Title

Biodiversity monitoring for a just planetary future

Journal

Science, 383(6678)

ISSN

0036-8075

Authors

Chapman, Melissa
Goldstein, Benjamin R
Schell, Christopher J
et al.

Publication Date

2024-01-05

DOI

10.1126/science.adh8874

POLICY FORUM

BIODIVERSITY

Biodiversity monitoring for a just planetary future

Data that influence policy and major investment decisions risk entrenching social and political inequities

By **Melissa Chapman**[1,2], **Benjamin R. Goldstein**[2], **Christopher J. Schell**[2], **Justin S. Brashares**[2], **Neil H. Carter**[3], **Diego Ellis-Soto**[4], **Hilary Oliva Faxon**[5], **Jenny E. Goldstein**[6], **Benjamin S. Halpern**[1], **Joycelyn Longdon**[7], **Kari E. A. Norman**[8], **Dara O'Rourke**[2], **Caleb Scoville**[9], **Lily Xu**[10], **Carl Boettiger**[2]

Ecologists and conservation scientists have long acknowledged that biodiversity data reflect legacies of social inequity (see the figure). Although the ramifications of these disparities were easy to dismiss when the application of large-scale biodiversity data was limited to academic biogeography and theoretical conservation prioritizations, the stakes have changed. Biodiversity data carry more influence than ever before (*1*), guiding the implementation of massive multilateral commitments and global investments that will affect nature and people for decades to come—from informing priorities for more than doubling the global area under conservation management to creating international biodiversity offset markets. We examine two contentious questions that arise as we consider the disparities in biodiversity data and their consequences in the wake of contemporary biodiversity policy: Are the best available data really a suitable standard? Can more data and better statistical methods ensure that inequities aren't entrenched when implementing data-driven solutions?

With hundreds of billions of dollars being invested in conserving, restoring, and sustainably managing ecosystems in the wake of the post-2020 Kunming-Montreal Global Biodiversity Framework (GBF) (*2*), an understanding of the ways in which data biases propagate through decision-making is critical for the effective creation and communication of data-driven solutions to global biodiversity loss. The path forward will require more than technocratic fixes. Interdisciplinary collaboration and inclusive, bottom-up processes will be critical for leveraging past, present, and future biodiversity data in a way that aligns with the equity goals of global biodiversity policy.

A GLIMPSE INTO GLOBAL BIODIVERSITY DATA

The systems that generate biodiversity data are complex, uneven, and ultimately human. Species observations reflect human processes across space and time: from the decadal impacts of colonialism to the weekly sway of work schedules in modern society, from geopolitical strife to neighborhood-scale disparities.

Take, for example, the Global Biodiversity Information Facility (GBIF). GBIF is a data repository that synthesizes billions of species observations across the globe (see the figure) and specifically aims to provide global-scale biodiversity data to underpin policy and inform decisions from invasive species management to priorities for conservation investment.

Even at first glance, GBIF data do not reflect latitudinal gradients of biodiversity, but more closely trace macroeconomic patterns (see the figure). These data disparities are unsurprising to most ecologists and, like the overrepresentation of population centers, roads, and protected areas in global species observations (*3*), are increasingly adjusted for, even if imperfectly, within existing modeling frameworks (*4*). But digging deeper into these data, strikingly uneven patterns of data availability reveal signatures of armed conflict (see the figure) (*5*), the legacy effects of racist policies (see the figure) (*6*), and changes in political regimes (*5*).

Although descriptions of how biodiversity data disparities trace social and political inequity are notable (*3*, *5*, *6*), they rarely provide the insights necessary to causally attribute mechanisms of those disparities. Human patterns captured by biodiversity data undoubtedly include observational biases but also reflect a reality of the Anthropocene: People—and our politics, economies, and histories—are major drivers of ecosystem composition. European colonial history is still detectable in the true distribution of alien floras worldwide (*7*). Armed conflict affects underlying ecological processes in a variety of complex ways (*8*). The legacy of residential segregation has influenced greenspace and tree cover across neighborhoods, which in turn affect habitat for and distribution of urban wildlife (*9*). To add complexity, environments most degraded by extractive infrastructure are often the most monitored—extractive infrastructures are often also (biodiversity) knowledge infrastructures. For example, the Sacramento–San Joaquin Delta in California is subject to a tremendous amount of ecological monitoring, established as a political compromise to assess the effects of building California's complex water infrastructure (*10*).

But will disparities in biodiversity data really translate to ineffective and inequitable decisions for nature and people? And if so, what can be done about it given the urgency of the biodiversity crisis and the immediacy of informing global policy implementation?

