Title: Global Threats to Biodiversity

Authors: Lucas Joppa[*1], Brian O'Connor[2], Piero Visconti[2], Cathy Smith[2], Jonas Geldmann[3,4], Michael Hoffmann[2,5], James E.M. Watson[6,7], Stuart H. M. Butchart[8,4], Malika Virah-Sawmy[9], Benjamin S. Halpern[10,11,12], Sadia E. Ahmed[1], Andrew Balmford[4], William J. Sutherland[4], Mike Harfoot[2], Craig Hilton-Taylor[13], Wendy Foden[14], Enrico Di Minin[15,16], Shyama Pagad[17], Piero Genovesi[18], Jon Hutton[2], Neil D. Burgess[2,3,4]

[1]Microsoft Research, One Microsoft Way, Redmond, WA 98052

[2]UNEP World Conservation Monitoring Centre, 219 Huntingdon Road, Cambridge, CB2 ODL, UK

[3]Center for Macroecology, Evolution and Climate, Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen E, Denmark

[4] Conservation Science Group, Department of Zoology, University of Cambridge, Downing, Cambridge CB2 3EJ, UK

[5] IUCN Species Survival Commission, IUCN, 28 Rue Mauverney, 1196 Gland, Switzerland

[6] School of Geography, Planning and Environmental Management, University of Queensland, Brisbane, QLD, Australia, 4072

[7] Wildlife Conservation Society, Global Conservation Program, Bronx, NY, USA, 10460

[8]BirdLife International, The David Attenborough Building, Pembroke Street, Cambridge CB2 3QZ, UK

[9]Luc Hoffmann Institute, WWF International, Avenue du Mont-Blanc, 1196 Gland, Switzerland

[10]Bren School of Environmental Science and Management, University of California, Santa Barbara, CA 93106, USA

[11]NCEAS, 735 State St., Santa Barabara, CA 93101, USA

[12]Imperial College London, Silwood Park, Buckhurst Rd., Ascot, UK

[13]IUCN Global Species Programme, The David Attenborough Building, Pembroke Street, Cambridge, CB2 3QZ, UK

[14]Department of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa and Chair IUCN Species Survival Commission, Climate Change Specialist Group

[15]Finnish Centre of Excellence in Metapopulation Research, Department of Biosciences, University of Helsinki, Helsinki, FI-00014, Finland

[16]School of Life Sciences, University of Kwa-Zulu-Natal, Private Bag X54001, Durban 4000, South Africa

[17]The IUCN Species Survival Commission Invasive Species Specialist Group, University of Auckland, Tamaki Campus, Auckland, New Zealand

[18]ISPRA, Institute for Environmental Protection and Research and Chair IUCN SSC Invasive Species Specialist Group, Rome, Italy

Abstract: Reducing rates of biodiversity loss and achieving environmental goals requires an understanding of what is threatening biodiversity, where and how fast the threats are changing in type and intensity, and appropriate actions needed to avert them. One might expect that the Information Age – typified by a deluge of data resulting from massive and widespread collection, digitization and dissemination of information – would have revolutionized our understanding of global threats to biodiversity. We examine the extent to which this is true, identify major data gaps for understanding threats to biodiversity, and suggest mechanisms for closing them. These recommendations include innovative partnerships with data providers of all kinds, ensuring relevant data sources are openly available and accessible, and a considerable investment of funding into scalable data gathering initiatives.

Text: The diversity of life on earth – which provides vital services to humanity (*1*) – stems from the difference between rates of evolutionary diversification and extinction. We live in the Anthropocene, an age where human activities have shifted the balance towards extinction (*2*). Species extinction rates are estimated at ~1,000 times higher than the 'background' rate (*3*) but could increase to 10,000 times higher should those species currently threatened with extinction succumb to the pressures they face (*4*). Reversing these trends in species extinctions is a focus of the Convention on Biological Diversity's 2020 Strategic Plan for Biodiversity and its 20 Aichi Targets, and is explicitly incorporated in the United Nation's (UN) 2030 Agenda for Sustainable Development and its 17 Sustainable Development Goals (SDGs).

Reducing rates of biodiversity loss and achieving environmental goals requires understanding of what is threatening biodiversity, where risks occur, how fast the threats are changing in type and intensity, and what the most appropriate actions are needed to avert them (*5*). At face value this seems straightforward. The Anthropocene overlaps with the Information Age, typified by a deluge of data resulting from massive and widespread collection, digitization and dissemination of information. The combination of crowd-sourced data, large-scale ground-based monitoring schemes, and satellite earth observation missions is capable of providing unprecedented insight into global threats to biodiversity, and how human interventions are altering those threats [eg. (*6*)]. Ensuring that the current era of 'big data' does not overlook the sustainability agenda has led to a UN report's specific policy recommendations for mobilizing the data revolution for sustainable development and environmental protection (*7*).

