Representation of global change across biodiversity databases

Authors: Gergana N. Daskalova[1]*, Diana E. Bowler[2], Isla H. Myers-Smith[1], Maria Dornelas[3]

Affiliations:

[1] School of GeoSciences, University of Edinburgh, West Mains Road, Edinburgh EH9 3FF, Scotland

[2] German Centre for Integrative Biodiversity Research (iDiv), Deutscher Pl. 5E, Leipzig 04103, Germany

[3] Centre for Biological Diversity, University of St Andrews, Greenside Place, St Andrews KY16 9TF, Scotland

*Correspondence to:

Gergana Daskalova, Room 401, Crew Building, School of GeoSciences

University of Edinburgh, West Mains Road, Scotland

gndaskalova@gmail.com

Abstract

Global change has altered biodiversity and impacted ecosystem functions and services around the planet. Understanding the effects of anthropogenic drivers like human use and climate change and predicting trajectories of future biodiversity change have become key challenges for science and policy. However, our knowledge of biodiversity change is limited by the available data and their biases. In both the terrestrial and marine realms, we test the representation of three worldwide biodiversity databases (Living Planet, BioTIME and PREDICTS) across geographic and temporal variation in global change and across the tree of life. We find that variation in global change drivers is better captured over space than over time and in the marine realm versus on land. We provide recommendations to improve the use of existing data, better target future ecological monitoring and capture different combinations of global change.

One sentence summary

Biodiversity data capture over half of the global change spectrum, yet substantial data gaps remain and need to be filled to capture ongoing change and better predict future trajectories for Earth's biota.

Introduction

Global change is reshaping the planet from the tropics to the poles and across land and sea[1–3], and the Earth's biodiversity is shifting in response[4]. Paralleled with this rapid biotic reorganisation, an ecological data revolution is underway with more open-access data available now than ever before[5,6]. Large-scale data compilations (e.g., Living Planet[7], BioTIME[8], PREDICTS[9], GBIF[10], TetraDensity[11]) have been analyzed to test general patterns of biodiversity change across the world and impacts of anthropogenic drivers[4,12–16]. Such studies have revealed a wide spectrum of biodiversity change, including both increases and decreases, with trends quantified over time[17–20], space[16,21] and taxa[22,23]. The biodiversity data underlying many of these syntheses (e.g., time series, occurrence records and space-for-time surveys) have already been shown to be biased geographically and taxonomically[24–27], but much less attention has been given to whether the data are also biased with respect to global change within and across biodiversity databases (but see [28] for spatio-taxonomic biases in North American butterfly occurrence records). The uneven sampling of global change drivers, like other biases, can hinder our ability to detect real patterns in syntheses of biodiversity databases.

Insights from large-scale data syntheses inform trajectories of past, current and future changes in the Earth's biota[4,7,29,30]. Such syntheses usually estimate the mean biodiversity change across all data as well as the variation in detected trends and their relationships with different global change drivers. Both the mean and variance of biodiversity estimates in syntheses findings can be affected by database biases related to global change variation over space and time, as well as spatial and taxonomic sampling[24,26,31–34]. To assess mean biodiversity change and its variation across different regions, taxa and time periods, it is essential to sample representatively spatially, taxonomically and temporally but also representing drivers of change[35–38]. To capture ecological tipping points and identify non-linear and lagged responses, we require sufficiently large samples over time that capture extreme driver intensity[35,39]. For example, we need to consider if studies include a range of different driver intensities, solely heavily impacted sites or exclusively intact wilderness areas. Additionally, we need to determine if global change is sampled representatively. A dataset that includes a higher proportion of intact grassland sites relative to the global average intensity of human impact in grassy biomes could underestimate the impact of global change on grassland biodiversity. Inferred or projected regional or global-scale change from data compilations of local sites will be biased unless we account for the representativeness of the underlying data across multiple axes, including global change.

Here, we present our perspective on quantifying the representativeness of biodiversity data in large-scale syntheses and discuss implications for interpreting their findings and implications for policy and conservation. To illustrate our views, we focus on four aspects of representativeness – global change intensity over space, global change intensity over time, geography, and taxonomy. Particularly, we highlight the patterns in spatial and temporal sampling of global change drivers as under-explored types of bias, but whose consideration can improve our understanding of the re-organisation and change in biological communities[4,35,36]. Our perspective serves two important purposes: 1) to highlight the variation in global change drivers that are already captured by global datasets and hence the drivers that we can test in ongoing studies, and 2) to identify the gaps in data representativeness that future surveys, monitoring and data mobilization actions should target with further data collection. With our recommendations on capturing current and novel global change, we can work towards biodiversity data that are representative across multiple axes of natural and anthropogenic variation and produce more accurate future predictions.

We linked three of the largest currently existing open-access biodiversity databases (Living Planet[7] - marine and terrestrial; BioTIME[8] - marine and terrestrial; and PREDICTS[9] - terrestrial) with maps of global change drivers[3]. To measure how well these databases capture and sample variation in global change intensity over space, we estimated driver variation around the world and determined the 'global change space' (similar to the quantification of trait space in [40]). We then mapped the location of marine and terrestrial sites onto the global change space, highlighting under- and over- representation and sampling intensity. To quantify representation of global change intensity over time, we focused on climate change and land cover change across terrestrial sites with time series data from the Living Planet and BioTIME databases. At each site, we compared the amounts of change that occurred before relative to during the periods of biodiversity monitoring. To estimate geographic representation, we mapped site locations of the three databases and determined sampling intensity across a spatial grid covering the planet. Finally, to estimate taxonomic representation, we calculated the percentage of known species included in the three databases. As we launch the next stage of biodiversity syntheses, scenarios and conservation goals, we argue that biodiversity science needs to move towards improved

representation in biodiversity data, including the heterogeneous distribution and sampling of global change.

Biodiversity data capture most of the spatial variation in global change space

Overall, we found that biodiversity data from the Living Planet, BioTIME and PREDICTS databases capture a surprisingly high amount of the spatial variation in global change intensity around the planet, particularly in the marine realm where 20% of the possible values along the global change spectrum were represented by the BioTIME database, versus 2% in the terrestrial realm by the PREDICTS database (Figure 1, Table S1). Among the five global change drivers we tested over land and sea, climate change and pollution in the marine realm were sampled the most representatively, suggesting that we can test the effects of these drivers with higher confidence and the underlying data could be used when creating global scenarios for the future (Figure 2). In contrast, human use intensity on land ranged between 0 and 1, with a relatively uniform distribution of 0 to 0.7 values across the world, but the three databases we tested predominantly sampled places with higher human use from 0.4 onwards. In the marine realm, the BioTIME database captured more global change variation than the Living Planet database, but over land, the sites in the Living Planet and PREDICTS databases occupied more of the global change space. Particularly, the Living Planet database sampled both positive and negative extremes of human use, pollution, invasion pressure and human population density. The different relative positions of these databases within global change space can be used to explain some of the differences in findings (e.g. predominantly negative impacts of intensifying land-use change based on PREDICTS[16], both negative and positive influences of forest loss based on Living Planet and BioTIME[13] and stronger impacts of temperature change on richness, composition and abundance trends in BioTIME[14]). Across both realms but particularly strongly over land, all databases were lacking sites that have experienced high amounts of climate change, reflecting geographic gaps in data collection in places like the Arctic. The consistent gaps in how global change is sampled in biodiversity data can bias the inferences we draw from data syntheses as well as the scenarios for the future of Earth's biota which often include large databases as a base upon which predictions are built. Drawing a parallel with experimental design, manipulative studies often include low, medium and high levels of treatment to have sufficient power to statistically detect the effect of the imposed change[41], and we suggest extending this perspective to observational studies and data syntheses. Using global change space, we can better

