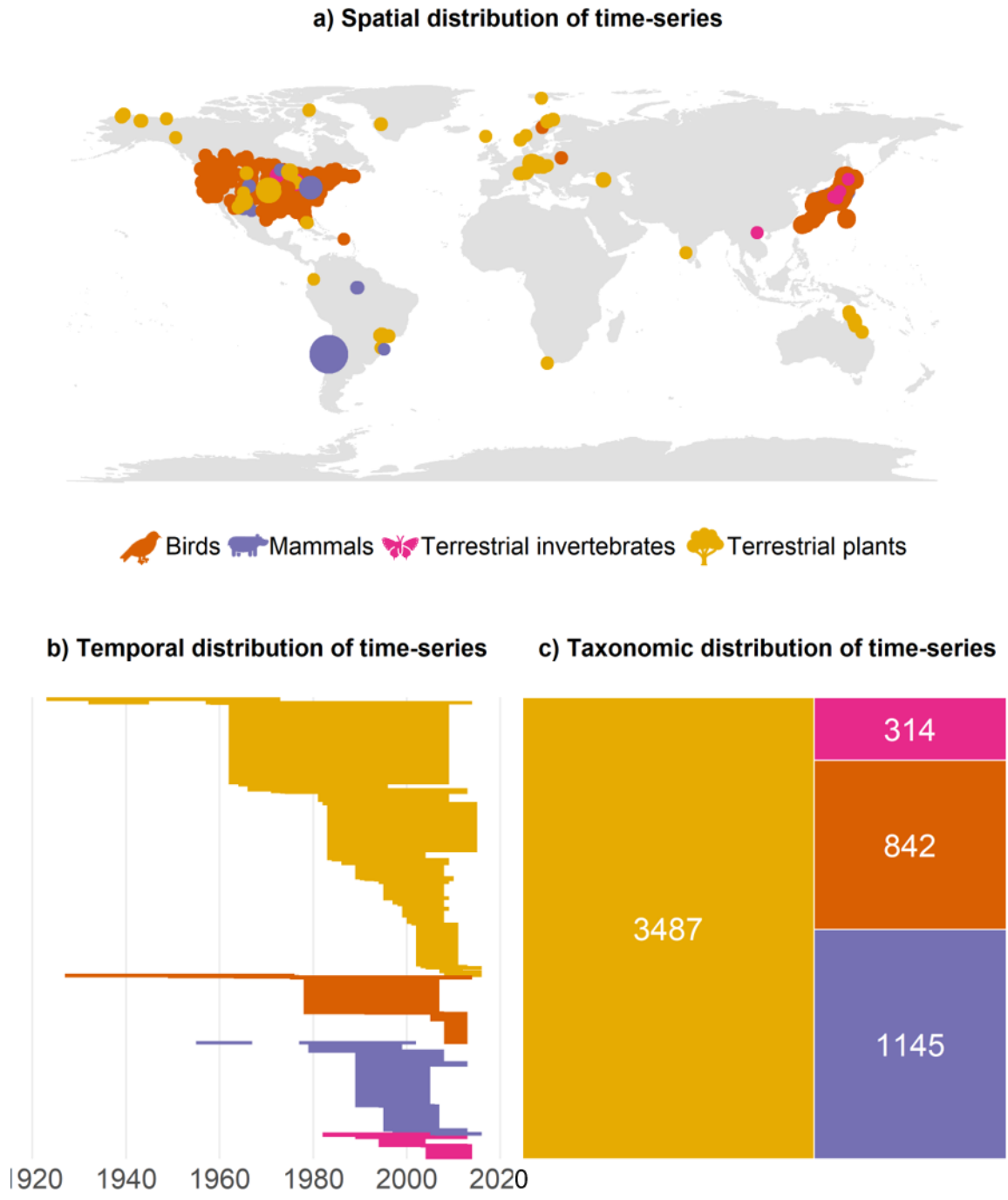


Figure 2: My study incorporated biodiversity time-series data across space, time and taxa (figure inspired by Daskalova et al., 2019).

### Accessibility to cities 2015 data

To quantify large scale human impact, I extracted an accessibility to cities metric (Weiss et al., 2018) from the Malariaatlas project Accessibility to cities 2015 global database (malariaatlas.org, 2020). This database calculates least-cost-path land-based travel time to the nearest densely populated area

(between 85 degrees north and 60 degrees south) at a 30 arc seconds resolution (approximately 1km at equator). Areas with  $>1,500$  inhabitants/km<sup>2</sup>, or a majority of built-up land cover types coincident with a population centre of at least 50,000 inhabitants, are defined as densely populated areas. This dataset is available for the nominal year of 2015.

In the main figures, I used the inverses of travel time to nearest densely populated area, so that high values correspond to higher hypothesized human effect.

### **Human population density 2015 data**

To estimate human population density, I derived data from the Gridded Population of the World, Version 4: Population Density, Revision 11 database (CIESIN, 2018). Human population density was defined as number of persons per square kilometre, based on national censuses counts and population registers. Data is available at a 30 arc-second resolution. To match the timing of the accessibility dataset, I used the dataset for the nominal year of 2015.

## **2.2 Data processing**

I conducted all data processing and analyses in R v. 3.6.1. (R Core Team, 2019).

I quantified temporal changes in community composition within sites, as the turnover component of beta diversity (species replacement rather than abundance) at the last data point available relative to the first observation data point within each time-series using {vegan} (Oksanen et al., 2019) and {betapart} (Baselga et al., 2018). Turnover, as opposed to nestedness, was the dominant part of biodiversity change in the BioTIME dataset (Dornelas et al., 2014). I converted density and biomass records into presence/absence for study records that were not count data, so no data needed to be excluded. Temporal turnover is bound between zero and one, where zero is no change in community composition and one indicates that all original species have been replaced.

I harmonized both accessibility and human population density dataset to a standard global grid size of 25km<sup>2</sup> by taking the mean value of the grid cell when extracting the values with the {raster} (Hijmans, 2020) package. I considered an area  $> 1\text{km}^2$  to be more representative of the impact-influence relationship. My analyses were not sensitive to cell sizes over 25km<sup>2</sup> (tests from 25km<sup>2</sup> to 100km<sup>2</sup>, but analysis showed higher values with similar slope for 1km<sup>2</sup>, see Figure 6 in Appendix A). I scaled the scores extracted between zero and one, where zero is not accessible/low human population density and one is very accessible/ high human population density.

To account for spatial autocorrelation, I created and assigned global grid cells to each time-series

using {dggridr} (Barnes, 2017). The grid cells covered areas of resolution 12, which is equivalent to an area of around 100km<sup>2</sup>.

## 2.3 Data analysis

To quantify the relationship of accessibility and human population density across taxa on temporal turnover of ecological communities, I used mixed effects models in a Bayesian modelling framework based on a Stan computational framework and accessed through the {brms} package (Bürkner, 2018). I based the models on a zero one inflated beta distribution to reflect the properties of turnover (bound between, and including, zero and one). I assumed a Bernoulli distribution for values of zero and one, and a beta error distribution for values between zero and one.

I modelled temporal turnover as my response variable. Fixed effects were accessibility (proportion) and duration of the time-series (RQ 1). I did not consider centring of duration necessary as the variation was relatively little within. Model convergence could not be achieved with the interaction term between accessibility and human population density (proportion), so latter was added as a fixed effect without an interaction (RQ 2). Neither could I achieve model convergence when allowing each taxon to have a random slope (nor intercept) with accessibility, so I added taxa as a fixed effect (RQ 3). I adjusted research question 2 and 3 accordingly when presenting my results and discussion. I included Study ID as a random effect to account for confounding effects within the same study as the methods were consistent within but not among studies. I added the global grid cell variable as a random effect to account for spatial autocorrelation. I used ggpredict from {ggeffects} (Lüdtke, 2018) to calculate model predictions. I assumed the said model set-up to be the optimal balance between requirements of the data and feasibility.

I used the default priors which are weakly informative as I had no ecological explanation to assume otherwise. I ran the models with 4000 iterations, with a warmup of 1000 iterations on 4 chains. I assessed convergence visually by examining trace plots and Rhat values.

As I used a Bayesian framework, my inferences were based on the posterior distribution of each fixed effect. I considered the effects significant if the lower and upper 95% credible intervals (CI) did not overlap zero. I concluded all results, regardless of the direction or magnitude of the effect size.

The full R script, the link to my GitHub repository and preregistration, and description of sensitivity analyses can be found in the appendices.

### 3 Results

#### 3.1 Temporal turnover and accessibility (Research Question 1)

Contrary to my predictions, I found that temporal turnover decreased as accessibility increases across the 5787 time-series surveyed (slope = -1.20, CI = -2.09 to -0.32, Figure 3, see Table 1 for full model outputs). On average, for every 10% increase in accessibility, turnover decreases by 1.6 %. Notably, when taking the last census point as a baseline (lowest point of slope located at high accessibility), the ecological assemblages across levels of accessibility have experienced average 24% replacement of their species (as turnover is bound between 0 and 1) over their time-period monitored (average duration of 19 years). 63% of time-series experienced some turnover (turnover > zero) and only ~ 7% of time-series analysed experienced complete turnover of communities (turnover = one). As 90.5% of data points had an accessibility score of 0.9 and above, there is higher uncertainty around the main effect at lower values of accessibility than at higher values, indicating the need to take the high value of temporal turnover at low accessibility with caution.

In line with my predictions, the duration of the observation influenced the magnitude of the detected temporal turnover trends, with higher temporal turnover observed for longer monitoring duration of sites (Figure 3, also see Figure 12 in appendix B).



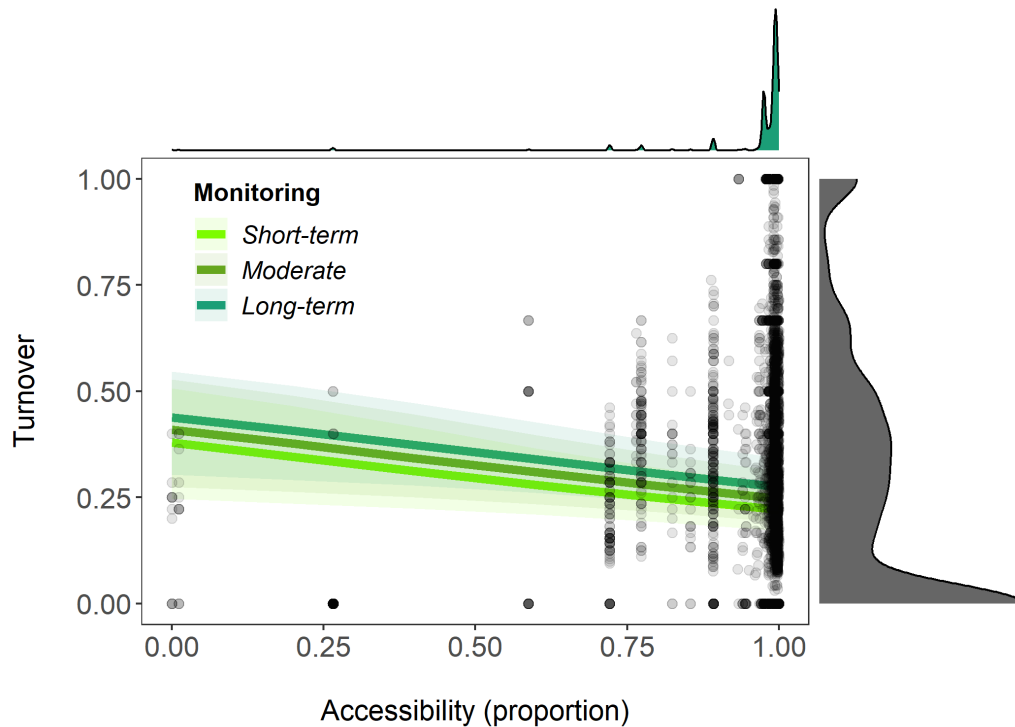


Figure 3: Temporal turnover decreased as accessibility increases across the 5787 time-series surveyed using a Bayesian mixed model (slope = -1.20, CI = -2.09 to -0.32, see Table 1 for full model outputs). Grey points represent raw data. Lines and error bands represent model predictions and 95 percent credible intervals, respectively. Colour coding of lines represent differing duration of monitoring of time-series. The marginal density plots represent accessibility (top density plot) and temporal turnover (right density plot) distributions across all time-series surveyed. See model Rhat values in appendix confirming model convergence.