We examine the extent to which the Information Age is generating data on threats to biodiversity. We show that the UN report's conclusion (*7*) – "too much that needs to be known remains unknown" – applies to threats to biodiversity. As a fundamental contribution to meeting the global policy community's aspirations for data relevant to the SDGs, we identify the major data gaps for biodiversity threats, and suggest mechanisms for closing them.

Data Deluge or Data Drought?

We used a threat classification scheme (*8*; Figure 1), that while not without shortcomings (*9, 10*), has been widely deployed for tens of thousands of conservation assessments for species, sites, and projects. By 'threat', we mean 'The proximate human activities or processes that have caused, are causing, or may cause the destruction, degradation, and/or impairment of biodiversity targets' (*8*). Under this definition, determining the impact of a threat on a species or ecosystem is a separate process often included in a conservation assessment. We followed a structured data collection procedure and associated each dataset with one or more classes of threat (see SOM for details). We omit three threat classes from our analysis: two (Geological Events; Other Options) are not exclusively anthropogenic and one (Climate Change and Severe Weather) due to comprehensive treatment by the Fifth Assessment Report for the Intergovernmental Panel on Climate Change. We restricted our search to spatial datasets with a global extent. We assume the datasets identified by this initial search will grow as additional datasets and metadata become known or are created.

We identified 290 unique datasets (Table S1, S2) across nine threat classes from data sources ranging from remote sensing via satellites to citizen science initiatives (Figure S1). Six data providers account for over a fifth of the entire catalogue of datasets. That so many datasets exist, from a variety of sources, shows that the Information Age is generating geospatial data that may be of use to biodiversity conservation policy and action. However, this apparent data deluge is misleading: overall, our analysis reveals how little is actually available, at the global level, about the spatial and temporal distribution of anthropogenic threats to biodiversity.

In order to assess whether data on different threats were available in proportion to their importance for biodiversity, we used threat information (for threatened taxa which have been comprehensively assessed) from the International Union for Conservation of Nature's Red List of Threatened Species (IUCN Red List), the repository of information on the global extinction risk of species. We find that the frequency of threats to marine or terrestrial and inland water species on the Red List is disproportionate to the availability of datasets on those threats (Figure 1, Table S2). Biological Resource Use (including direct and indirect impacts of hunting, fishing, and logging) is one of the most common threats to species, yet accounts for just 5% of threat datasets.

To assess how much threat information is actionable and available we examined our datasets with respect to five desirable data attributes (Table 1; see SOM for details). We note that determining accurate attribute values was often difficult due to lack of formal metadata, creating uncertainty in the absolute number of datasets which might satisfy all criteria. Regardless, our filters remove all but fourteen datasets (5%) and not all threat classes are represented (Figure 2, SOM and Table S1 for details). Further, datasets which do comply are often applicable to only a few taxa or habitats. We provide dataset attributes as filters in Table S1.

Business Models for Data Acquisition

Our findings highlight major data gaps in understanding global threats to biodiversity. The conservation community should aspire to at least one 'gold standard' dataset - that meets at a minimum all five desirable attributes in Table 1 and is applicable to as many taxa as possible - for each class and subclass of threat. Achieving this will require working with appropriate data providers to develop business models for data acquisition that leverage new, longer-term funding mechanisms and partnerships with government and the private sector.

Partnerships with Data Owners and Creators

In certain instances, the data required for effective conservation policy already exists, but are not accessible (e.g. due to access cost, commercial considerations, or intellectual property arrangements) to organizations or agencies mandated to conserve biodiversity. Sometimes these data result from taxpayer-funded initiatives that, thanks to advocacy by the scientific community, can result in major success stories (*6*). In 2008 NASA announced the free public release of the Landsat image archive, dating back to 1978. Landsat imagery subsequently empowered the scientific community to begin studies of land cover change at an actionable resolution. Since then the European Space Agency opened the Sentinel Scientific Data hub, a free and open access data portal for imagery from the Copernicus Sentinel missions, and the French Space Agency declared five years or older SPOT satellite data, free of charge to noncommercial users.