interpret the findings of large-scale syntheses and target future ecological monitoring to fill in the gaps in global change space and capture the variety of ways in which humans are altering the planet.
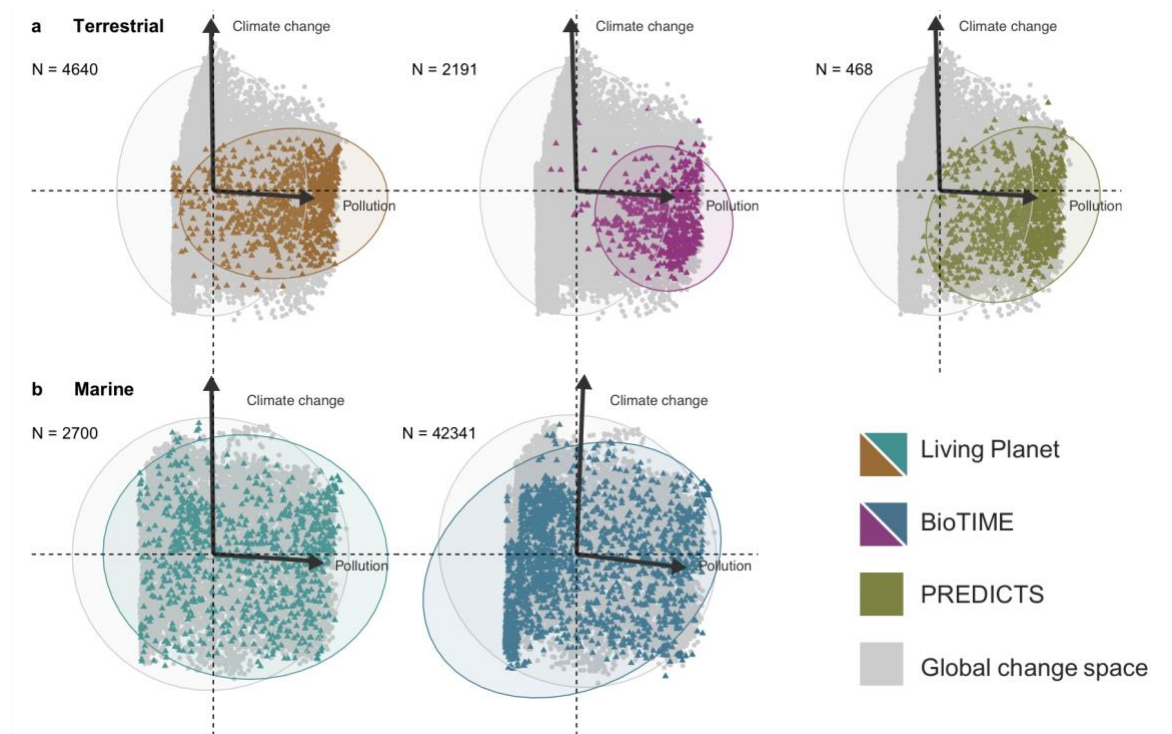

Figure 1. Biodiversity data capture spatial variation in global change space better in the marine versus terrestrial realm. Figure shows Principal Component Analysis of the magnitudes of human use, climate change, human population density, pollution and invasion potential across the locations of the Living Planet, BioTIME and PREDICTS databases as well as one million randomly sampled locations across the full extent of the globe (in grey). PCA axes omitted for visualisation purposes. Arrows show direction and magnitude of PCA scores. Human use, human population density and invasion potential were very correlated with pollution and their arrows were omitted for clarity. For details on the global change driver layers, see Bowler et al. 2020. Annotations on figure show sample size (N).
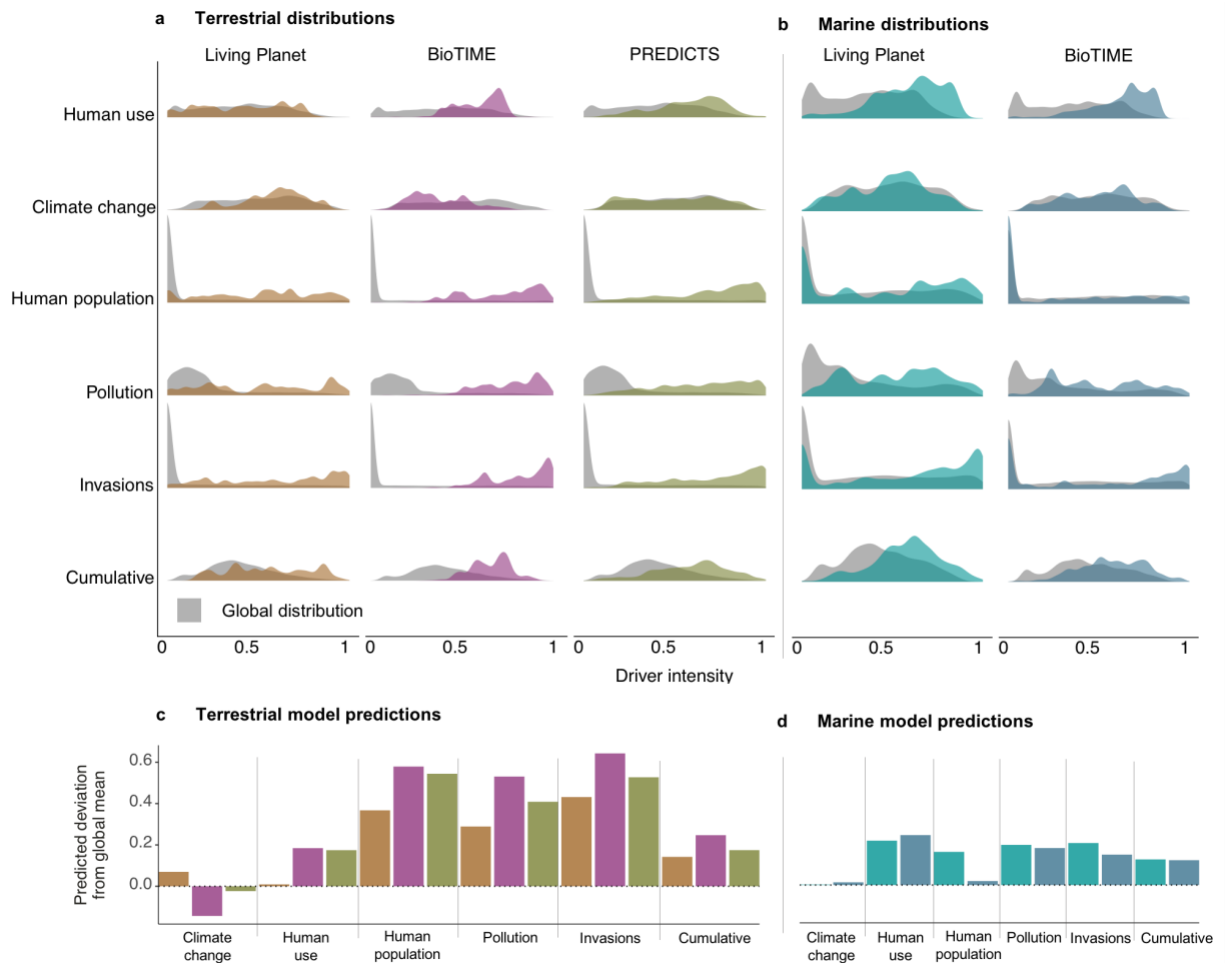
Figure 2. Higher magnitudes of global change drivers are overrepresented in biodiversity data. Panels **a** and **b** show distributions of the raw global change driver data from random sampling spanning the globe and sites from existing biodiversity databases. Panels **c** and **d** show effect sizes of statistical models comparing the individual and cumulative magnitudes of global change drivers across the Living Planet, BioTIME and PREDICTS databases relative to the global average. Positive effect sizes indicate higher magnitudes than the global average and negative effect sizes indicate lower magnitudes. Because of the large sample sizes included in the models, the errors around the effect sizes were too small to be visualised in the figure.