### 3.2 Temporal turnover and human population density (Research Question 2 modified)

Looking at the relationship between temporal turnover and human population density (without the interaction as that model did not converge), I found that temporal turnover neither increased nor decreased as human population density increases across the 5787 time-series surveyed (slope = -0.22, CI = -0.49 to 0.05, Figure 4, see Table 1 for more model outputs). Notably, when taking the last census point as a baseline (lowest point of slope located at high human population density), the ecological assemblages have experienced 22% replacement of their species (as turnover is bound between 0 and 1) over their time-period monitored (average duration of 19 years). Contrary to the distribution of accessibility falling into higher spectrum, 96.4% of data points have human population density score of 0.1 and below.

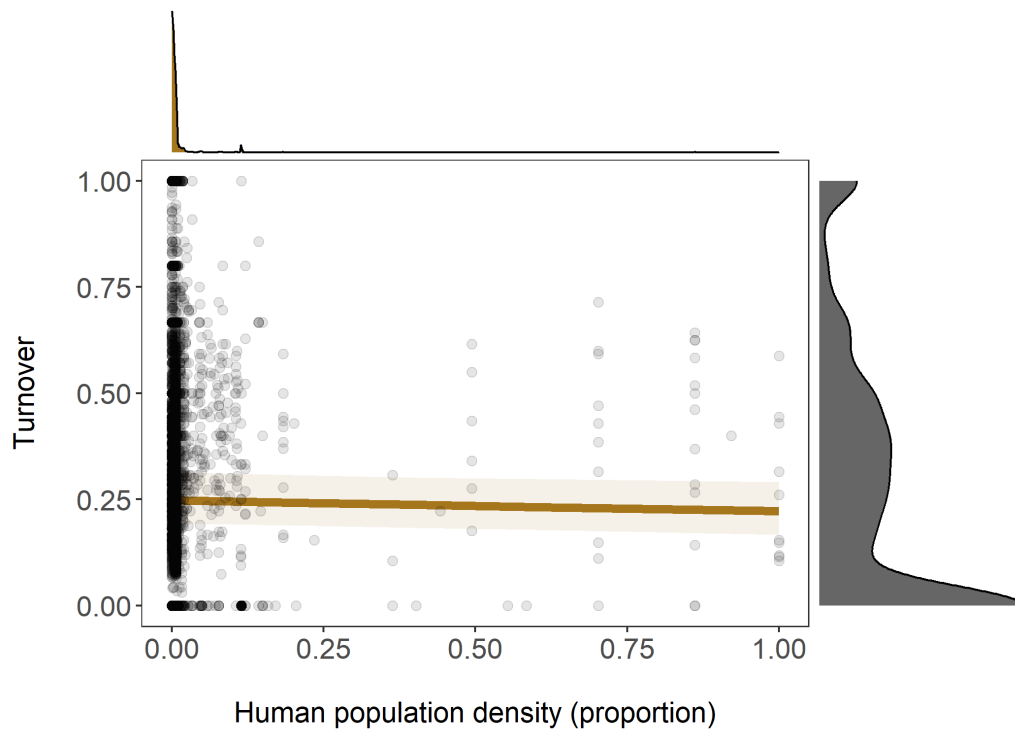


Figure 4: Temporal turnover showed no directional relationship with human population density across the 5787 time-series surveyed using a Bayesian mixed model (slope = -0.22, CI = -0.49 to 0.05, see Table 1 for full model outputs). Grey points represent raw data. Lines and error bands represent model predictions and 95 percent credible intervals, respectively. The marginal density plots represent human population density (top density plot) and temporal turnover (right density plot) distributions across all time-series surveyed. See model Rhat values in appendix confirming model convergence.

### 3.3 Temporal turnover and taxa (Research Question 3 modified)

Looking at the relationship between temporal turnover and taxa (not considering the relationship across levels of accessibility as that model did not converge), I found that all taxa experienced temporal turnover (see Figure 5, Table 1 for full model outputs). Terrestrial invertebrates (slope = 0.99, CI = 0.31 – 1.7,  $n = 314$ ) and mammals (slope = 0.71, CI = 0.01 – 1.42,  $n = 1145$ ) experienced higher amounts of turnover relative to birds (slope = 0.11, CI = -0.87 – 1.14,  $n = 841$ ) and plants (slope = 0, CI = -0.56 – 0.55,  $n = 3487$ ). Terrestrial invertebrates experienced highest amount of turnover (0.54 times more than plants). The raw data demonstrates high variability of turnover experienced within mammals and lowest variability within birds. The number of time-series sampled for each taxon differed: of all time-series analysed 60% are made up of plants, followed by 20% mammals, followed by 14% birds and only 6% invertebrates. Out of the 37% time-series that showed no turnover at all, plants accounted for 73%.

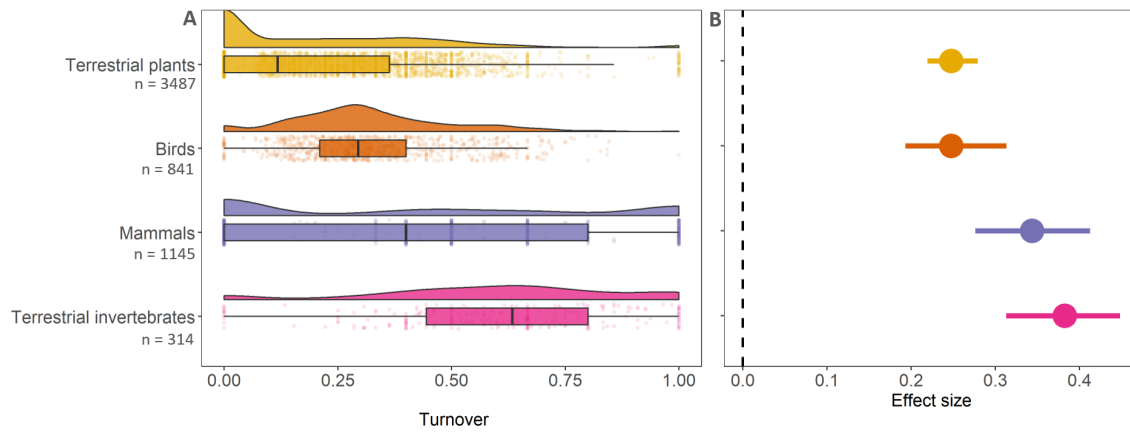


Figure 5: Temporal turnover was apparent across all taxa surveyed using a Bayesian mixed model. **A** Raw data visualisation of turnover across taxa. Split violins represent data distribution, boxplots with mean values and points are raw data. Numbers are sample sizes per taxon. **B** Model visualisation for temporal turnover and taxa. Terrestrial invertebrates (slope = 0.99, CI = 0.31 – 1.7) and mammals (slope = 0.71, CI = 0.01 – 1.42) experienced higher amounts of turnover relative to birds (slope = 0.11, CI = -0.87 – 1.14) and plants (slope = 0, CI = -0.56 – 0.55). Error bars represent 95 percentage credible intervals. Dashed line marks the zero-line threshold. See model Rhat values in appendix confirming model convergence.

## 4 Discussion

### 4.1 Overview

My global analysis of 5787 terrestrial time-series demonstrates complex heterogeneity in the relationships between large-scale human activity and biodiversity changes. Contrary to my predictions, I revealed that as accessibility increases, temporal turnover declined (Figure 3), suggesting a rejection of my alternative hypothesis of a positive relationship. This highlights the importance of analysing multiple drivers and its interacting effects on biodiversity change. Simultaneously, I demonstrated that even the lower range of predicted temporal turnover rates are above baseline levels, highlighting widespread temporal turnover across levels of human activity (Figure 3). I found a very similar baseline magnitude of temporal turnover across levels of human population density, although in this case without any directional trend (Figure 4), pointing towards the importance of capturing both direct and indirect effects of human activities. I uncovered distinct taxonomic signals of temporal turnover, with terrestrial invertebrates experiencing the highest amount, followed by mammals, whereas plants and birds showed relatively lower levels of turnover (Figure 5). The lack of sufficient data prevented me to answer my original questions of an interaction between accessibility and human population. Similarly, I could not quantify the relationship between temporal turnover and taxa across levels of accessibility which highlights the need for more comprehensive sampling in the future. On a global scale, human activities are causing a decline in biodiversity, but considering a more local scale, I uncovered heterogeneous responses to large-scale anthropogenic activity challenging the assumption that the absence of human activity is sufficient to conserve compositional biodiversity.

Global datasets have been criticised for various reasons with important implications for the interpretation of results. The first criticism concerns global biodiversity datasets often spatially underrepresenting areas modified by human activities (Gonzalez et al., 2016). This prevents an accurate reflection of relationships between human modification and environmental changes. I found that locations of biodiversity monitoring in my data covered a representative range of accessibility scores, as many parts of our planet are highly accessible (see Figure 8 in appendix B). My data suffered from the opposite, as it underrepresented areas with low accessibility, causing high uncertainty of the model at low levels (Figure 3). Additionally, 33% of the time-series monitored fell into protected areas while having an accessibility level of 0.9 and above (see Figure 11 in appendix B). This indicates that turnover rates at high levels of accessibility might have been influenced by the presence of protected areas which are not strictly representative of other highly accessible areas. In a post-hoc sensitivity analysis, I used a subset of plants-only data because it showed the most variation across levels of accessibility (see Fig-

ure 10 in appendix B). I found similar directional relationships between accessibility and turnover (see Figure 7 in appendix A), as well as relatively high coincidence of time-series in protected areas (39%). Model convergence could not be achieved with a subset of data excluding protected areas, making it difficult to exclude this potential influence on turnover. The second criticism concerns mismatches in timing of disturbance and monitored time-series, as only then an attribution of a global change driver to biodiversity change is possible (Gonzalez et al., 2016; De Palma et al., 2018). One robust method to do this is the Before-After-Control-Impact (BACI) method, which compares changes before and after disturbance. In my case, human activities related to accessibility have often occurred before the start of the biodiversity monitoring (Ibisch et al., 2016), nor did I have a quantification of large-scale human activities over time which prevented me from doing a BACI. Therefore, my results can be interpreted as the time-accumulated impact of human activities which captured both disturbances and recoveries at the same time and therefore also represented the real world (Vellend, Dornelas, et al., 2017). While biodiversity changes and their attribution to large-scale human activities is somewhat limited with my method, I still observed an accurate picture of the cumulative and interactive responses captured across levels of accessibility corresponded with human transformations of the natural environment.

## **4.2 Temporal turnover and accessibility (Research Question 1)**

Contrary to my first prediction, I found that temporal turnover decreased as accessibility increases (Figure 3). I also found that 63% of all time-series monitored showed some degree of turnover (turnover > zero). Furthermore, the lowest model prediction at the highest level of accessibility pointed towards 25% species replacement across the monitored time periods (Figure 3), which is higher than hypothetical baseline turnover rates (Dornelas et al., 2014), indicating widespread turnover rates across levels of accessibility. The negative relationship observed could be due to various reasons. Firstly, heterogeneity exists in species' responses to impacts caused by human activities as it is also reflected in the ongoing debate about habitat fragmentation (Fahrig, 2017; Damschen et al., 2019). Secondly, potential interactions of multiple drivers play an important role. Simultaneous drivers can relate to each other in different ways, being either additive, antagonist or amplified (Côté, Darling and Brown, 2016; Sirami et al., 2017). Research on multiple drivers and their interactions has been limited (Sirami et al., 2017; Mazar et al., 2018) with previous biodiversity change research mostly focussed on single drivers of environmental change such as forest loss (Daskalova et al., 2019), land use change (Newbold et al., 2015) or warming (Spooner, Pearson and Freeman, 2018). A marine study has found antagonistic effects of multiple drivers on biodiversity changes resulting in no net biodiversity change (Dunic et al., 2017). Accessibility is a cumulative variable, representing many different drivers and their respective

environmental impacts. Most research has focussed only on a singular impact of accessibility so far. Land-use change, urbanisation and roads are associated with both higher rates of extinction, due to increased resource and habitat constraints and immigration due to increased alien potential (Hulme, 2009; Seebens et al., 2018). We currently only have a limited understanding how the magnitude of exposure to these drivers differ, which is likely to also influence biodiversity change observed. However, many of the impacts described co-occur and can affect the intensity and interaction of one another (Geary et al., 2019). Furthermore, other drivers such as climate change can bring additional complexity in patterns of exposure, as these drivers can cooccur and interact with human activity-related drivers as well (Chazal and Rounsevell, 2009; Gonzalez et al., 2016). One determining coincidence of my study could be that the very inaccessible places I studied coincide with areas that are in very high and low latitudes, which are often sensitive to climate change (Myers-Smith et al., 2019) and thus might show a higher level of turnover. My findings highlight that accounting for multiple drivers and potential interactions is essential when attributing biodiversity change.