Data held by the private sector also have the potential to fill major gaps. Gaining access will require partnerships that respect the intellectual property of companies and the right of conservation organizations to use these data for conservation actions. One such agreement between the UNEP World Conservation Monitoring Center and the IHS Company enables detailed and comprehensive data on oil and gas activity worldwide to be used for biodiversity assessments. More broadly, the conservation community should emulate the UN's *Data for Climate Action* initiative, which is laying the groundwork for working with the private sector to access big data – with options ranging from companies making data freely available to specific arrangements for scientists to access data within the company's protected network.

Funding Mechanisms

Successfully delivering the UN's 2030 Agenda for Sustainable Development will require exceptional financial support. In July 2015, the UN's Third International Conference for Financing for Development produced a comprehensive framework – the Addis Ababa Action Agenda (AAAA). The AAAA specifies >100 measures for how to finance the sustainable development agenda, and explicitly recognizes the need to fund 'science, technology, innovation, and capacity building', as well as 'data, monitoring, and follow-up' (*11*). In a separate note, the UN Secretariat details linkages between each measure and the individual SDGs (*12*).

The AAAA 'encourage(s) the mobilization of financial resources from all sources and at all levels to conserve and sustainably use biodiversity and ecosystems'. This is an important recognition of the need to finance the achievement of SDG 15 (the most relevant to halting the loss of biodiversity). Yet a lack of specific mention of the need to fund the data required to achieve that goal is a critical omission.

Enabling the Data Pipeline

We note that for many threat classes the creation of a 'gold standard' dataset need not start from scratch. Existing datasets and data pipelines, if provided with appropriate resources or mandates, can be scaled up. We highlight this potential with two conservation issues where data scarcity on threats is a major obstacle for achieving SDG 15.

Specific Threat Class: Invasive and other problematic species

Invasive alien species homogenize global biodiversity and are a significant threat to native species, particularly those endemic to islands and specific ecosystems. National and regional policy mechanisms (e.g. United States' *National Invasive Species Management Plan*, European Union *Regulation 1143*) are in place to prevent, control, and minimize the impact of alien species. Effective policy must be empowered with comprehensive data on which species are where, and pathways by which they move (as the European Union's legal framework explicitly requires). These data allow implementation agencies to monitor transmission routes, prevent invasive species entry or departure, and respond rapidly to early detections. The Threatened Island Biodiversity Database and the IUCN's Global Invasive Species Database are backed by international institutions and networks of experts and, if appropriately resourced, are capable of scaling up substantially to meet the five key data attributes we identify above. Such a database would significantly enhance the capabilities of agencies to plan for, prevent, and respond to the arrival of invasive species or manage or eradicate those already established.

Data Gaps on Threats from Land Use and Cover Change

Habitat loss is a leading cause of biodiversity decline, and most countries have local, regional, and national legislation protecting natural landscapes. Yet globally we still do not have a standard land use and cover change assessment tool for biodiversity conservation end users. New and standardized land cover change detection approaches for the 2000-2010 interval are emerging, at both high (30m) (e.g. *13*) and moderate (300m) resolution (*14*). Although these products have promise, it is still impossible to obtain a global and standardized overview of how natural landscapes are changing on a time scale that allows appropriate conservation action. Changing this clearly requires breaking the practice of repeatedly modifying remote-sensing algorithms – interesting for the field itself but exasperating for end-users – and instead agreeing to a series of global maps comparable through time and space.

Discussion

To be useful, threat datasets must be integrated with conservation assessment processes. The IUCN Red List compiles input from >10,000 species experts into easily and freely available conservation assessments for nearly 80,000 species that influence international and national policy mechanisms. Connecting such efforts to 'gold standard' datasets for each major class of threat will help bring actionable insights into what conservation actions are needed, and where, for the most imperiled species and populations on the planet. Our metadatabase is intended as an initial and growing information portal to datasets relevant to biodiversity threats. Over time, we recommend the inclusion of the significantly more numerous and available regional datasets (even if they do not meet the dataset attributes identified here), and their integration to create more globally representative information.

Leveraging the technology of the Information Age to counter biodiversity loss – a defining feature of the Anthropocene, can help make a fundamental contribution towards the success of the UN's 2030 Agenda for Sustainable Development. Successfully doing so will require new and innovative partnerships with data providers of all kinds, making sure relevant data sources are made openly available and accessible, and a considerable investment of funding into scalable data gathering initiatives.