Biodiversity data often miss the temporal peaks of land cover change but capture those of climate change

We found mismatches between global change and the timing of biodiversity data collection that were more frequent for land-use change than for climate change (Figure 3). While it is well-known that peak land-use conversion often predates ecological monitoring by centuries to millennia (e.g., [1,42,43]), studies rarely quantify the magnitude of this mismatch or account for the long-term trajectory and historic baseline of global change drivers (but see [13,21]). For drivers such as forest loss, the peak often occurred tens to hundreds of years before the start of most biodiversity monitoring[13]. In contrast for climate warming, we found that the majority of the Living Planet and BioTIME time series (76% and 56% of terrestrial time series, and 64% and 59% of marine time series, respectively) have experienced larger magnitudes of warming during the period of monitoring when compared to the period preceding data collection, suggesting that biodiversity data capture contemporary warming (Figure 3c-e). This variation in global change driver intensity over time can influence the strength of relationships detected in attribution analyses[13,35,39] and can obscure assessment of biodiversity trends in ecosystems with tipping points[44]. Monitoring schemes that start well after the peak magnitude of a global change driver will likely underestimate its impacts on biodiversity[35], but equally, lagged biodiversity change might mean that the effects of drivers like agriculture persist decades after farming has ceased[39]. Such interactions between lagged biodiversity responses to disturbance and temporal variability of global change have produced heterogeneous and often non-linear biodiversity trends, as has already been reported for many taxa, including birds[45], moths[20] and wasps[46]. Additionally, analyses of observational datasets with small sample sizes over time have reduced statistical power and thus might not be able to detect the effect of global change drivers, a phenomenon also found in experimental studies[47]. It is critical to incorporate the temporal match and mismatch of ecological monitoring and global change drivers to improve attribution analyses of biodiversity change.

a **Forest loss over time (Living Planet)**

b **Forest loss over time (BioTIME)**

c **Climate warming over time (Living Planet)**

Warming before, Warming during

Cooling before, Warming during

Cooling before, cooling during

Cooling before, warming during

d **Climate warming over time (BioTIME terrestrial)**

e **Climate warming over time (Living Planet marine)**

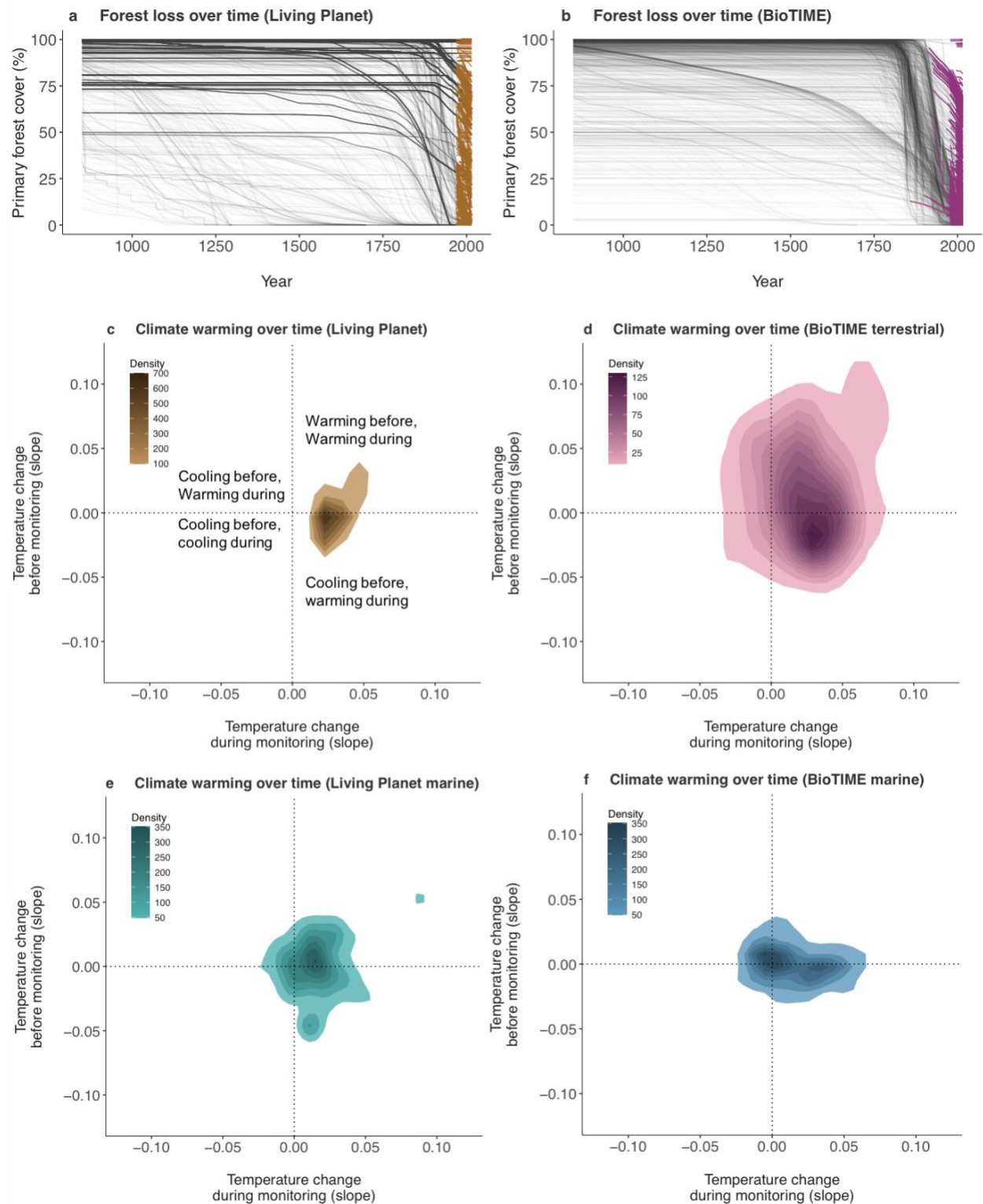**f** **Climate warming over time (BioTIME marine)**