#### **4.3 Temporal turnover and human population density (Research question 2 modified)**

I found a very similar baseline magnitude of temporal turnover and human population density as with accessibility, although in this case without any directional trend. The lowest model prediction at the highest level of human population density pointed towards a turnover level of 22% across the monitored time periods (Figure 4), making it very similar to that of accessibility (24%) (Figure 3), confirming a generally increased turnover rate above natural baseline levels. Interestingly, when looking at the coincidence of the distribution of accessibility and human population density, they were located on opposite ends (see Figure 9 in appendix B). That indicates that high accessibility does not coincide with high human population density, but many places on earth have been modified even without the immediate or dense presence of humans. This is further emphasized by the finding that many highly accessible places coincide with protected areas (see Figure 11 in appendix B). Therefore, accessibility seems to be an important measure that can also capture indirect impacts. Much of our planet has been altered in such a way that patches of habitat are exposed to both direct and indirect influences of human activity. While land-use change from pristine to urban centres and agriculture and forest loss represent more direct impacts, neighbouring habitats fragments are also impacted indirectly and this might be reflected in a metric such as accessibility (Haddad et al., 2015). As pointed out by Gonzalez et al. (2016), understanding and contrasting both direct and indirect effects of human activity on biodiversity change is a challenge that has not been addressed adequately yet. My findings highlight the

importance of capturing both direct and indirect effects of human activities when attributing biodiversity change to global change drivers.

#### **4.4 Temporal turnover and taxa (Research Question 3 modified)**

My findings highlight variation of temporal turnover among taxa, with terrestrial invertebrates experiencing the highest amount, followed by mammals, whereas plants and birds showed relatively lower levels of turnover (Figure 5). Invertebrates' relationship should be treated carefully, as the number of sampling was disproportionally lower for time series-studied (6%) and coincidence with potential confounding protected areas (4%), as well as known difficulties in sampling invertebrates (Andersen et al., 2004). Species' responses ultimately depend on their traits such as life history traits, generation time, trophic levels and reproduction strategy (Purvis et al., 2000). Concerning species' vulnerability to human impacts and consequently setting conservation priorities, research often focussed on traits such as mobility and specialisation. While some research has found more persistence of more mobile species and of species that are more generalist in human impacted areas (Newbold et al., 2014), other research has found both populations increases, and declines related to range size and specialisation (Daskalova, Myers-Smith and Godlee, 2019). However, both studies did not consider community composition in their analysis. I did not find a relationship between the turnover and mobility, assuming birds to be more-mobile and plants to be less-mobile species. Despite there being no relationship with traits in this study, phylogenetic relatedness might predict extinction over longer timescales (Jetz and Pyron, 2018). Shorter timescales might be influenced by other processes such as the interaction of effects and ecological lags. I demonstrated heterogeneous levels of turnover experienced among taxa, pointing out the need to quantify how these relationships are moderated by species' traits and vulnerabilities in order to effectively target conservation efforts.

#### **4.5 Study Limitations**

Analyses of biodiversity change can be limited by insufficient and imbalanced taxonomic, spatial and temporal data. The lack of model convergence which prevented answering two of my original research questions, highlighted the need to collect data more comprehensively. Tropical locations and thus species were underrepresented in the biodiversity dataset (see figure 2), as were invertebrates in my subset of data. Spatial scales that are of species-specific relevance can differ heavily among species but also in relation to human impact. In a sensitivity analysis I found that turnover trends only differed in their magnitude but not in the direction of the relationship with accessibility and only

between very small scales (1km<sup>2</sup>) but not across bigger scales (see Figure 6 in appendix A). Lastly, temporal mismatches and missing accurate accounts of past human activities make the attribution of biodiversity change challenging. Recognising heterogeneous effects including disturbance and recovery of human impacts and focussing on understanding the interaction of current impacts are thus gaining even more importance. Analysis of biodiversity change and its global change driver can benefit from more comprehensive taxonomic, spatial and temporal data collection of environmental impact and biodiversity data.

#### **4.6 Future directions**

Biodiversity change and attributing global change drivers associated with human activities requires a step beyond previous studies which have focussed on singular drivers. Exploring patterns of exposure of multiple influences (Bowler et al., 2020) will give us a more differentiated view than a singular cumulative metric can give. Analysing interactive and cumulative effects of those various drivers, will allow us to better understand the complex varieties with which our activities alter the environment and consequently ecological communities. Understanding the role better that species' traits play in determining vulnerability to multiple human impacts will lead to more targeted conservation efforts. Also going beyond one singular community, we need to understand better how large-scale reorganization is unfolding across communities as well and how biotic homogenization (Eriksson and Hillebrand, 2019) is influencing ecosystem functioning across scales. Analysing biodiversity change related to multiple drivers across both terrestrial and marine realms will allow development of better conservation strategies.



## 5 Conclusions

My analysis revealed complex relationships of biodiversity change and large-scale human activities across taxa with a decrease of 1.6% of temporal turnover for every 10% increase in accessibility. These findings challenge the assumption that the absence of large-scale human activity is sufficient to conserve compositional biodiversity (Di Marco et al., 2019). I support the latest research, calling for the importance of using real-world data to detect longer term dynamics that are revealing more complex biodiversity changes (Daskalova et al., 2019). I further highlighted the additional complexity in biodiversity changes arising through interactions between different drivers and environmental changes. To fully comprehend biodiversity changes it is not sufficient to consider singular environmental changes in isolation. More differentiated research that distinguishes the relative contributions of each driver is needed, as a metric like accessibility, albeit useful in one sense, is too simplistic. My findings highlight the complexity of biodiversity change as well as the heterogeneity in responses. Consequently, I am challenging calls for establishing wilderness areas (Di Marco et al., 2019) as a conservation priority, unless we gain a better understanding of biodiversity changes both inside and outside of human-impacted areas. Being able to adequately quantify wide ranging responses of biodiversity change to large-scale human activities will improve predictions for better international policy making in the light of our rapidly changing Anthropocene.

## 6 References

- Andersen, A. N. et al. (2004) 'Use of terrestrial invertebrates for biodiversity monitoring in Australian rangelands, with particular reference to ants', *Austral Ecology*, 29(1), pp. 87–92. doi: 10.1111/j.1442-9993.2004.01362.x.
- Barnes, Richard (2017). dggridR: Discrete Global Grids for R. R package version 2.0.4. "<https://github.com/r-barnes/dggridR/>" doi:10.5281/zenodo.1322866
- Barnosky, A. D. et al. (2011) 'Has the Earth's sixth mass extinction already arrived?', *Nature*. Nature Publishing Group, 471(7336), pp. 51–57. doi: 10.1038/nature09678.
- Baselga A. et al.(2018). betapart: Partitioning Beta Diversity into Turnover and Nestedness Components. R package version 1.5.1. <https://CRAN.R-project.org/package=betapart>
- Blowes, S. A. et al. (2019) 'The geography of biodiversity change in marine and terrestrial assemblages', *Science*, 366(6463), pp. 339–345. doi: 10.1126/science.aaw1620.
- Bowler, D. E. et al. (2020) 'Mapping human pressures on biodiversity across the planet uncovers anthropogenic threat complexes', *People and Nature*, n/a(n/a). doi: 10.1002/pan3.10071.
- Bürkner, P. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, 10(1), 395–411. doi:10.32614/RJ-2018-017
- Butchart, S. H. M. et al. (2010) 'Global Biodiversity: Indicators of Recent Declines', *Science*, 328(5982), pp. 1164–1168. doi: 10.1126/science.1187512.
- Chase, J. M. et al. (2019) 'Species richness change across spatial scales', *Oikos*, 128(8), pp. 1079–1091. doi: 10.1111/oik.05968.
- Chazal, J. and Rounsevell, M. (2009) 'Land-Use and Climate Change within Assessments of Biodiversity Change: A Review', *Global Environmental Change*, 19, pp. 306–315. doi: 10.1016/j.gloenvcha.2008.09.007.
- CIESIN (2018) 'Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11'. NASA Socioeconomic Data and Applications Center (SEDAC). Available at: <https://doi.org/10.7927/H49C6VHW>.
- Concepción, E. D. et al. (2015) 'Impacts of urbanisation on biodiversity: the role of species mobility, degree of specialisation and spatial scale', *Oikos*, 124(12), pp. 1571–1582. doi: 10.1111/oik.02166.
- Côté, I. M., Darling, E. S. and Brown, C. J. (2016) 'Interactions among ecosystem stressors and their

importance in conservation', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 283(1824), p. 20152592. doi: 10.1098/rspb.2015.2592.

Damschen, E. I. et al. (2019) 'Ongoing accumulation of plant diversity through habitat connectivity in an 18-year experiment', *Science*. American Association for the Advancement of Science, 365(6460), pp. 1478–1480. doi: 10.1126/science.aax8992.

Daskalova, G. N. et al. (2019) 'Landscape-scale forest loss as a catalyst of population and biodiversity change', *bioRxiv*, p. 473645. doi: 10.1101/473645.

Daskalova, G. N., Myers-Smith, I. H. and Godlee, J. L. (2019) 'All is not decline across global vertebrate populations', *bioRxiv*, p. 272898. doi: 10.1101/272898.

De Palma, A. et al. (2018) 'Challenges with inferring how land-use affects terrestrial biodiversity: Study design, time, space and synthesis', *Advances in Ecological Research*. doi: De Palma, Adriana, Sanchez Ortiz, Katia, Martin, Phillip A., Chadwick, Amy, Gilbert, Guillermo, Bates, Amanda E., Börger, Luca, Contu, Sara, Hill, Samantha L.L. and Purvis, Andy (2018) Challenges with inferring how land-use affects terrestrial biodiversity: Study design, time, space and synthesis. *Advances in Ecological Research*. (doi:10.1016/bs.aecr.2017.12.004 <http://dx.doi.org/10.1016/bs.aecr.2017.12.004>).

Di Marco, M. et al. (2019) 'Wilderness areas halve the extinction risk of terrestrial biodiversity', *Nature*. Nature Publishing Group, 573(7775), pp. 582–585. doi: 10.1038/s41586-019-1567-7.

Dornelas, M. et al. (2014) 'Assemblage Time Series Reveal Biodiversity Change but Not Systematic Loss', *Science*, 344(6181), pp. 296–299. doi: 10.1126/science.1248484.

Dornelas, M. et al. (2018) 'BioTIME: A database of biodiversity time series for the Anthropocene', *Global Ecology and Biogeography*, 27(7), pp. 760–786. doi: 10.1111/geb.12729.

Dunic, J. C. et al. (2017) 'Attributing the variability in direction and magnitude of local-scale marine biodiversity change to human activities', *bioRxiv*. Cold Spring Harbor Laboratory, p. 162362. doi: 10.1101/162362.

Ehrlén, J. and Morris, W. F. (2015) 'Predicting changes in the distribution and abundance of species under environmental change', *Ecology Letters*, 18(3), pp. 303–314. doi: 10.1111/ele.12410.

Eriksson, B. K. and Hillebrand, H. (2019) 'Rapid reorganization of global biodiversity', *Science*, 366(6463), pp. 308–309. doi: 10.1126/science.aaz4520. Fahrig, L. (2017) 'Ecological Responses to Habitat Fragmentation Per Se', *Annual Review of Ecology, Evolution, and Systematics*, 48(1), pp. 1–23. doi: 10.1146/annurev-ecolsys-110316-022612.