References and Notes

1. R. S. De Groot, A. W. Matthew, R. M. J. Boumans. A typology for the classification, description and valuation of ecosystem functions, goods and services. *Ecological Economics* **41**, 393-408 (2002).
2. W. Steffen, P. J. Crutzen, J. R. McNeill. The Anthropocene: are humans now overwhelming the great forces of nature. *AMBIO: A Journal of the Human Environment* **36**, 614-621 (2007).
3. J. M. De Vos, *et al*. Estimating the normal background rate of species extinction. *Conservation Biology* **29**, 452-462 (2015).
4. S. L. Pimm, *et al*. The biodiversity of species and their rates of extinction, distribution, and protection. *Science* **344**, 1246752 (2014).
5. J. Geldmann, L. Joppa, N.D. Burgess. Mapping Change in Human Pressure Globally on Land and within Protected Areas. *Conservation Biology* **28**, 1604-1616 (2014).
6. W. Turner. Sensing Biodiversity. *Science* **346**, 301-302 (2014).
7. Independent Expert Advisory Group. A World that Counts: Mobilising the Data Revolution for Sustainable Development. http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf (2015).
8. N. Salafsky *et al*. A standard lexicon for biodiversity conservation: Unified classifications of threats and actions. *Conservation Biology* **22**, 897-911 (2008).
9. A. Balmford, *et al*. Capturing the many dimensions of threat: comment on Salafsky et al. *Conservation Biology* **23**, 482-487 (2009).
10. N. Salafsky, *et al*. Pragmatism and practice in classifying threats: a reply to Balmford et al. *Conservation Biology* **23**, 488-493 (2009).
11. Addis Ababa Action Agenda. Third International Conference on Financing for Development. http://www.un.org/ga/search/view_doc.asp?symbol=A/CONF.227/L.1 (Accessed 1/20/2016).
12. Informal Note by the Secretariat. Linkages between the means of implementation of the Sustainable Development Goals and the Addis Ababa Action Agenda. http://www.un.org/esa/ffd/ffd3/wp-content/uploads/sites/2/2015/07/SDG-MoIs_AAAA.pdf (Accessed 1/20/2016).
13. P. Gong, *et al.* Finer resolution observation and monitoring of global land cover: first mapping results 444 with Landsat TM and ETM+ data. *International Journal of Remote Sensing*, **34**, 2607-2654. (2013)
14. ESA Climate Change Initiative (CCI), 2016: http://www.esa-landcover-cci.org/?q=node/1. Accessed 13th January 2016

Acknowledgements

This paper was supported by funding from the Cambridge Conservation Initiative (CCI) project "Toward an integrated approach to the mapping and mitigating of global pressures on biodiversity " involving the following CCI partners: the UNEP World Conservation Monitoring Centre (UNEP WCMC), BirdLife International , International Union for Conservation of Nature (IUCN) and the Department of Zoology, University of Cambridge. The non-CCI partners in the project are: Microsoft Research, University of Copenhagen, University of California (Santa Barbara), the Wildlife Conservation Society, University of Queensland and James Cook University. LNJ conceived the study and wrote the manuscript with contributions from all co-authors. PV, BOC and LNJ designed the analyses, CS, BOC, and PV collected the data. All authors and their affiliated institutions, as well as Ben Phalan (University of Cambridge), Graeme Buchanan (Royal Society for the Protection of Birds), Maria Dias (BirdLife International), Victoria Veach (University of Helsinki), and Steffen Fritz (International Institute for Applied Systems Analysis), were involved in reviewing the datasets included here. BOC, PV and LJ analyzed the data,

All data analyzed in this paper is presented in Tables S1 and S2.