Figure 3. The majority of primary forest was lost by the time ecological monitoring began whereas high magnitudes of climate warming predominantly occurred during the time series. Panels **a** and **b** show

the temporal trajectory of primary forest loss across sites part of the Living Planet (N = 4640) and BioTIME (N = 2191) databases. The primary forest cover estimates show proportions based on the LUH database (Hurtt et al. 2013) and were calculated for cells of approximately ˜96km² around the centerpoint of each site. Historic human use time series data of sufficient duration were not available for the marine realm. The periods for comparison in panels **c-e** were the same as the duration of each time series and were always more than five years (for example for a time series starting in 2000 and ending in 2010, we used 1990-2000 as the comparison period). Slope values on **b** and **d** show changes in temperature in degrees Celsius per year, derived from general linear models estimating temperature as a function of year. For the terrestrial realm, surface air temperature was obtained from the CRU TS v4.05 database[48] and for the marine realm, the sea surface temperature data was extracted from the NOAA Extended Reconstructed SST v5 database[49].

Biodiversity data are skewed towards Europe and North America

Geographic gaps exist across both time series data (Living Planet and BioTIME databases) and space-for-time data (PREDICTS database), particularly in tropical and high latitudes. Regions like Northern Asia, Africa and South America had fewer sample sites than Europe and North America across the three databases. Ecoregions in the marine realm were better represented than those in the terrestrial realm, including records from 69% and 48% of marine ecoregions in the Living Planet and BioTIME databases, and 16% to 32% of terrestrial ecoregions in the Living Planet, BioTIME and PREDICTS databases, respectively (Figure 4). Such geographic biases are well-known caveats of biodiversity data (e.g., [24,26,50]) and can be particularly problematic when extrapolating from patchy local-scale data to broad macroecological patterns[4]. For example, studies of insect biodiversity trends from a limited sample of geographic locations have found steep declines (e.g. [51,52]), but the nuance of geographic representation is often lost in media and public communication of population and biodiversity change[53–56]. Geographic gaps in biodiversity data can also translate into a perceived lack of localized conservation evidence, which can then hinder management and policy implementation[4]. Underrepresentation in geographic space, however, did not directly translate into gaps in global change space and thus an incomplete geographic sample can capture a surprising amount of variation in global change driver intensity (Figures 1 and 2). With continued calls for more biodiversity data (e.g. [4,57,58]), we advocate for a broader

perspective on the design and synthesis of future ecological monitoring to target not just geographic gaps, but also underrepresentation of global change.
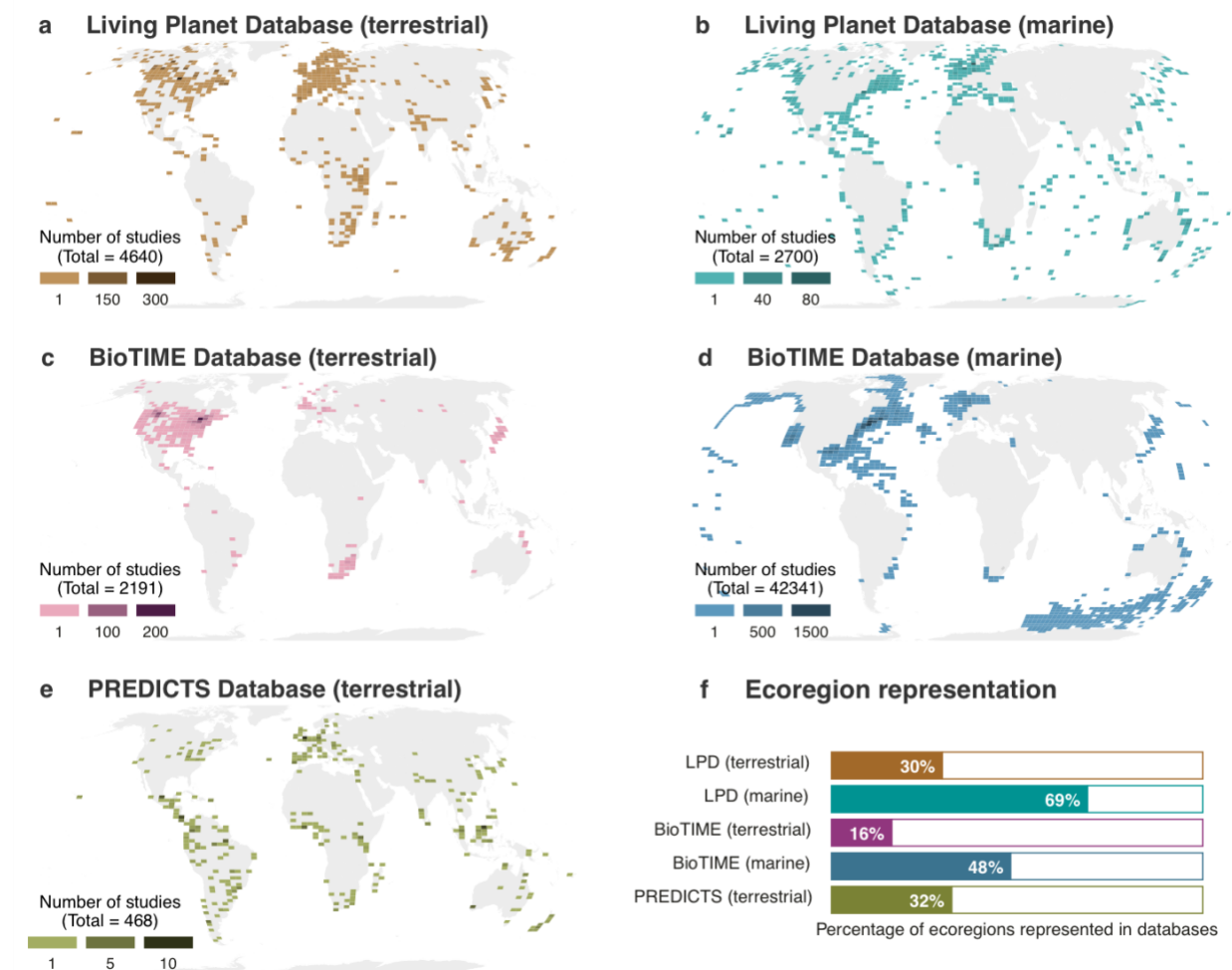

Figure 4. Geographic gaps in biodiversity data exist in both the marine and terrestrial realms but ecoregion representation was higher over sea than over land. Maps show locations of sites from the Living Planet, BioTIME and PREDICTS databases with darker colours indicating higher numbers of sites. Ecoregions are based on the classification of [59].