- Foden, W. B. et al. (2019) 'Climate change vulnerability assessment of species', *WIREs Climate Change*, 10(1), p. e551. doi: 10.1002/wcc.551.
- Gaston, K. J. and Fuller, R. A. (2008) 'Commonness, population depletion and conservation biology', *Trends in Ecology & Evolution*, 23(1), pp. 14–19. doi: 10.1016/j.tree.2007.11.001.
- Geary, W. L. et al. (2019) 'Threat webs: Reframing the co-occurrence and interactions of threats to biodiversity', *Journal of Applied Ecology*, 56(8), pp. 1992–1997. doi: 10.1111/1365-2664.13427.
- Gonzalez, A. et al. (2016) 'Estimating local biodiversity change: a critique of papers claiming no net loss of local diversity', *Ecology*, 97(8), pp. 1949–1960. doi: 10.1890/15-1759.1.
- Gotelli, N. J. et al. (2017) 'Community-level regulation of temporal trends in biodiversity', *Science Advances*, 3(7), p. e1700315. doi: 10.1126/sciadv.1700315.
- Haddad, N. M. et al. (2015) 'Habitat fragmentation and its lasting impact on Earth's ecosystems', *Science Advances*. American Association for the Advancement of Science, 1(2), p. e1500052. doi: 10.1126/sciadv.1500052.
- Hijmans, Robert J. (2020). raster: Geographic Data Analysis and Modeling. R package version 3.1-5. <https://CRAN.R-project.org/package=raster>
- Hillebrand, H. et al. (2018) 'Biodiversity change is uncoupled from species richness trends: Consequences for conservation and monitoring', *Journal of Applied Ecology*, 55(1), pp. 169–184. doi: 10.1111/1365-2664.12959.
- Hulme, P. E. (2009) 'Trade, transport and trouble: managing invasive species pathways in an era of globalization', *Journal of Applied Ecology*, 46(1), pp. 10–18. doi: 10.1111/j.1365-2664.2008.01600.x.
- Ibisch, P. L. et al. (2016) 'A global map of roadless areas and their conservation status', *Science*. American Association for the Advancement of Science, 354(6318), pp. 1423–1427. doi: 10.1126/science.aaf7166.
- IPBES (2019) 'Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services'. Available at: <https://uwe-repository.worktribe.com/output/1493508/summary-for-policymakers-of-the-global-assessment-report-on-biodiversity-and-ecosystem-services-of-the-intergovernmental-science-policy-platform-on-biodiversity-and-ecosystem-services> (Accessed: 24 January 2020).
- Isaac, N. J. B. and Cowlshaw, G. (2004) 'How species respond to multiple extinction threats', *Proceedings of the Royal Society of London. Series B: Biological Sciences*. Royal Society, 271(1544),

pp. 1135–1141. doi: 10.1098/rspb.2004.2724.

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 & Evolution*. Nature Publishing Group, 2(5), pp. 850–858. doi: 10.1038/s41559-018-0515-5.

Kareiva, P. et al. (1990) 'Population dynamics in spatially complex environments: theory and data', *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. Royal Society, 330(1257), pp. 175–190. doi: 10.1098/rstb.1990.0191.

Kilpatrick, A. M. and Ives, A. R. (2003) 'Species interactions can explain Taylor's power law for ecological time series', *Nature*. Nature Publishing Group, 422(6927), pp. 65–68. doi: 10.1038/nature01471.

Lande, R. (1993) 'Risks of Population Extinction from Demographic and Environmental Stochasticity and Random Catastrophes', *The American Naturalist*. The University of Chicago Press, 142(6), pp. 911–927. doi: 10.1086/285580.

Longton, R. and Hedderson, T. (2000) 'What Are Rare Species and Why Conserve Them?', *Lindbergia*, 25, pp. 53–61. doi: 10.2307/20150038.

Lüdtke D (2018). "ggeffects: Tidy Data Frames of Marginal Effects from Regression Models."(26), 772. doi:10.21105/joss.00772

Magurran, A. et al. (2018) 'Divergent biodiversity change within ecosystems', *Proceedings of the National Academy of Sciences*, 115, p. 201712594. doi: 10.1073/pnas.1712594115.

malariatlas.org (2020) Accessibility to Cities, MAP. Available at: [https://malariatlas.org/research-project/accessibility\\_to\\_cities/](https://malariatlas.org/research-project/accessibility_to_cities/) (Accessed: 5 May 2020).

Mazor, T. et al. (2018) 'Global mismatch of policy and research on drivers of biodiversity loss', *Nature Ecology & Evolution*. Nature Publishing Group, 2(7), pp. 1071–1074. doi: 10.1038/s41559-018-0563-x.

McGill, B. J. et al. (2015) 'Fifteen forms of biodiversity trend in the Anthropocene', *Trends in Ecology & Evolution*, 30(2), pp. 104–113. doi: 10.1016/j.tree.2014.11.006.

Myers-Smith, I. H. et al. (2019) 'Eighteen years of ecological monitoring reveals multiple lines of evidence for tundra vegetation change', *Ecological Monographs*, 89(2), p. e01351. doi: 10.1002/ecm.1351.

Newbold, T. et al. (2014) 'A global model of the response of tropical and sub-tropical forest biodiversity to anthropogenic pressures', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society,

281(1792), p. 20141371. doi: 10.1098/rspb.2014.1371.

Newbold, T. et al. (2015) 'Global effects of land use on local terrestrial biodiversity', *Nature*, 520(7545), pp. 45–50. doi: 10.1038/nature14324.

Oksanen, F. et al. (2019) *vegan: Community Ecology Package*. R package version 2.5-6. <https://CRAN.R-project.org/package=vegan>

Pereira, H. M., Navarro, L. M. and Martins, I. S. (2012) 'Global Biodiversity Change: The Bad, the Good, and the Unknown', *Annual Review of Environment and Resources*, 37(1), pp. 25–50. doi: 10.1146/annurev-environ-042911-093511.

Purvis, A. et al. (2000) 'Predicting extinction risk in declining species', *Proceedings of the Royal Society of London. Series B: Biological Sciences*. Royal Society, 267(1456), pp. 1947–1952. doi: 10.1098/rspb.2000.1234.

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>

Seebens, H. et al. (2018) 'Global rise in emerging alien species results from increased accessibility of new source pools', *Proceedings of the National Academy of Sciences of the United States of America*, 115(10), pp. E2264–E2273. doi: 10.1073/pnas.1719429115.

Sirami, C. et al. (2017) 'Impacts of global change on species distributions: obstacles and solutions to integrate climate and land use', *Global Ecology and Biogeography*, 26(4), pp. 385–394. doi: 10.1111/geb.12555.

Spooner, F. E. B., Pearson, R. G. and Freeman, R. (2018) 'Rapid warming is associated with population decline among terrestrial birds and mammals globally', *Global Change Biology*. John Wiley & Sons, Ltd, 24(10), pp. 4521–4531. doi: 10.1111/gcb.14361.

UN (2015) *Transforming our world: the 2030 Agenda for Sustainable Development*, DISD. Available at: <https://www.un.org/development/desa/dspd/2015/08/transforming-our-world-the-2030-agenda-for-sustainable-development/> (Accessed: 26 January 2020).

Vellend, M., Dornelas, M., et al. (2017) 'Estimates of local biodiversity change over time stand up to scrutiny', *Ecology*, 98(2), pp. 583–590. doi: 10.1002/ecy.1660.

Vellend, M., Baeten, L., et al. (2017) 'Plant Biodiversity Change Across Scales During the Anthropocene', *Annual Review of Plant Biology*, 68(1), pp. 563–586. doi: 10.1146/annurev-arplant-042916-040949.

Weiss, D. J. et al. (2018) 'A global map of travel time to cities to assess inequalities in accessibility in 2015', *Nature*, 553(7688), pp. 333–336. doi: 10.1038/nature25181.

## 7 Appendices

### 7.1 Appendix A: Sensitivity analyses

I conducted two sensitivity analyses to explore my data further. Firstly, I conducted an analysis to test the sensitivity of scale, modelling accessibility extractions at 1km<sup>2</sup>, 50km<sup>2</sup> and 100km<sup>2</sup>. Secondly, in a post-hoc analysis I used a subset of my data that only included plants, as this was the taxon that was most equally distributed across levels of accessibility. All model terms were kept equal to the original model, outlined above, except when I modelled the plants-subset, I removed the taxa fixed effect.

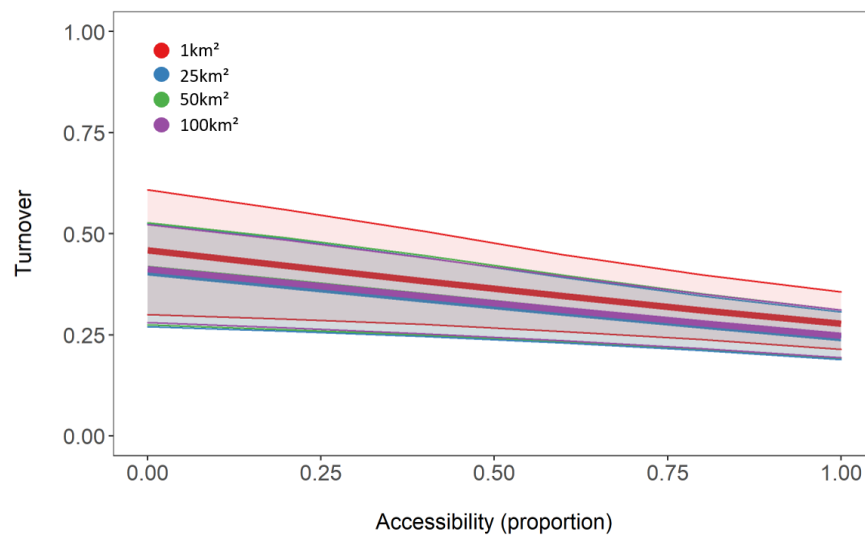


Figure 6: All different scales of extracting accessibility (1km<sup>2</sup>, 25km<sup>2</sup>, 50km<sup>2</sup>, 100km<sup>2</sup>) displayed a similar negative relationship with temporal turnover using Bayesian mixed models. Accessibility scores extracted at 1km<sup>2</sup> displayed the highest intercept and the biggest error at lower levels of accessibility (slope = -1.13, CI = -2.02 - -0.29, n = 4738, see Table 3 for full model outputs). Accessibility scores extracted at 50km<sup>2</sup> and 100km<sup>2</sup> were very similar to the one at 25km<sup>2</sup> which I used for the rest of my analyses (slope = -1.22, CI = -2.1 - -0.33, n = 5787, see Table 4 for full model outputs; slope = -1.21, CI = -2.06 - -0.37, n = 5787, see Table 5 for full model outputs, respectively). Lines and error bands represent model predictions and 95 percent credible intervals, respectively. Colour coding of lines represent differing scales of extraction of accessibility values which were calculated taking the mean. See model Rhat values in appendix confirming model convergence.



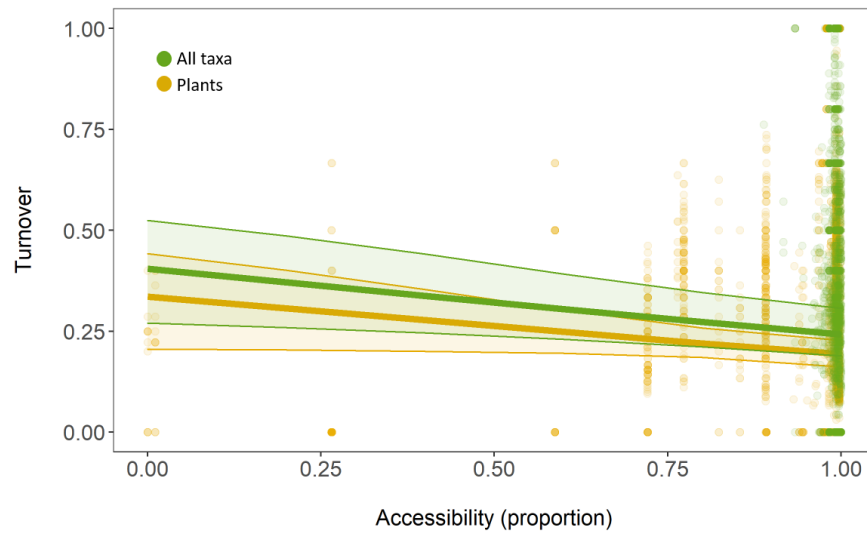


Figure 7: Plants displayed a similar negative relationship with temporal turnover as all taxa combined did using Bayesian mixed models (slope = -1.11, CI = -2.1 - -0.07,  $n = 3487$ , see Table 2 for full model outputs). The plant intercept was slightly lower than the one including all taxa. Lines and error bands represent model predictions and 95 percent credible intervals, respectively. Colour coding of lines represent all data (green) and plant subset of data used for model (yellow). Points are raw data. See model Rhat values in appendix confirming model convergence.