Figures

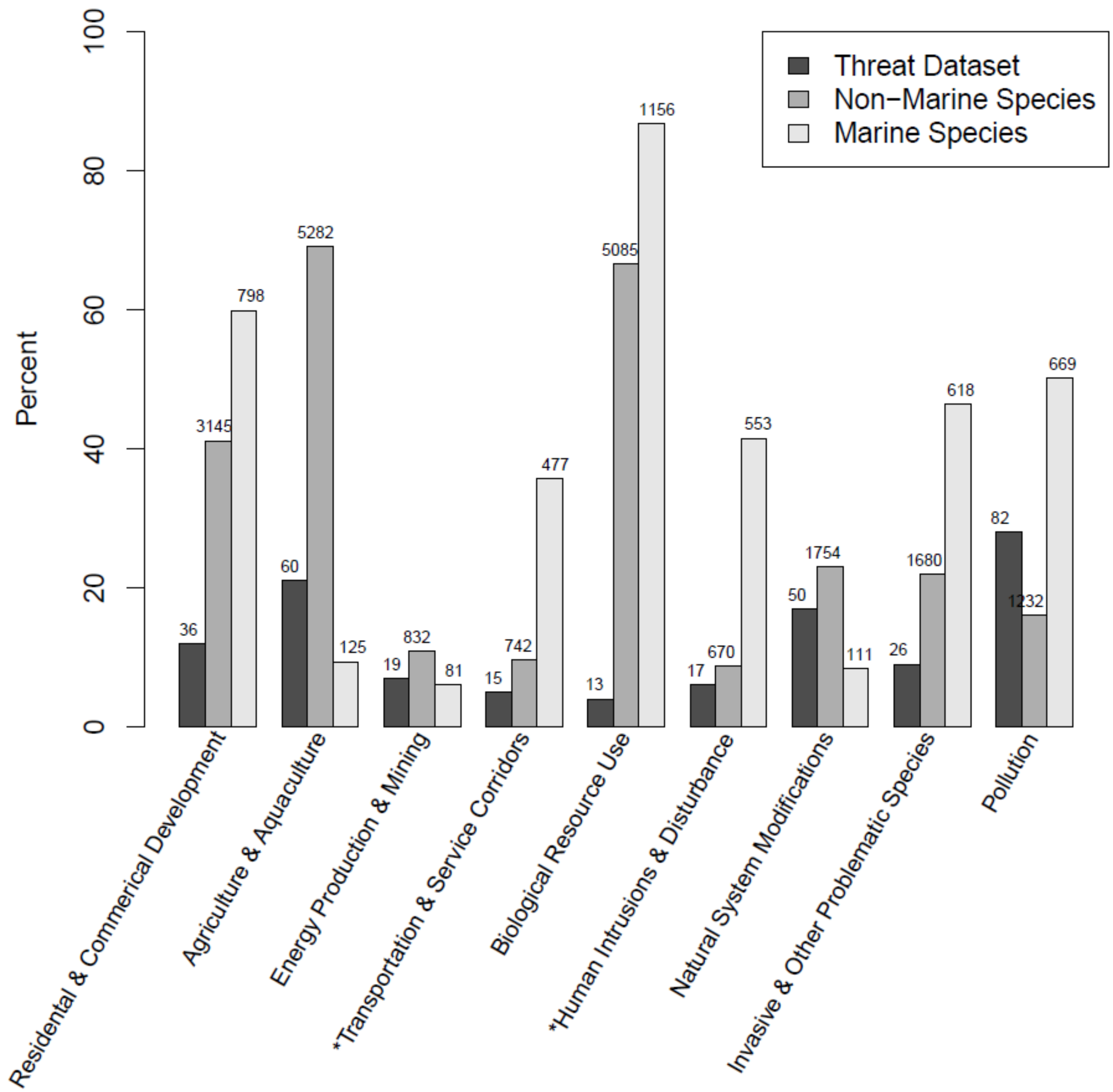

Figure 1: The percent of all threat datasets (black) that relate to each threat class and the percent of threatened terrestrial and inland water (grey) and marine (white) species on the IUCN Red List impacted by each threat class. Number of datasets or species in each class is indicated above each bar. Threat classes not covered by a single dataset are denoted by an * in the figure labels. Details on species included can be found in Table S2.