Biodiversity data are most representative of bird diversity and least representative of arthropods

We found that birds were the best-, and arthropods the worst-represented taxa across the Living Planet, BioTIME and PREDICTS databases (Figure 5). Such patterns in taxonomic representation of biodiversity data are well-established in ecology and are a common feature of many biodiversity datasets[8,9,60]. Within the scientific community, there are frequent calls for better sampling across the

tree of life to capture the variety of ways in which species from the smallest ant to the biggest sequoias and more are being impacted by the Anthropocene (e.g. [61–65]. Recently, invertebrates and in particular insects have been highlighted as a taxon experiencing potential steep declines in abundance and biomass[51,52], yet such findings are confounded by the general paucity of invertebrate data[54–56]. In contrast, birds are the focus of many national and international monitoring schemes and for many species, scientists have established how they are changing over time. These differences in taxonomic representation influence our confidence in biodiversity studies, as well as the interpretation and communication of taxon- and species-specific trends. For less-studied taxa, additional challenges arise when synthesising across data sources and the lack of sufficient data can hinder testing large-scale biogeographic patterns. Species differ in both the magnitude and pace at which they respond to global change[66–68], which suggests that ecological communities will not shift in unison following global change. Consequently, we could be missing signals of compositional change in ecological communities as well as the winners and losers of particular global change drivers due to a lack of representative taxonomic coverage. Understanding and improving taxonomic representation is important for interpreting global patterns as well as scenarios for biodiversity which generalise across several different taxa[69–71]. Extending findings from a limited representation of the planet's diversity to cross-taxa scenarios of change should be done with caution and placed in the context of which species dominate the data.
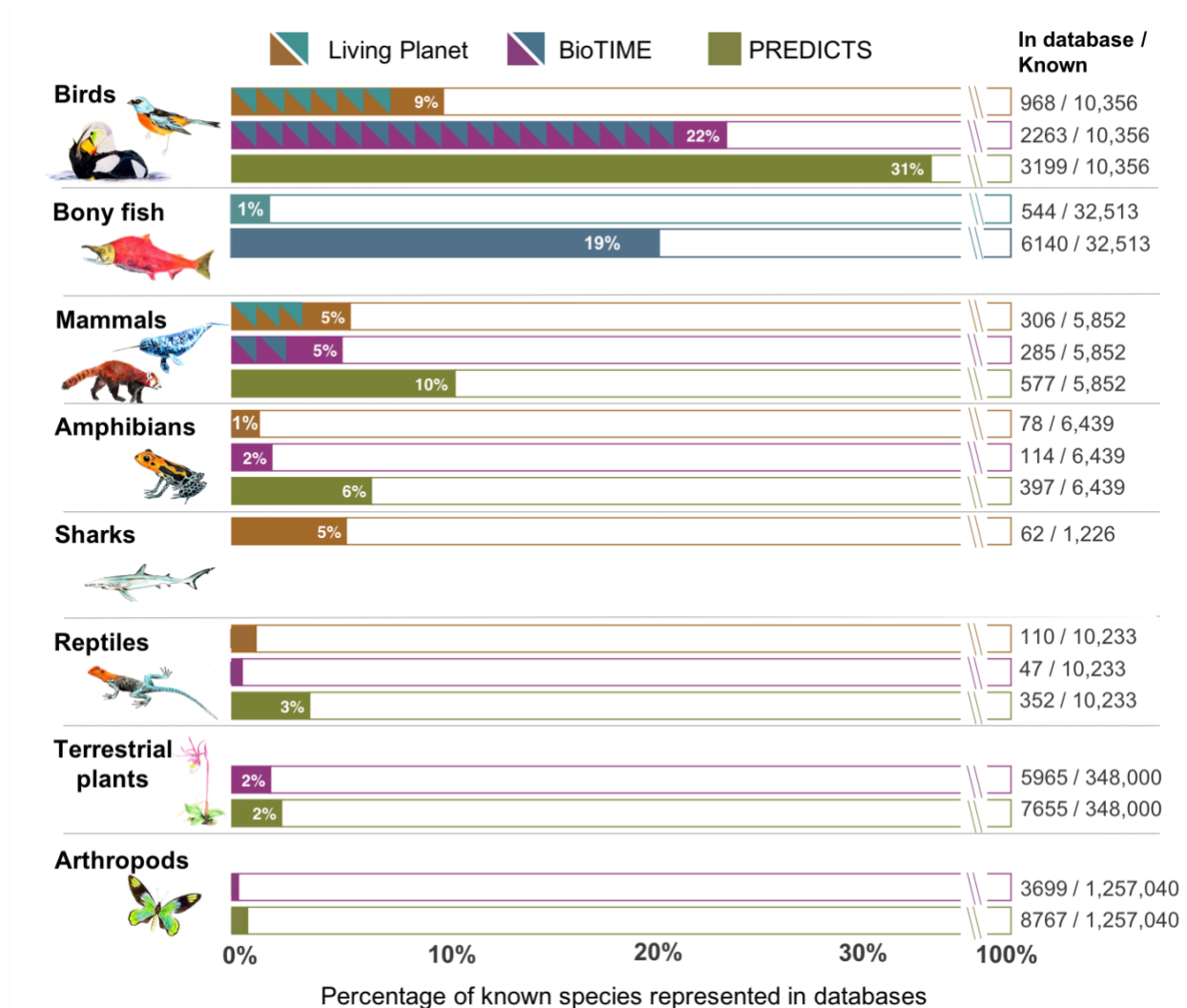
Figure 5. Taxonomic representation of biodiversity data is highest for birds and mammals and lowest for arthropods. Percentages show how many of the known species in each taxon are represented by at least one record in the Living Planet, BioTIME and PREDICTS databases. The numbers of known species per taxa were extracted from the 2019 edition of the Catalogue of Life (http://www.catalogueoflife.org). The values for the birds and mammals in the Living Planet and BioTIME database include both marine and terrestrial species.

Recommendations for capturing the spectrum and distribution of global change across space, time and the tree of life

Understanding ongoing and future biodiversity change can be improved by quantitatively accounting for the representation of biodiversity data across global change space, over the temporal trajectory of drivers, across geographic regions and across the tree of life. Together, our three recommendations provide guidance on using existing observational data, determining where to locate future ecological monitoring and designing experimental studies of novel global change space without modern day analogues.

Recommendation 1: Account for data representation across multiple axes in existing observational data

Extending our thinking beyond just geographic and taxonomic bias to incorporate global change space can contextualise research findings from biodiversity data. We have demonstrated a framework to test data representation across different global change drivers over space and time which can be applied to other datasets. Randomised subsampling can help balance uneven data where certain types of global change are overrepresented while others are underrepresented. Additional practical ways forward include using continuous variables of driver intensity versus categories that explain less variance, as well as adding a variable in statistical models capturing whether the biodiversity data from a given site come from before, during or after peak driver impact. Statistical models can also include uncertainty and global change weightings to account for higher confidence in biodiversity trends derived from sites that better sample global change space (similar to approaches used in citizen science data, e.g. [32,72]). Presenting the results of such tests can help interpret heterogeneous biodiversity trends and guide decisions about what data to include as the base of scenarios for future shifts in the Earth's biota.