## 7.2 Appendix B: Additional figures

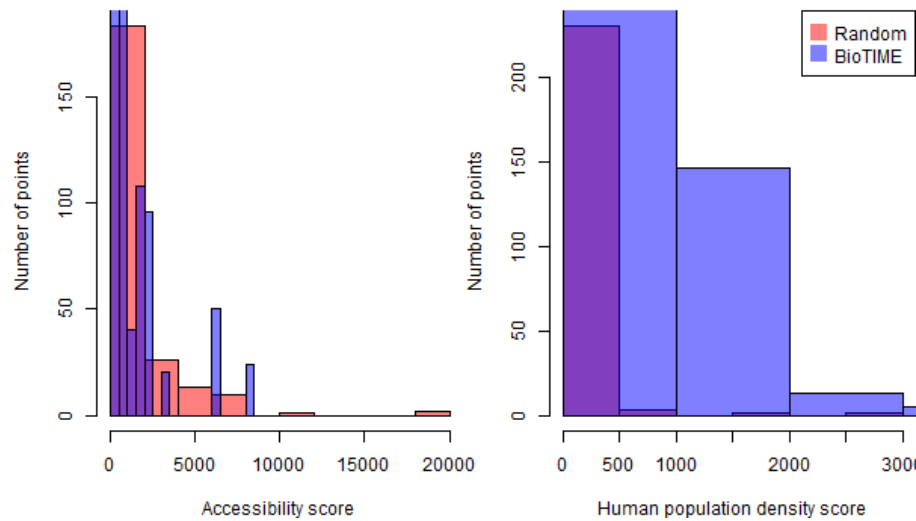


Figure 8: My extracted values of accessibility and human population density coincided with levels of randomly extracted values. Random values were generated with a latitude/longitude number generator and subsequently treated equally as I did with the BioTIME values (but note that I did not yet inverse the scores of accessibility).

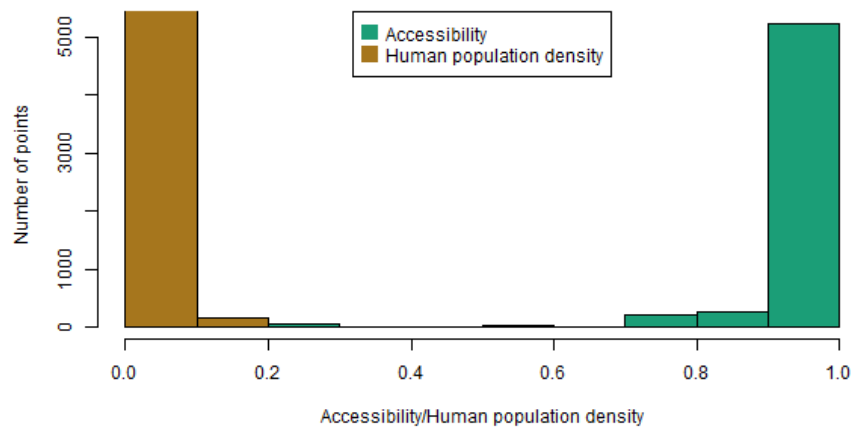


Figure 9: Using the locations of BioTIME, the extracted accessibility and human population density proportions fell on opposite ends of the spectrum (note the inverse of accessibility scores as also used in the figures in the main text).

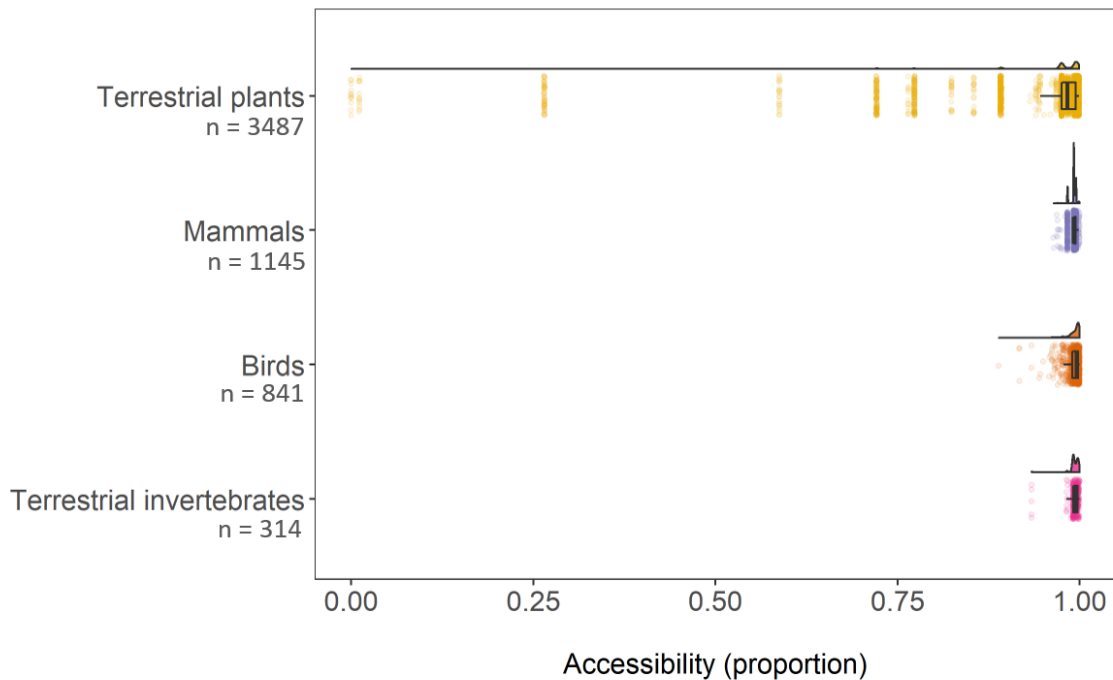


Figure 10: Accessibility scores are not distributed equally across taxa. Plants showed the greatest representation across levels of accessibility. All are raw data visualisation of accessibility across taxa. Split violins represent data distribution, boxplots with mean values and points are raw data. Numbers are sample size per taxon.

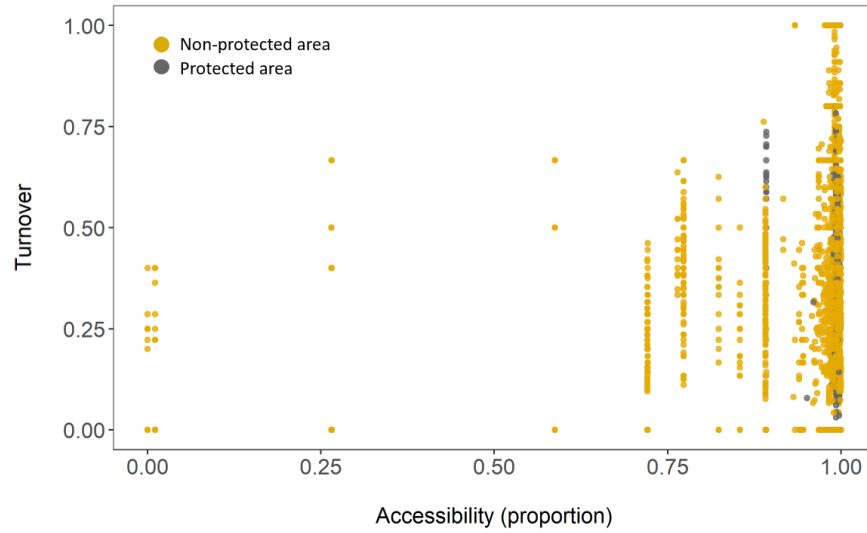


Figure 11: All of the time-series coinciding with protected areas (1888 out of 5787) had an accessibility score of 0.9 and above. Points are raw data.

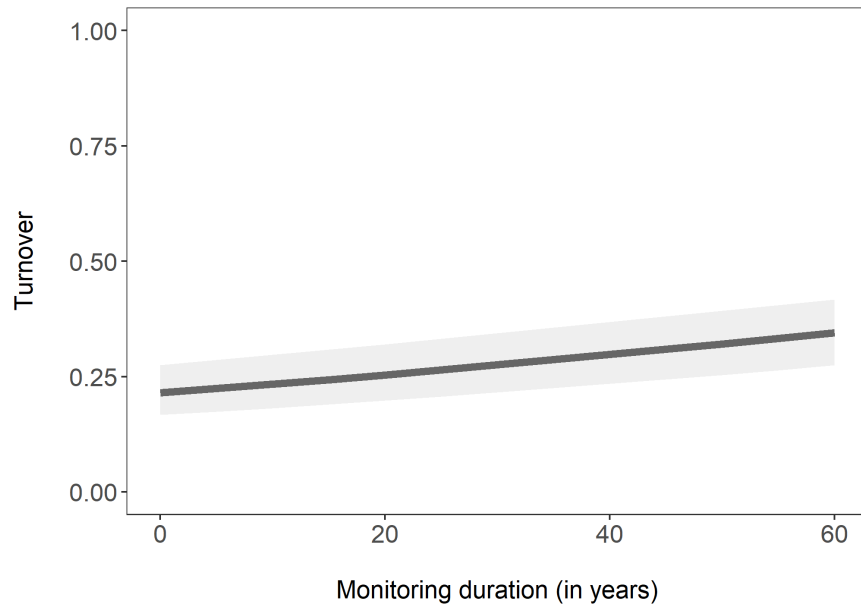


Figure 12: Temporal turnover increased as monitoring duration increases across the 5787 time-series surveyed using a Bayesian mixed model (slope = 0.02, CI = 0.01 to 0.02, see Table 1 for full model outputs). Lines and error bands represent model predictions and 95 percent credible intervals, respectively. See model Rhat values in appendix confirming model convergence.

### 7.3 Appendix C: Model outputs

Random effects were not included in the model outputs for ease of overview. Full model outputs can be found in my GitHub repository (<https://github.com/DaniGargya/dissertation.git>).

Table 1: Outputs of the model testing the relationship between temporal turnover and accessibility, human population density, duration and taxa (using random intercepts for Study ID and grid cell). Accessibility and human population density scores were extracted at 25km<sup>2</sup>.

Cleaned_Parameter	Parameter	Group	Estimate	Std.Error	HDI(95%)	pd	ESS	Rhat	MCSE
(Intercept)	b_Intercept		0.1052413	0.5122943	[-0.87, 1.14]	0.5824167	7381	0.9998629	0.0059429
scaleacc_25	b_scaleacc_25		-1.1990326	0.4564273	[-2.11, -0.35]	0.9964167	10563	0.9998509	0.0044085
scalehpd_25	b_scalehpd_25		-0.2199697	0.1367625	[-0.50, 0.04]	0.9445833	18318	0.9997444	0.0010175
duration_plot	b_duration_plot		0.0168001	0.0020429	[0.01, 0.02]	1.0000000	16841	1.0000213	0.0000157
TAXAMammals	b_TAXAMammals		0.7112776	0.3532237	[-0.02, 1.39]	0.9764167	4775	1.0013285	0.0051446
TAXATerrestrialinvertebrates	b_TAXATerrestrialinvertebrates		0.9874206	0.3491712	[0.31, 1.70]	0.9977500	4837	1.0006850	0.0050756
TAXATerrestrialplants	b_TAXATerrestrialplants		0.0002394	0.2761790	[-0.56, 0.55]	0.5002500	4307	1.0007951	0.0042641

Table 2: Outputs of the model using a subset of plant-only data testing the relationship between temporal turnover and accessibility, human population density and duration (using random intercepts for Study ID and grid cell). Accessibility and human population density scores were extracted at 25km<sup>2</sup>.