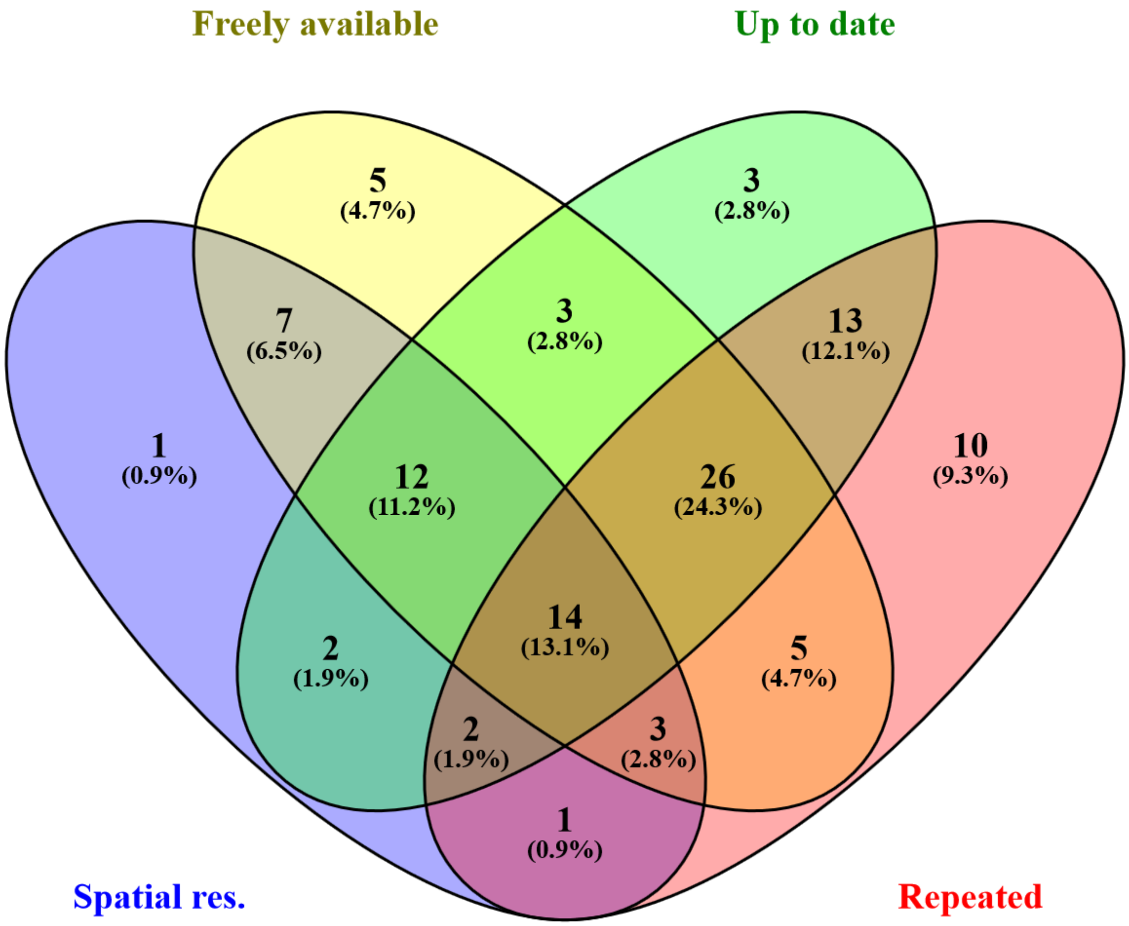

Figure 2: The number of datasets that meet each of four desirable dataset attributes outlined in Table 1 as well as being global in coverage and representing either models assessed for accuracy or empirical observations. Numbers in each intersection represent the number of datasets that meet those constraints. See Table S1 for a full list of datasets and their quality attributes.

Table 1: The rational for each of the five datasets attributes considered key for use in biodiversity threat assessments. These results are available as filters in Table S1.

Box 1 Desirable Dataset Attributes		
Attribute	**Definition and justification**	**Datasets Available**
Freely Available	Dataset is freely available (at least for non-commercial use). We note that being freely available is a necessary but insufficient criterion as depending on the skills and technical capacity of experts a free dataset may still be impossible to access.	153 (53%) datasets are freely accessible for non-commercial use.
Spatial Resolution	Dataset has a spatial resolution of ≤10kmx10km. Approximately one-quarter (23%) of species on the IUCN Red List have ranges smaller than 1,000 km^2. A spatial dataset of a threat with a resolution of 10kmx10km would cover such a species' range with no more than 10 grid cells – a minimum desirable resolution for most analyses.	124 (43%) datasets are either of vector format or at a gridded resolution greater than 10km x 10km (results are 171 (59%) when using a gridded resolution of 100km x 100km).
Up to Date	Dataset has been produced within the last decade: sufficiently recent to inform current and future policy.	149 (51%) datasets have been created or updated since 2006 with 195 (67%) since 2001.
Repeated	Dataset is available for at least two time-points. Monitoring and reporting require an understanding of changes over time. Such trends are fundamental for many conservation assessment criteria and without them it is impossible to understand the impact of regulatory policies.	163 (56%) datasets have been repeated at least once since they were created.
Assessed for Accuracy	Datasets that are modelled have been assessed for accuracy through a validation exercise, so that they can be used with confidence. Conservation assessments are generally subject to independent review, and the datasets used to create them must themselves be of sufficient scientific rigour.	112 (39%) datasets are likely either direct observations or model outputs that have been assessed for accuracy at a global scale. (172 (59%) when including regional validation).