Recommendation 2: Prioritise new data collection for underrepresented parts of the global change spectrum and increase data representation

Extrapolating from an uneven geographic, temporal and taxonomic sampling of the world to global trends can be improved by targeting future ecological monitoring programmes in areas of the world

that are experiencing undersampled types of global change. For example, we are currently lacking biodiversity data from places with high magnitudes of climate change. Such data are important not just for understanding current effects of climate change, but also for projecting future impacts as high-climate change areas give a preview of what shifts in the climate might mean for the rest of the planet. Another underrepresented part of the global change spectrum is relatively intact sites with low human impact which are an important counterpoint for testing the impacts of human use, pollution and other global change drivers. Future ecological monitoring can overcome site-selection bias by filling in the gaps in global change space.

Recommendation 3: Design experiments to study novel global change space

Global change space is not static and to make scenarios for future biodiversity trends, we need to sample not only current variation in global change drivers but also future magnitudes. Predicting biodiversity change in novel environments poses numerous challenges and is a common problem is species distribution modelling with implications for ecology and conservation more widely[73]. We suggest that using projections for climate change and human impact, such as IPCC[2] and HYDE[43], we can compute future global change space and determine the novel environments without analogues today. Designing experiments that mirror gaps in future global change space can provide a preview of biodiversity responses to unprecedented environmental conditions and inform policy and conservation to prevent negative impacts.

Conclusion and ways forward

Predicting future biodiversity change and its consequences for ecosystem functions and services to society is an urgent scientific challenge. Global biodiversity monitoring needs to capture a representative sample of the world over both space and time, as well as the full spectrum of global change drivers. In this study, we quantified four types of representativeness - global change intensity over space, global change intensity over time, geography, and taxonomy. Together, our findings demonstrated that global biodiversity datasets do capture the intensity of global change, yet with gaps that would not be predicted from geographic, temporal and taxonomic representativeness alone. We provide a framework to aid the design of future studies and the interpretation of patterns inferred from global biodiversity syntheses. Over space, existing data capture the spatial variance in global change

drivers. Over time, monitoring often starts after the peak intensity in environmental change for drivers like primary forest loss, while the period of biodiversity monitoring coincided with historic extremes for climate warming. We identify three recommendations to account for gaps, prioritise new data collection and design future biodiversity monitoring. The biodiversity literature must progress beyond merely discussing bias to instead quantifying representativeness of biodiversity data to better target key research questions, improve interpretation of findings and integrate diverse data streams. By accounting for the range and sampling frequency of biodiversity data along all axes of the global change spectrum, we can strengthen the empirical evidence for the next stage of IPBES global biodiversity assessments and develop more representative projections of ecological change.

Methods

Databases of ecological monitoring

We combined three of the largest biodiversity databases - Living Planet, BioTIME and PREDICTS. The Living Planet database[7] includes time series data of individual species' abundance for vertebrate taxa for the terrestrial, marine and freshwater realms (freshwater realm data were excluded for the purposes of this analysis because of lack of global change driver data for freshwater environments). The BioTIME database[8] is also a compilation of time series but of ecological assemblages for vertebrate, invertebrate and plant taxa across the marine and terrestrial realms. The PREDICTS database[9] includes space-for-time studies testing the effects of land-use change on vertebrates, invertebrates and plants and thus focuses on the terrestrial realm.

Databases of global change

We used the marine and terrestrial global change driver layers compiled by Bowler et al.[3]. The drivers we focused on were human use, climate change, human population density, pollution and invasion potential. The driver data are spatially-explicit, but current data limitations prevent us from extracting temporally-explicit values of the magnitudes of these drivers over time. For details on the individual layers forming the global change data, including their resolutions and temporal coverage, see Table S1 in Bowler et al. 2020. We used the Land Use Harmonisation (LUH) database[42] to extract land cover estimates over a long historic period (from the year 800 to 2014). For the terrestrial realm, we obtained

surface air temperature from the CRU TS v4.05 database[48] and for the marine realm, we extracted sea surface temperature data from the NOAA Extended Reconstructed SST v5 database[49].

Mapping ecological monitoring in global change space

We combined the geographical coordinates of all spatially-explicit monitoring sites in the Living Planet, BioTIME and PREDICTS databases. For each location, we extracted the intensity of 16 global change layers, summarised under five categories (human use, climate change, human population density, pollution, invasion potential) as well as their cumulative magnitudes. The estimates for the magnitudes of each driver were standardised between 0 and 1 to make them comparable. We used a Principal Component Analysis (PCA) to map global change space (similar to trait space[40]) and visualised where in this global change space current monitoring occurs. We extracted driver intensity for one million simulated random locations spanning the globe which represent a scenario in which we monitor the whole world. We used the values from the random sampling as a baseline for quantifying the representation of global change variation in biodiversity data. We visualised marine and terrestrial global change space separately because of known differences in human impact across realms[3].

To statistically compare the intensity of global change drivers around the world and on sites with biodiversity data, we used two general linear models (one for the marine and one for the terrestrial realm) in a Bayesian framework with the following structure:

$$driver\ intensity \sim location,$$

where location represented a four-level categorical variable (Random global sampling, Living Planet, BioTIME or PREDICTS database; in the marine model the PREDICTS database was omitted since it only covers the terrestrial realm). We fitted our model using the package brms[74]. We used the default weakly informative priors. Credible intervals that overlap zero around the effect size (posterior mean) for the intensity of each driver across monitoring sites indicate that existing monitoring is capturing the variation in the intensity of the driver in question. When the credible intervals do not overlap zero,

that suggests that existing monitoring underestimates driver intensity (negative effect size) or overestimates it (positive effect size).

Quantifying mismatches between peak driver intensity and ecological monitoring

To quantify how well biodiversity captured variation in global change over time, we focused on changes in primary forest cover, derived from the LUH database[42] and in temperature, derived from the CRU TS v.4.05 database[48] for the terrestrial realm and from the NOAA Extended Reconstructed SST v5 database[49] for the marine realm. We chose these focal drivers because they explain high amounts of the variation in global change in the terrestrial realm[3] and they have long-enough temporal data to allow us to determine their trajectory and how well it matches with the timing of biodiversity time series. We visualised primary forest cover from the year 800 to 2014 for the location of each terrestrial site in the Living Planet and BioTIME databases and indicated when the monitoring began. We were unable to complete a similar analysis for the marine realm because there are no temporal data for human use drivers like fishing of sufficiently high temporal and spatial resolution. We extracted monthly mean temperature data for the same locations and summarised it as yearly averages. We then compared the slopes of temperature change during the biodiversity monitoring with the slopes of temperature change in the period preceding the monitoring (the two comparison periods were of equal length and always more than five years).