Cleaned_Parameter	Parameter	Group	Estimate	Std.Error	HDI(95%)	pd	ESS	Rhat	MCSE
(Intercept)	b_Intercept		-0.0403534	0.4881842	[-1.03, 0.92]	0.5384167	5679	1.0000343	0.0065298
scaleacc_25	b_scaleacc_25		-1.1102855	0.5049306	[-2.09, -0.05]	0.9825833	5453	1.0001567	0.0070112
scalehpd_25	b_scalehpd_25		0.0177177	3.7895687	[-6.91, 7.93]	0.5021667	8242	1.0003289	0.0417429
duration_plot	b_duration_plot		0.0194180	0.0021272	[0.02, 0.02]	1.0000000	22421	0.9997537	0.0000142

Table 3: Outputs of the model testing sensitivity of scale in the relationship between temporal turnover and accessibility, human population density, duration and taxa (using random intercepts for Study ID and grid cell). Accessibility and human population density scores were extracted at 1km<sup>2</sup>.

Cleaned_Parameter	Parameter	Group	Estimate	Std.Error	HDI(95%)	pd	ESS	Rhat	MCSE
(Intercept)	b_Intercept		-0.0068572	0.5082247	[-0.99, 0.97]	0.5052500	7963	1.0003496	0.0056571
scaleacc_1	b_scaleacc_1		-1.1330821	0.4386837	[-1.98, -0.26]	0.9956667	10911	1.0001544	0.0042309
scalehpd_1	b_scalehpd_1		-0.1937612	0.1298192	[-0.43, 0.07]	0.9319167	20191	0.9998427	0.0009080
duration_plot	b_duration_plot		0.0168902	0.0020260	[0.01, 0.02]	1.0000000	17825	0.9997086	0.0000152
TAXAMammals	b_TAXAMammals		0.7609157	0.3593235	[0.03, 1.45]	0.9805000	5523	0.9998649	0.0048876
TAXATerrestrialinvertebrates	b_TAXATerrestrialinvertebrates		0.9821865	0.3604289	[0.26, 1.67]	0.9967500	5599	0.9999022	0.0048462
TAXATerrestrialplants	b_TAXATerrestrialplants		0.0417673	0.2829950	[-0.51, 0.63]	0.5584167	5017	1.0003132	0.0040734

Table 4: Outputs of the model testing sensitivity of scale in the relationship between temporal turnover and accessibility, human population density, duration and taxa (using random intercepts for Study ID and grid cell). Accessibility and human population density scores were extracted at 50km<sup>2</sup>.

Cleaned_Parameter	Parameter	Group	Estimate	Std.Error	HDI(95%)	pd	ESS	Rhat	MCSE
(Intercept)	b_Intercept		0.1356536	0.5051079	[-0.87, 1.14]	0.6070833	6834	1.0002099	0.0062096
scaleacc_50	b_scaleacc_50		-1.2192630	0.4466915	[-2.09, -0.32]	0.9965000	9651	0.9998305	0.0046045
scalehpd_50	b_scalehpd_50		-0.1817154	0.1198487	[-0.41, 0.06]	0.9338333	18096	0.9997597	0.0008968
duration_plot	b_duration_plot		0.0168613	0.0020752	[0.01, 0.02]	1.0000000	17172	0.9997405	0.0000157
TAXAMammals	b_TAXAMammals		0.7053864	0.3429454	[0.04, 1.39]	0.9795833	4701	1.0009269	0.0050495
TAXATerrestrialinvertebrates	b_TAXATerrestrialinvertebrates		0.9737655	0.3492443	[0.28, 1.65]	0.9963333	4707	1.0007377	0.0051086
TAXATerrestrialplants	b_TAXATerrestrialplants		-0.0148072	0.2780921	[-0.58, 0.51]	0.5205833	3964	1.0010940	0.0044433

Table 5: Outputs of the model testing sensitivity of scale in the relationship between temporal turnover and accessibility, human population density, duration and taxa (using random intercepts for Study ID and grid cell). Accessibility and human population density scores were extracted at 100km<sup>2</sup>.

Cleaned_Parameter	Parameter	Group	Estimate	Std.Error	HDI(95%)	pd	ESS	Rhat	MCSE
(Intercept)	b_Intercept		0.1301411	0.4888914	[-0.85, 1.06]	0.6035000	7531	1.0000554	0.0056413
scaleacc_100	b_scaleacc_100		-1.2075882	0.4326068	[-2.05, -0.37]	0.9973333	10343	0.9998170	0.0042539
scalehpd_100	b_scalehpd_100		-0.2778054	0.1294950	[-0.53, -0.03]	0.9859167	16268	0.9999648	0.0010041
duration_plot	b_duration_plot		0.0168349	0.0020136	[0.01, 0.02]	1.0000000	16713	0.9997272	0.0000155
TAXAMammals	b_TAXAMammals		0.7160233	0.3471979	[0.04, 1.43]	0.9779167	5159	0.9998328	0.0048779
TAXATerrestrialinvertebrates	b_TAXATerrestrialinvertebrates		0.9979777	0.3474504	[0.31, 1.65]	0.9968333	5004	1.0001845	0.0049072
TAXATerrestrialplants	b_TAXATerrestrialplants		-0.0172238	0.2719903	[-0.56, 0.53]	0.5293333	4636	1.0004054	0.0040314

## 7.4 Appendix D: R Code

R Code, including outputs, can be accessed through GitHub <https://github.com/DaniGargya/dissertation.git>.

My preregistration can be accessed through this link at the Open Science Framework [https://osf.io/pua5m/?view\\_only=575f6a48587245f3b6971235bcf32b3f](https://osf.io/pua5m/?view_only=575f6a48587245f3b6971235bcf32b3f).

### Initial set-up

```
# Full R script analysing biodiversity changes across levels of accessibility  
# Dissertation project 2020, University of Edinburgh  
# Daniela Gargya  
# May 2020  
  
# Load libraries ----  
library(tidyverse)  
library(ggplot2)  
library(vegan)  
library(betapart)  
library(rgdal)  
library(raster)  
library(labdsv)  
library(dggridR)  
library(brms)  
library(ggeffects)  
library(tidybayes)  
library(bayesplot)  
library(maps)  
library(ggthemes)  
library(treemap)  
library(RColorBrewer)  
library(treemapify)  
library(ggpubr)
```

```

library(mapdata)
library(gridExtra)

# Load data ----
# BioTIME
biotime_full <- read.csv("data/BioTIMEQuery02_04_2018.csv")
biotime_meta <- read.csv("data/BioTIMEMetadata_02_04_2018.csv")
# accessed through website: http://biotime.st-andrews.ac.uk/downloadArea.php

# accessibility data
aa <- raster("data/2015_accessibility_to_cities_v1.0/2015_accessibility_to_cities_v1.0.tif")
# accessed through website: https://malariaatlas.org/explorer/#/

# human population density data
hpd <- raster("data/gpw_v4_population_density_rev11_2015_30_sec.tif")
# accessed through website:
# https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-rev11

### filter biotime data ----
### data inclusion criteria
# terrestrial realm
# minimum study duration 5 years
# at least 15 studies per taxa
# no more than 5000 plots per study
# at least 2 survey points in each plot
# filter by fixed plot

bio <- biotime_meta %>%
  filter(REALM == "Terrestrial") %>% # only terrestrial species
  group_by(STUDY_ID) %>%
  mutate(duration = END_YEAR - START_YEAR) %>%
  filter(!duration < 5) %>% # minimum duration of 5 years
  #filter(!DATA_POINTS < 5) %>%

```

```

group_by(TAXA) %>%
mutate(studies_taxa=length(unique(STUDY_ID))) %>%
filter(!studies_taxa < 15) %>% # minimum 15 studies per taxa
ungroup() %>%
# merge datasets
left_join(biotime_full, by = "STUDY_ID") %>%
group_by(STUDY_ID) %>%
unite(STUDY_ID_PLOT, STUDY_ID, PLOT, sep = "_", remove=F) %>%
group_by(STUDY_ID_PLOT) %>%
mutate(maxplots = length(unique(STUDY_ID_PLOT))) %>%
filter(!maxplots > 5000) %>%
ungroup() %>%
group_by(STUDY_ID_PLOT) %>%
filter(HAS_PLOT == "Y") %>%
filter(YEAR %in% c(max(YEAR), min(YEAR))) %>%
mutate(number_plots = length(unique(YEAR))) %>%
filter(number_plots == 2) %>% # have min and max year per plot only
dplyr::select(STUDY_ID, PLOT, STUDY_ID_PLOT, START_YEAR, END_YEAR, duration, TAXA,
               LATITUDE, LONGITUDE, YEAR, sum.allrawdata.ABUNDANCE, GENUS_SPECIES,
               AREA_SQ_KM, NUMBER_OF_SAMPLES, ABUNDANCE_TYPE, AB_BIO, BIOMASS_TYPE,
               PROTECTED_AREA) %>%
ungroup()

```

## Data processing

```

### calculating temporal turnover as component of beta diversity ----

# data manipulation ----
# change biomass and density data to 1 -> absence/presence
bio$sum.allrawdata.ABUNDANCE[bio$sum.allrawdata.ABUNDANCE < 1] <- 1

# summarise abundances per year per species per plot
bio_turnover <- bio %>%

```

```

dplyr::select(STUDY_ID_PLOT, YEAR, GENUS_SPECIES, sum.allrawdata.ABUNDANCE) %>%
group_by(STUDY_ID_PLOT, YEAR, GENUS_SPECIES) %>%
summarise(Abundance = sum(sum.allrawdata.ABUNDANCE)) %>%
ungroup()

# creating empty dataframe ----
beta_Jaccard <- data.frame(matrix(ncol = 6, nrow = length(unique(bio_turnover$STUDY_ID_PLOT))))
names(beta_Jaccard) <- c("STUDY_ID_PLOT", "duration_plot",
                        "richness_change", "Jbeta", "Jtu", "Jne")

i = 1

# for loop with betapart (code help received from Isla) ----
for (i in 1:length(unique(bio_turnover$STUDY_ID_PLOT))) {
  StudyIDPlot <- as.character(unique(bio_turnover$STUDY_ID_PLOT)[i])
  sub_bio_abundance <- filter(bio_turnover,
                              STUDY_ID_PLOT == StudyIDPlot)
  duration_plot <- (max(sub_bio_abundance$YEAR) - min(sub_bio_abundance$YEAR)) + 1
  sub_bio_abundance_wider <- pivot_wider(sub_bio_abundance, names_from = GENUS_SPECIES,
                                         values_from = Abundance,
                                         values_fill = list(Abundance = 0))
  sub_bio_abundance_matrix <- dplyr::select(sub_bio_abundance_wider, -STUDY_ID_PLOT, -YEAR)
  sub_bio_presence_matrix <- with(sub_bio_abundance_matrix,
                                  ifelse(sub_bio_abundance_matrix > 0, 1, 0))
  J_components <- beta.pair(sub_bio_presence_matrix, index.family='jaccard') # distance
  richness_change <- rowSums(sub_bio_presence_matrix)[2] - rowSums(sub_bio_presence_matrix)[1]
  Jbeta <- J_components$beta.jac
  Jtu <- J_components$beta.jtu
  Jne <- J_components$beta.jne
  beta_Jaccard[i,] <- c(StudyIDPlot, duration_plot, richness_change, Jbeta, Jtu, Jne)

  i = i+1
}

```



```

### extracting values accessibility ----
# df of all unique lat/long values
bio_short <- bio %>%
  distinct(STUDY_ID_PLOT, .keep_all = TRUE)

SP <- bio_short %>%
  dplyr::select(LATITUDE, LONGITUDE, STUDY_ID_PLOT) %>%
  distinct(LATITUDE, .keep_all = TRUE)

# adjusting to crs
points <- cbind(SP$LONGITUDE, SP$LATITUDE)
sppoints <- SpatialPoints(points, proj4string=CRS('+proj=longlat +datum=WGS84'))
tp <- spTransform(sppoints, crs(aa))

# extract numbers at different scales
e <- extract(aa, tp)
e_25 <- extract(aa, tp, buffer = 25000, fun = mean)
e_50 <- extract(aa, tp, buffer = 50000, fun = mean)
e_100 <- extract(aa, tp, buffer = 100000, fun = mean)

bio_aa_2 <- cbind(SP, e, e_2, e_5, e_25, e_50, e_75, e_100)

# drop NA according to scale I am looking at
# scale extracted values between zero and one
bio_aa_short <- bio_aa_2 %>%
  drop_na(e_25) %>%
  mutate(scaleacc_25= 1 - ((e_25 -min(e_25))/(max(e_25)-min(e_25))))

# join to full dataset
bio_full_acc <- bio_aa_short %>%
  right_join(bio_short, by = "LATITUDE") %>%
  dplyr::select(-STUDY_ID_PLOT.x, -LONGITUDE.y) %>%
  rename(STUDY_ID_PLOT = STUDY_ID_PLOT.y,