Determining taxonomic representation

To quantify taxonomic representation, we extracted the numbers of known species per taxa from the 2019 edition of the Catalogue of Life (http://www.catalogueoflife.org) and then compared them to the numbers of distinct species recorded in the Living Planet, BioTIME and PREDICTS databases. We quantified taxonomic representation as percentages of species which have at least one record in the respective databases.

Code and data availability

All data and code are publicly available. Population and biodiversity time-series data are freely available in the Living Planet and BioTIME Databases (see references for details on data collection). The Living Planet Database can be accessed on http://www.livingplanetindex.org/data_portal. The BioTIME

Database can be accessed on Zenodo (https://doi.org/10.5281/zenodo.1211105) or through the BioTIME website (http://biotime.st-andrews.ac.uk/). PREDICTS can be downloaded from https://www.predicts.org.uk/pages/outputs.html. The database of ecological monitoring locations we compiled is available on GitHub **insert link** and our analyses can be reproduced by downloading the following GitHub code repository **insert link**.

References

1. Ellis, E. C. *et al.* Used planet: A global history. *Proceedings of the National Academy of Sciences* **110**, 7978–7985 (2013).

2. IPCC. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp. (2014).

3. Bowler, D. E. *et al.* Mapping human pressures on biodiversity across the planet uncovers anthropogenic threat complexes. *People and Nature* 2, 380–394 (2020).

4. IPBES. Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. (2019).

5. Culina, A. *et al.* Navigating the unfolding open data landscape in ecology and evolution. *Nat Ecol Evol* 2, 420–426 (2018).

6. Hampton, S. E. *et al.* Big data and the future of ecology. *Frontiers in Ecology and the Environment* 11, 156–162 (2013).

7. WWF. Living Planet Report - 2018: Aiming Higher. Grooten, M. and Almond,

R.E.A.(Eds). WWF, Gland, Switzerland. (2018).

8. Dornelas, M. *et al.* BioTIME: A database of biodiversity time series for the Anthropocene. *Global Ecology and Biogeography* 27, 760–786 (2018).

9. Hudson, L. N. *et al.* The database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) project. *Ecology and Evolution* 7, 145–188 (2017).

10. GBIF. GBIF Home Page. Available from: https://www.gbif.org [28 April 2021]. (2021).

11. Santini, L., Isaac, N. J. B. & Ficetola, G. F. TetraDENSITY: A database of population density estimates in terrestrial vertebrates. *Global Ecol Biogeogr* 27, 787–791 (2018).

12. McCallen, E. *et al.* Trends in ecology: shifts in ecological research themes over the past four decades. *Front Ecol Environ* 17, 109–116 (2019).

13. Daskalova, G. N. *et al.* Landscape-scale forest loss as a catalyst of population and biodiversity change. *Science* 368, 1341–1347 (2020).

14. Antão, L. H. *et al.* Temperature-related biodiversity change across temperate marine and terrestrial systems. *Nat Ecol Evol* 4, 927–933 (2020).

15. Millette, K. L. *et al.* No consistent effects of humans on animal genetic diversity worldwide. *Ecol Lett* 23, 55–67 (2020).

16. Newbold, T. *et al.* Global effects of land use on local terrestrial biodiversity. *Nature* 520, 45–50 (2015).

17. Vellend, M. *et al.* Global meta-analysis reveals no net change in local-scale plant biodiversity over time. *Proceedings of the National Academy of Sciences* 110, 19456–

19459 (2013).

18. Dornelas, M. *et al.* Assemblage Time Series Reveal Biodiversity Change but Not
    Systematic Loss. *Science* 344, 296–299 (2014).

19. Blowes, S. A. *et al.* The geography of biodiversity change in marine and terrestrial
    assemblages. *Science* 366, 339–345 (2019).

20. Macgregor, C. J., Williams, J. H., Bell, J. R. & Thomas, C. D. Moth biomass increases and
    decreases over 50 years in Britain. *Nat Ecol Evol* 3, 1645–1649 (2019).

21. Betts, M. G. *et al.* Global forest loss disproportionately erodes biodiversity in intact
    landscapes. *Nature* 547, 441–444 (2017).

22. Outhwaite, C. L., Gregory, R. D., Chandler, R. E., Collen, B. & Isaac, N. J. B. Complex
    long-term biodiversity change among invertebrates, bryophytes and lichens. *Nat Ecol
    Evol* 4, 384–392 (2020).

23. Daskalova, G. N., Myers-Smith, I. H. & Godlee, J. L. Rare and common vertebrates span a
    wide spectrum of population trends. *Nat Commun* 11, 4394 (2020).

24. Gonzalez, A. *et al.* Estimating local biodiversity change: a critique of papers claiming no
    net loss of local diversity. *Ecology* 97, 1949–2960 (2016).

25. Amano, T., Lamming, J. D. L. & Sutherland, W. J. Spatial Gaps in Global Biodiversity
    Information and the Role of Citizen Science. *BioScience* 66, 393–400 (2016).

26. Boakes, E. H. *et al.* Distorted Views of Biodiversity: Spatial and Temporal Bias in Species
    Occurrence Data. *PLoS Biol* 8, e1000385 (2010).

27. Meyer, C., Kreft, H., Guralnick, R. & Jetz, W. Global priorities for an effective

information basis of biodiversity distributions. *Nat Commun* 6, 8221 (2015).

28. Shirey, V., Belitz, M. W., Barve, V. & Guralnick, R. A complete inventory of North American butterfly occurrence data: narrowing data gaps, but increasing bias. *Ecography* 44, 537–547 (2021).

29. Hill, S. L. L. *et al. Worldwide impacts of past and projected future land-use change on local species richness and the Biodiversity Intactness Index.* http://biorxiv.org/lookup/doi/10.1101/311787 (2018) doi:10.1101/311787.

30. Schipper, A. M. *et al.* Projecting terrestrial biodiversity intactness with GLOBIO 4. *Glob Change Biol* gcb.14848 (2019) doi:10.1111/gcb.14848.

31. Buckland, S. T. & Johnston, A. Monitoring the biodiversity of regions: Key principles and possible pitfalls. *Biological Conservation* 214, 23–34 (2017).

32. Johnston, A., Moran, N., Musgrove, A., Fink, D. & Baillie, S. R. Estimating species distributions from spatially biased citizen science data. *Ecological Modelling* 422, 108927 (2020).

33. Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R. & Legendre, F. Taxonomic bias in biodiversity data and societal preferences. *Sci Rep* 7, 9132 (2017).

34. Mentges, A., Blowes, S. A., Hodapp, D., Hillebrand, H. & Chase, J. M. Effects of site-selection bias on estimates of biodiversity change. *Conservation Biology* 35, 688–698 (2021).

35. Mihoub, J.-B. *et al.* Setting temporal baselines for biodiversity: the limits of available monitoring data for capturing the full impact of anthropogenic pressures. *Scientific*

Reports 7, 41591 (2017).

36. De Palma, A. *et al.* Challenges With Inferring How Land-Use Affects Terrestrial
    Biodiversity: Study Design, Time, Space and Synthesis. in *Advances in Ecological
    Research* vol. 58 163–199 (Elsevier, 2018).

37. Ferraro, P. J., Sanchirico, J. N. & Smith, M. D. Causal inference in coupled human and
    natural systems. *Proc Natl Acad Sci USA* 116, 5311–5318 (2019).

38. Korell, L., Auge, H., Chase, J. M., Harpole, S. & Knight, T. M. We need more realistic
    climate change experiments for understanding ecosystems of the future. *Glob Change
    Biol* 26, 325–327 (2020).

39. Isbell, F., Tilman, D., Reich, P. B. & Clark, A. T. Deficits of biodiversity and productivity
    linger a century after agricultural abandonment. *Nat Ecol Evol* 3, 1533–1538 (2019).

40. Díaz, S. *et al.* The global spectrum of plant form and function. *Nature* 529, 167–171
    (2016).

41. Osenberg, C. W., Schmitt, R. J., Holbrook, S. J., Abu-Saba, K. E. & Flegal, A. R. Detection
    of Environmental Impacts: Natural Variability, Effect Size, and Power Analysis.
    Ecological Applications 4, 16–30 (1994).

42. Hurtt, G. C. *et al.* Harmonization of land-use scenarios for the period 1500–2100: 600
    years of global gridded annual land-use transitions, wood harvest, and resulting
    secondary lands. *Climatic Change* 109, 117–161 (2011).

43. Klein Goldewijk, K., Beusen, A., Doelman, J. & Stehfest, E. Anthropogenic land use
    estimates for the Holocene – HYDE 3.2. *Earth Syst. Sci. Data* 9, 927–953 (2017).

44. Dakos, V. *et al.* Ecosystem tipping points in an evolving world. *Nat Ecol Evol* 3, 355–362 (2019).

45. Jarzyna, M. A. & Jetz, W. Taxonomic and functional diversity change is scale dependent. *Nature Communications* 9, (2018).

46. Jönsson, G. M., Broad, G. R., Sumner, S. & Isaac, N. J. B. A century of social wasp occupancy trends from natural history collections: spatiotemporal resolutions have little effect on model performance. *Insect Conserv Divers* icad.12494 (2021) doi:10.1111/icad.12494.

47. Jennions, M. D. A survey of the statistical power of research in behavioral ecology and animal behavior. *Behavioral Ecology* 14, 438–445 (2003).

48. Harris, I., Osborn, T. J., Jones, P. & Lister, D. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci Data* 7, 109 (2020).

49. Huang, B. *et al.* Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons. *Journal of Climate* 30, 8179–8205 (2017).

50. Titley, M. A., Snaddon, J. L. & Turner, E. C. Scientific research on animal biodiversity is systematically biased towards vertebrates and temperate regions. *PLoS ONE* 12, e0189577 (2017).

51. Hallmann, C. A. *et al.* More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLOS ONE* 12, e0185809 (2017).

52. Sánchez-Bayo, F. & Wyckhuys, K. A. G. Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation* 232, 8–27 (2019).

53. Dornelas, M. & Daskalova, G. N. Nuanced changes in insect abundance. *Science* 368, 368–369 (2020).

54. Didham, R. K. *et al.* Interpreting insect declines: seven challenges and a way forward. *Insect Conserv Divers* 13, 103–114 (2020).

55. Montgomery, G. A. *et al.* Is the insect apocalypse upon us? How to find out. *Biological Conservation* 108327 (2019) doi:10.1016/j.biocon.2019.108327.

56. Daskalova, G. N., Phillimore, A. B. & Myers-Smith, I. H. Accounting for year effects and sampling error in temporal analyses of invertebrate population and biodiversity change: a comment on Seibold *et al*. 2019. *Insect Conserv Divers* 14, 149–154 (2021).

57. Kissling, W. D. *et al.* Towards global data products of Essential Biodiversity Variables on species traits. *Nat Ecol Evol* 2, 1531–1540 (2018).

58. Jetz, W. *et al.* Essential biodiversity variables for mapping and monitoring species populations. *Nat Ecol Evol* 3, 539–551 (2019).

59. Olson, D. M. & Dinerstein, E. The Global 200: Priority Ecoregions for Global Conservation. *Annals of the Missouri Botanical Garden* 89, 199–224 (2002).

60. Rocha-Ortega, M., Rodriguez, P. & Córdoba-Aguilar, A. Geographical, temporal and taxonomic biases in insect GBIF data on biodiversity and extinction. *Ecol Entomol* een.13027 (2021) doi:10.1111/een.13027.

61. Collen, B., Ram, M., Zamin, T. & McRae, L. The Tropical Biodiversity Data Gap: Addressing Disparity in Global Monitoring. *Tropical Conservation Science* 1, 75–88 (2008).

62. Cameron, E. K. *et al.* Global gaps in soil biodiversity data. *Nat Ecol Evol* 2, 1042–1043 (2018).

63. Pereira, H. M. *et al.* Essential Biodiversity Variables. *Science* 339, 277–278 (2013).

64. Geijzendorffer, I. R. *et al.* Bridging the gap between biodiversity data and policy reporting needs: An Essential Biodiversity Variables perspective. *J Appl Ecol* 53, 1341–1350 (2016).

65. Wetzel, F. T. *et al.* Unlocking biodiversity data: Prioritization and filling the gaps in biodiversity observation data in Europe. *Biological Conservation* 221, 78–85 (2018).

66. Isaac, N. J. B. & Cowlishaw, G. How species respond to multiple extinction threats. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271, 1135–1141 (2004).

67. Barnagaud, J.-Y. *et al.* Relating Habitat and Climatic Niches in Birds. *PLoS ONE* 7, e32819 (2012).

68. Frishkoff, L. O. *et al.* Climate change and habitat conversion favour the same species. *Ecology Letters* 19, 1081–1090 (2016).

69. Mace, G. M. *et al.* Aiming higher to bend the curve of biodiversity loss. *Nat Sustain* 1, 448–451 (2018).

70. Leclère, D. *et al.* Bending the curve of terrestrial biodiversity needs an integrated strategy. *Nature* (2020) doi:10.1038/s41586-020-2705-y.

71. Pereira, L. M. *et al.* Developing multiscale and integrative nature–people scenarios using the Nature Futures Framework. *People and Nature* 2, 1172–1195 (2020).

72. Harris, S. J. *et al.* The Breeding Bird Survey 2016. BTO Research Report 700. British Trust for Ornithology, Thetford, UK. https://www.bto.org/volunteer-surveys/bbs/bbs-publications/bbs-reports [Accessed 13/9/2018]. (2017).

73. Zurell, D., Elith, J. & Schröder, B. Predicting to new environments: tools for visualizing model behaviour and impacts on mapped distributions: Predicting to new environments. *Diversity and Distributions* 18, 628–634 (2012).

74. Bürkner, P.-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* 80, (2017).

Acknowledgements

Competing interests: The authors declare no conflicts of interest.

Author contributions: G.N.D. … conceptualized the study. G.N.D. wrote the first draft, conducted analyses and visualized results with input from all co-authors. I.M.S. was the primary supervisor, M.A.D. the co-supervisor. D.B. compiled the global change driver layers (Bowler et al. 2020). All authors contributed to editing the manuscript.