```

```

    LONGITUDE = LONGITUDE.x) %>%
dplyr::select(STUDY_ID_PLOT, scaleacc_25, e, e_25, e_50, e_100)

### extracting values hpd ----

# turn lat/long values into right CRS format
tp_hpd <- spTransform(sppoints, crs(hpd))

# extract long/lat from raster at different scales
e_hpd <- extract(hpd, tp_hpd)
e_hpd25 <- extract(hpd, tp_hpd, buffer = 25000, fun = mean)
e_hpd50 <- extract(hpd, tp_hpd, buffer = 50000, fun = mean)
e_hpd100 <- extract(hpd, tp_hpd, buffer = 100000, fun = mean)

# bind extracted values to dataframe
bio_hpd2 <- cbind(SP, e_hpd, e_hpd2, e_hpd5, e_hpd25, e_hpd50, e_hpd75, e_hpd100)

# drop NA according to scale I am looking at
# scale extraccted values between zero and one
bio_hpd_short <- bio_hpd2 %>%
  drop_na(e_hpd25) %>%
  mutate(scalehpd_25= (e_hpd25-min(e_hpd25))/(max(e_hpd25)-min(e_hpd25)))

# join to full dataset
bio_full_hpd <- bio_hpd_short %>%
  right_join(bio_short, by = "LATITUDE") %>%
dplyr::select(-STUDY_ID_PLOT.x, -LONGITUDE.y) %>%
  rename(STUDY_ID_PLOT = STUDY_ID_PLOT.y,
    LONGITUDE = LONGITUDE.x) %>%
dplyr::select(STUDY_ID_PLOT, scalehpd_25, e_hpd e_hpd25, e_hpd50, e_hpd100)

```

```

### creating global grid cell variable ----
#Construct a global grid with cells approximately 100km2
#(res= 12 equivalates 95.97785km2 cell area)
dggs <- dgconstruct(res = 12, metric=FALSE, resround='down')

#Get the corresponding grid cells for each lat-long pair
bio_short$cell <- dgGEO_to_SEQNUM(dggs,bio_short$LONGITUDE,bio_short$LATITUDE)$seqnum

#Get the number of time-series in each equally-sized cell
biocounts <- bio_short %>% group_by(cell) %>% summarise(count=n())

### joining dataset with all important variables ----
data1 <- beta_Jaccard %>%
  left_join(bio_full_acc, by = "STUDY_ID_PLOT") %>%
  left_join(bio_full_hpd, by = "STUDY_ID_PLOT") %>%
  left_join(bio_short, by = "STUDY_ID_PLOT")

```

## Data Analysis

```

### Data analysis ----
# Models ----
# main model
mo_tu_simp6 <- brm(bf(Jtu ~ scaleacc_25 + scalehpd_25 + duration_plot + TAXA +
  (1|cell) + (1|STUDY_ID)),
  family = zero_one_inflated_beta(),
  data = data1,
  iter = 4000,
  warmup = 1000,
  inits = '0',
  control = list(adapt_delta = 0.85),
  cores = 4, chains = 4)

```

```

save(mo_tu_simp6, file = "outputs/mo_tu_simp6.RData")

# sensitivity analysis plants
data1_plants <- data1 %>%
  filter(TAXA == "Terrestrial plants")

mo_tu_simp_plants <- brm(bf(Jtu ~ scaleacc_25 + scalehpd_25 + duration_plot +
  (1|STUDY_ID)), # or with (1|cell)?
  family = zero_one_inflated_beta(),
  data = data1_plants,
  iter = 4000,
  warmup = 1000,
  inits = '0',
  control = list(adapt_delta = 0.85),
  cores = 4, chains = 4)

save(mo_tu_simp_plants, file = "outputs/mo_tu_simp_plants.RData")

# sensitivity analysis scale (1km)
mo_tu_simp8 <- brm(bf(Jtu ~ scaleacc_1 + scalehpd_1 + duration_plot + TAXA +
  (1|cell) + (1|STUDY_ID)),
  family = zero_one_inflated_beta(),
  data = data1,
  iter = 4000,
  warmup = 1000,
  inits = '0',
  control = list(adapt_delta = 0.85),
  cores = 4, chains = 4)

save(mo_tu_simp8, file = "outputs/mo_tu_simp8.RData")

# sensitivity analysis scale (50km)

```

```

mo_tu_simp7 <- brm(bf(Jtu ~ scaleacc_50 + scalehpd_50 + duration_plot + TAXA +
                    (1|STUDY_ID)), # or with (1|cell)?
                  family = zero_one_inflated_beta(),
                  data = data1,
                  iter = 4000,
                  warmup = 1000,
                  inits = '0',
                  control = list(adapt_delta = 0.85),
                  cores = 4, chains = 4)

save(mo_tu_simp7, file = "outputs/mo_tu_simp7.RData")

# sensitivity analysis scale 100km
mo_tu_scale100 <- brm(bf(Jtu ~ scaleacc_100 + scalehpd_100 + duration_plot + TAXA +
                        (1|cell) + (1|STUDY_ID)),
                    family = zero_one_inflated_beta(),
                    data = data1,
                    iter = 4000,
                    warmup = 1000,
                    inits = '0',
                    control = list(adapt_delta = 0.85),
                    cores = 2, chains = 4)

save(mo_tu_scale100, file = "outputs/mo_tu_scale100.RData")

```

## Data Visualisation

```

# create colour palette ----
display.brewer.pal(n = 8, name = 'Dark2')
brewer.pal(n = 8, name = "Dark2")
display.brewer.pal(n=4, name = "Set1")
brewer.pal(n=4, name = "Set1")
taxa.palette <- c("#D95F02", "#7570B3", "#E7298A", "#E6AB02")

```

```

names(taxa.palette) <- levels(data1$TAXA)

# clean theme ----
theme_clean <- function(){
  theme_bw() +
    theme(axis.text.x = element_text(size = 14),
          axis.text.y = element_text(size = 14),
          axis.title.x = element_text(size = 14, face = "plain"),
          axis.title.y = element_text(size = 14, face = "plain"),
          panel.grid.major.x = element_blank(),
          panel.grid.minor.x = element_blank(),
          panel.grid.minor.y = element_blank(),
          panel.grid.major.y = element_blank(),
          plot.margin = unit(c(0.5, 0.5, 0.5, 0.5), units = , "cm"),
          plot.title = element_text(size = 15, vjust = 1, hjust = 0.5),
          legend.text = element_text(size = 12, face = "italic"),
          legend.title = element_text(size = 12, face = "bold"),
          legend.position = c(0.2, 0.8))
}

#### panel time-series ----
# spatial distribution of time-series ----
(map_studies2 <- ggplot(bio_short,
                        aes(x = LONGITUDE, y = LATITUDE,
                            colour = TAXA, size = NUMBER_OF_SAMPLES), alpha = I(0.7)) +
  borders("world", colour = "gray88", fill = "gray88", size = 0.3) +
  coord_cartesian(xlim = NULL, ylim = NULL, expand = TRUE) +
  theme_map() +
  geom_point(range = c(7,15)) +
  scale_size_continuous(range = c(3,10)) +
  guides(size = FALSE) +

```

```

scale_colour_manual(values = taxa.palette) +
scale_fill_manual(labels = c("Terrestrial plants",
                              "Birds",
                              "Mammals",
                              "Terrestrial invertebrates")) +
labs(title = ("\n\n a) Spatial distribution of time-series\n")) +
theme(legend.position= "bottom",
      legend.title = element_blank(),
      legend.text = element_text(size = 14),
      legend.justification = "top",
      plot.title = element_text(size = 14, hjust = 0.5, face = "bold"))

ggsave(map_studies2, filename = "outputs/map_studies2.png",
       height = 5, width = 8)

# temporal distribution of biodiversity time-series ----
# making id variable as factor
bio_short$STUDY_ID_PLOT <- as.factor(as.character(bio_short$STUDY_ID_PLOT))

# create a sorting variable
bio_short$sort <- bio_short$TAXA
bio_short$sort <- factor(bio_short$sort, levels = c("Terrestrial plants",
                                                    "Birds",
                                                    "Mammals",
                                                    "Terrestrial invertebrates"),
                        labels = c(1,2,3,4))

bio_short$sort <- paste0(bio_short$sort, bio_short$START_YEAR)
bio_short$sort <- as.numeric(as.character(bio_short$sort))

(timeline_studies2 <- ggplot() +
  geom_linerange(data = bio_short, aes(ymin = START_YEAR, ymax = END_YEAR,

```

```

        colour = TAXA,
        x = fct_reorder(STUDY_ID_PLOT, desc(sort))),
        size = 1) +
scale_colour_manual(values = taxa.palette) +
labs(x = NULL, y = NULL,
      title = ("\n\n b) Temporal distribution of time-series\n")) +
#theme_clean() +
coord_flip() +
guides(colour = F) +
theme_bw() +
theme(panel.grid.minor = element_blank(),
      panel.grid.major.y = element_blank(),
      panel.grid.major.x = element_line(),
      panel.border = element_blank(),
      panel.background = element_blank(),
      axis.ticks = element_blank(),
      legend.position = "bottom",
      legend.title = element_blank(),
      axis.title.y = element_blank(),
      axis.text.y = element_blank(),
      axis.ticks.y = element_blank(),
      plot.title = element_text(size = 14, hjust = 0.5, face = "bold"),
      axis.text = element_text(size = 16),
      axis.title = element_text(size = 20)))

ggsave(timeline_studies2, filename = "outputs/timeline_studies2.png",
       height = 5, width = 8)

# taxonomic distribution of biodiversity time-series ----
# calculating sample size for each taxa
taxa_sum <- bio_short %>% group_by(TAXA) %>% tally

(taxa_studies2 <- ggplot(taxa_sum, aes(area = n, fill = TAXA, label = n,

```



```

subgroup = TAXA)) +

geom_treemap() +
geom_treemap_subgroup_border(colour = "white", size = 1) +
geom_treemap_text(colour = "white", place = "center", reflow = T) +
scale_colour_manual(values = taxa.palette) +
scale_fill_manual(values = taxa.palette) +
labs(title = ("\n\n c) Taxonomic distribution of time-series\n")) +
theme(plot.title = element_text(size = 14, hjust = 0.5, face = "bold")) +
guides(fill= FALSE))

ggsave(taxa_studies2, filename = "outputs/taxa_studies2.png",
       height = 5, width = 8)

# panel ----
panel_b2 <- ggarrange(timeline_studies2, taxa_studies2, ncol = 2, align = c("h"))

panel_full2 <- grid.arrange(map_studies2, panel_b2, nrow = 2)

ggsave(panel_full2, filename = "outputs/panel_studies2.png",
       height = 10, width = 8)

#### RQ1: turnover ~ accessibility + duration ----
# prediction
pred_acc_d <- ggpredict(mo_tu_simp6, terms = c("scaleacc_25", "duration_plot"))

pred_acc_d$Monitoring <- factor(pred_acc_d$group, levels = c("6", "19.1", "32.2"),
                              labels = c("Short-term", "Moderate", "Long-term"))

# calculate change
pred_acc <- ggpredict(mo_tu_simp6, terms = c("scaleacc_25"))

# model vis

```

```

(graph_acc <- ggplot() +
  geom_line(data = pred_acc_d, aes(x = x, y = predicted, color = Monitoring),
    size = 2) +
  geom_ribbon(data = pred_acc_d, aes(ymin = conf.low, ymax = conf.high,
    x = x, fill = Monitoring), alpha = 0.1) +
  geom_point(data = data1, aes(x = scaleacc_25, y = Jtu),
    alpha = 0.1, size = 2) +
  theme_clean() +
  scale_fill_manual(values = c("#7CFC00", "#66A61E", "#1B9E77")) +
  scale_colour_manual(values = c("#7CFC00", "#66A61E", "#1B9E77")) +
  labs(x = "\nAccessibility (proportion)", y = "Turnover\n") +
  theme(legend.position = c(0.2, 0.8)))

ggsave(graph_acc, file = "outputs/graph_acc.png", width = 7, height = 5)

(acc_d <- ggMarginal(graph_acc, type="density", size = 3,
  fill = "#666666", xparams = list(fill = "#A6761D"))))

ggsave(acc_d, file = "outputs/graph_acc_den.png" , width = 7, height = 5)

#### RQ2: turnover ~ hpd ----
# prediction
pred_hpd <- ggpredict(mo_tu_simp6, terms = c("scalehpd_25"))

# model vis
(graph_hpd <- ggplot() +
  geom_line(data = pred_hpd, aes(x = x, y = predicted),
    size = 2, color = "#A6761D") +
  geom_ribbon(data = pred_hpd, aes(ymin = conf.low, ymax = conf.high,
    x = x), fill = "#A6761D", alpha = 0.1) +
  geom_point(data = data1, aes(x = scalehpd_25, y = Jtu),
    alpha = 0.1, size = 2) +
  theme_clean() +

```

```

labs(x = "\nHuman population density (proportion)", y = "Turnover\n") +
theme(legend.position = c(0.85, 0.8)))

ggsave(graph_hpd, file = "outputs/graph_hpd.png", width = 7, height = 5)

(hpd_d <- ggMarginal(graph_hpd, type="density", size = 3,
  fill = "#666666", xparams = list(fill = "#A6761D"))))

ggsave(hpd_d, file = "outputs/graph_hpd_den.png" , width = 7, height = 5)

#### RQ3: turnover ~ taxa ----
# raincloud plot
# by Ben Marwick
source("https://gist.githubusercontent.com/benmarwick/
  #2a1bb0133ff568cbe28d/raw/fb53bd97121f7f9ce947837ef1a4c65a73bffb3f/
  #geom_flat_violin.R")

# raincloud plot taxa jtu
(raincloud_taxa <-
  ggplot(data = data1,
    aes(x = reorder(TAXA, desc(Jtu)), y = Jtu, fill = TAXA)) +
  geom_flat_violin(position = position_nudge(x = 0.2, y = 0), alpha = 0.8) +
  geom_point(aes(y = Jtu, color = TAXA),
    position = position_jitter(width = 0.15), size = 1, alpha = 0.1) +
  geom_boxplot(width = 0.2, outlier.shape = NA, alpha = 0.8) +
  labs(y = "\n Turnover", x = NULL) +
  guides(fill = FALSE, color = FALSE) +
  scale_y_continuous(limits = c(0, 1)) +
  scale_fill_manual(values = taxa.palette) +
  scale_colour_manual(values = taxa.palette) +
  coord_flip() +
  theme_clean())

```

```

ggsave(raincloud_taxa, filename = "outputs/raincloud_taxa_graph.png",
       height = 5, width = 8)

# effect size graph
pred_taxa <- ggpredict(mo_tu_simp6, terms = c("TAXA"))

(graph_taxa <- ggplot(pred_taxa, aes(x = reorder(x, desc(predicted)),
                                   y = predicted, color = x)) +

  geom_point(aes(size = 1)) +
  geom_pointrange(aes(ymin = conf.low, ymax = conf.high), size = 2) +
  scale_fill_manual(values = taxa.palette) +
  scale_colour_manual(values = taxa.palette) +
  theme_clean() +
  theme(legend.position = "none",
        axis.text.y=element_blank(),
        axis.title.y=element_blank()) +
  geom_hline(yintercept=0, linetype="dashed", size=1) +
  labs(x = NULL, y = "Effect size\n") +
  coord_flip())

ggsave(graph_taxa, filename = "outputs/taxa_model.png",
       height = 5, width = 5)

panel_taxa <- ggarrange(raincloud_taxa, graph_taxa, labels = c("A", "B"), ncol = 2)

ggsave(panel_taxa, filename = "outputs/panel_taxa3.png",
       height = 5, width = 8)

#### sensitivtiy analysis ----
# plants vs all ----
pred_plants <- ggpredict(mo_tu_simp_plants, terms = c("scaleacc_25"))
pred_all <- ggpredict(mo_tu_simp6, terms = c("scaleacc_25"))

```

```

data1_plants <- data1 %>%
  filter(TAXA == "Terrestrial plants")

data1_nopl <- data1 %>%
  filter(!TAXA == "Terrestrial plants")

(graph_plants <- ggplot() +
  geom_line(data = pred_plants, color = "#E6AB02", aes(x = x, y = predicted),
    size = 2) +
  geom_ribbon(data = pred_plants, color = "#E6AB02", aes(ymin = conf.low,
    ymax = conf.high, x = x),
    alpha = 0.1, fill = "#E6AB02") +
  geom_point(data = data1_plants, color = "#E6AB02", aes(x = scaleacc_25, y = Jtu),
    alpha = 0.1, size = 2) +
  geom_line(data = pred_all, color = "#66A61E", aes(x = x, y = predicted),
    size = 2) +
  geom_ribbon(data = pred_all, color = "#66A61E", aes(ymin = conf.low, ymax = conf.high,
    x = x), alpha = 0.1,
    fill = "#66A61E") +
  geom_point(data = data1_nopl, color = "#66A61E", aes(x = scaleacc_25, y = Jtu),
    alpha = 0.1, size = 2) +
  theme_clean() +
  labs(x = "\nAccessibility", y = "Turnover\n"))

ggsave(graph_plants, filename = "outputs/graph_plants.png", height = 5, width = 8)

# scales ----
pred_1 <- ggpredict(mo_tu_simp8, terms = c("scaleacc_1"))
pred_25 <- ggpredict(mo_tu_simp6, terms = c("scaleacc_25"))
pred_50 <- ggpredict(mo_tu_simp7, terms = c("scaleacc_50"))
pred_100 <- ggpredict(mo_tu_scale100, terms = c("scaleacc_100"))

(graph_scale <- ggplot() +

```

```

geom_line(data = pred_1, color = "#E41A1C", aes(x = x, y = predicted),
          size = 2) +
geom_ribbon(data = pred_1, color = "#E41A1C", aes(ymin = conf.low, ymax = conf.high,
          x = x), alpha = 0.1,
          fill = "#E41A1C") +
geom_line(data = pred_25, color = "#377EB8", aes(x = x, y = predicted),
          size = 2) +
geom_ribbon(data = pred_25, color = "#377EB8", aes(ymin = conf.low, ymax = conf.high,
          x = x), alpha = 0.1,
          fill = "#377EB8" ) +
geom_line(data = pred_50, color = "#4DAF4A", aes(x = x, y = predicted),
          size = 2) +
geom_ribbon(data = pred_50, color = "#4DAF4A", aes(ymin = conf.low, ymax = conf.high,
          x = x), alpha = 0.1,
          fill = "#4DAF4A" ) +
geom_line(data = pred_100, color = "#984EA3", aes(x = x, y = predicted),
          size = 2) +
geom_ribbon(data = pred_100, color = "#984EA3", aes(ymin = conf.low, ymax = conf.high,
          x = x), alpha = 0.1,
          fill = "#984EA3") +
theme_clean() +
labs(x = "\nAccessibility", y = "Turnover\n") +
ylim(0, 1))

ggsave(graph_scale, filename = "outputs/graph_scale.png", height = 5, width = 8)
#### other additional graphs ----
# raincloud plot taxa accessibility ----
(raincloud_acc <-
  ggplot(data = data1,
          aes(x = reorder(TAXA, desc(scaleacc_25)), y = scaleacc_25, fill = TAXA)) +
  geom_flat_violin(position = position_nudge(x = 0.2, y = 0), alpha = 0.8) +
  geom_point(aes(y = scaleacc_25, color = TAXA),
             position = position_jitter(width = 0.15), size = 1, alpha = 0.1) +

```

```

geom_boxplot(width = 0.2, outlier.shape = NA, alpha = 0.8) +
labs(y = "\nAccessibility (proportion)", x = NULL) +
guides(fill = FALSE, color = FALSE) +
scale_fill_manual(values = taxa.palette) +
scale_colour_manual(values = taxa.palette) +
coord_flip() +
theme_clean())

ggsave(raincloud_acc, filename = "outputs/raincloud_taxa_acc_graph.png",
       height = 5, width = 8)

# protected areas ----
(graph_protected_areas <- ggplot() +
  geom_point(data = data1, aes(x = scaleacc_25, y = Jtu, colour = PROTECTED_AREA),
            alpha = 0.8, size = 1.5) +
  scale_colour_manual(values = c("#E6AB02", "#666666")) +
  theme_clean() +
  theme(legend.position = "none") +
  labs(x = "\nAccessibility", y = "Turnover\n"))

ggsave(graph_protected_areas, filename = "outputs/graph_protected_areas.png",
       height = 5, width = 8)

# graph duration and turnover ----
pred_duration <- ggpredict(mo_tu_simp6, terms = c("duration_plot"))

(graph_duration <- ggplot() +
  geom_line(data = pred_duration, aes(x = x, y = predicted),
           size = 2, color = "#666666") +
  geom_ribbon(data = pred_duration, aes(ymin = conf.low, ymax = conf.high,
                                       x = x), fill = "#666666", alpha = 0.1) +
  theme_clean() +
  labs(x = "\nMonitoring duration (in years)", y = "Turnover\n"))

```

```

ylim(0, 1))

ggsave(graph_duration, file = "outputs/graph_duration.png", width = 7, height = 5)

# histogram distribution of random and real values ----
# accessibility
bio_short <- bio %>%
  distinct(STUDY_ID_PLOT, .keep_all = TRUE)

SP <- bio_short %>%
  dplyr::select(LATITUDE, LONGITUDE, STUDY_ID_PLOT) %>%
  distinct(LATITUDE, .keep_all = TRUE)

library(generator)
fake_lat <- r_latitudes(1023)
fake_long <- r_longitudes(1023)

fake_ll <- SP %>%
  mutate(fake_lat = c(fake_lat),
         fake_long = c(fake_long)) %>%
  dplyr::select(- LATITUDE, -LONGITUDE)

f_points <- cbind(fake_ll$fake_long, fake_ll$fake_lat)

f_sppoints <- SpatialPoints(f_points, proj4string=CRS('+proj=longlat +datum=WGS84'))
f_tp <- spTransform(f_sppoints, crs(aa))

f_e <- extract(aa, f_tp)

f_bio_aa <- cbind(fake_ll, f_e)
f_bio_aa_short <- na.omit(f_bio_aa)

```



```

# hpd
f_tp_hpd <- spTransform(f_sppoints, crs(hpd))

f_e_hpd <- extract(hpd, f_tp_hpd)

f_bio_hpd <- cbind(fake_ll, f_e_hpd)
f_bio_hpd_short <- na.omit(f_bio_hpd)

# figure overall
png("outputs/fake_ll.png", width=600, height=350)
par(
  mfrow=c(1,2),
  mar=c(4,4,1,0)
)
hist(f_bio_aa_short$f_e, col=rgb(1,0,0,0.5), xlab="Accessibility score",
      ylab = "Number of points", main = "") # more normal distributed?
hist(data1$e, col=rgb(0,0,1,0.5), add=T)
hist(f_bio_hpd_short$f_e_hpd, col=rgb(1,0,0,0.5), xlab="Human population density score",
      ylab = "Number of points", main = "")
hist(data1$e_hpd, col=rgb(0,0,1,0.5), add=T)
legend("topright", legend=c("Random", "BioTIME"), col=c(rgb(1,0,0,0.5),
                                                         rgb(0,0,1,0.5)), pt.cex=2, pch=15)
dev.off()

# histogram acc and hpd ----
png("outputs/hist_acc_hpd.png", width=600, height=350)
hist(data1$scaleacc_25, col = "#1B9E77", xlab="Accessibility/Human population density",
      ylab = "Number of points", main = "")
hist(data1$scalehpd_25, col = "#A6761D", add = T)
legend("top", legend=c("Accessibility", "Human population density"),
      col=c("#1B9E77", "#A6761D"), pt.cex=2, pch=15)
dev.off()

```