FROM DATA TO DECISIONS

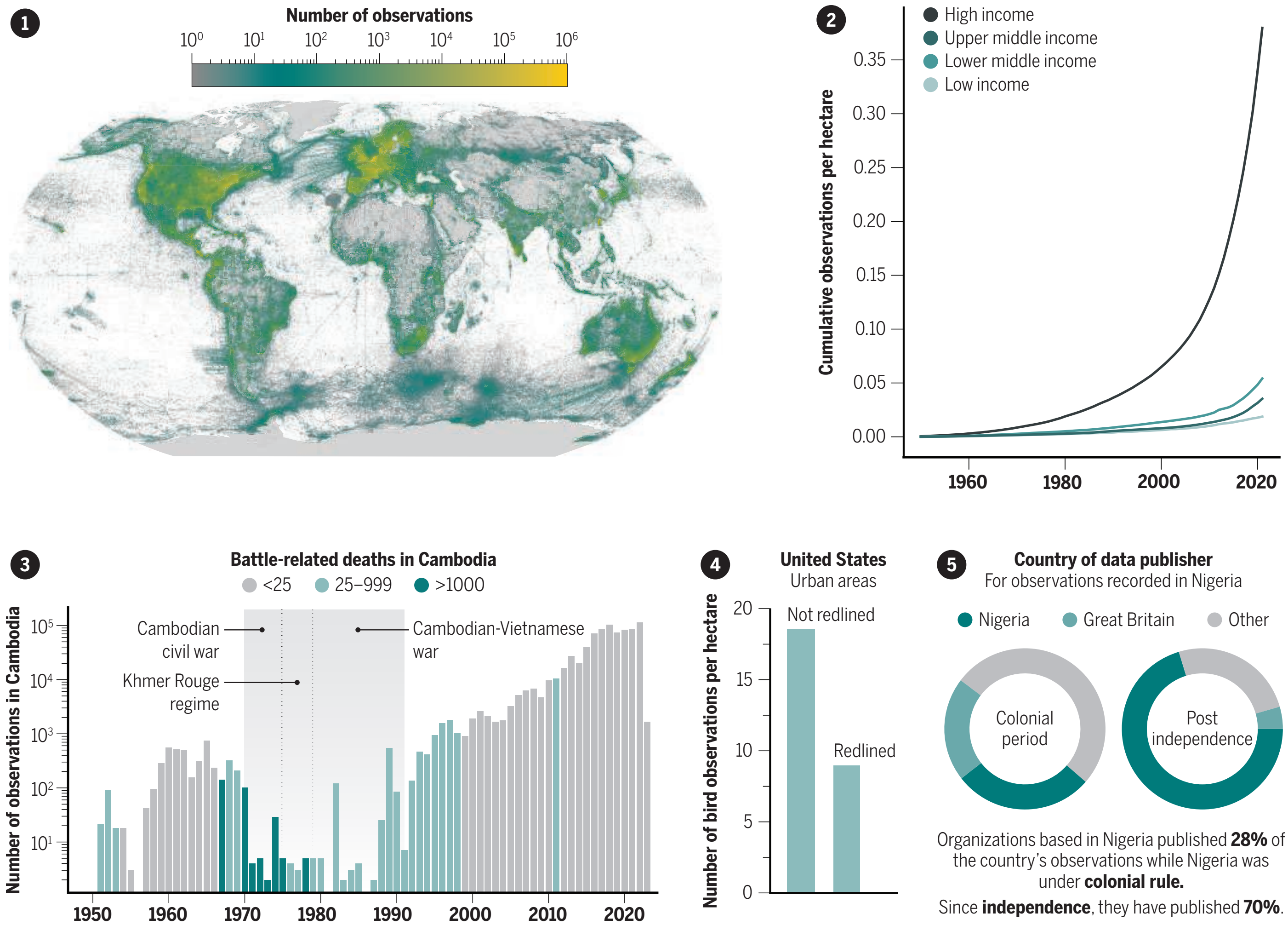
Although the impact of data disparities on decisions is central to discussions on data governance throughout society—from policing to finance to health care—the environmental domain has skirted many of these critiques under the guise that its data reflect and affect the natural world, not people, politics, and histories. The ecological community agrees that data disparities exist but has yet to assess how those disparities propagate through derived ecosystem indicators and policy and management decisions.

There are several ways in which data disparities might be reflected in science-informed decisions in the context of global biodiversity targets. For example, extensive data collected within government-managed parks compared to community-managed and Indigenous lands (*11*) might lead to systematic underestimates of biodiversity

[1]National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara, Santa Barbara, CA, USA. [2]Department of Environment Science, Policy, and Management, University of California, Berkeley, Berkeley, CA, USA. [3]School for Environment and Sustainability, University of Michigan, Ann Arbor, MI, USA. [4]Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA. [5]Department of Society and Conservation, University of Montana, Missoula, MT, USA. [6]Department of Global Development, Cornell University, Ithaca, NY, USA. [7]Department of Computer Science, University of Cambridge, Cambridge, UK. [8]Département de sciences biologiques, Universite de Montreal, Montreal, Canada. [9]Department of Sociology, Tufts University, Medford, MA, USA.[10]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. Email: mchapman@nceas.ucsb.edu

Biodiversity data reflect legacies of social inequity

(**1**) The >2.6 billion species observations in the Global Biodiversity Information Facility (GBIF) database are disproportionately from high-income countries. (**2**) These macroeconomic disparities in data density have become more pronounced through time. (**3**) There are fewer species observations in places and times of conflict (*5*). For example, biodiversity observations notably declined during the Cambodian Civil War, which began in 1970, and especially the Cambodian genocide (1975–1979), and remained low during the following decade of armed conflict. (**4**) In the United States, biodiversity observations unearth the legacy of the effects of racial and ethnic discrimination in housing policy in the 1930s ("redlining") (*6*). Neighborhoods that were redlined, or deemed "hazardous," have approximately half the bird observation density today of those neighborhoods that were deemed "safe" investments (*6*). (**5**) Human histories are reflected not only in where and when data are recorded but also who collects, published, and owns data. In Nigeria, shifts can be seen in the country of data-publishing organizations following independence from Great Britain in 1960 (*5*). See https://github.com/milliechapman/humanDim-gbif for data, code, and further information about each panel.



presence in the latter, misguiding ongoing dialogues about the impact of different land tenure, property rights, and management regimes on biodiversity outcomes. Invasive species might be detected earlier in more intensely surveyed areas, driving investment toward removal and restoration in areas most thoroughly monitored.

Without directly addressing and correcting for social and political disparities in data, the conservation community will likely fall into the same traps that other domains do—entrenching the inequities of the past and present in future decision-making through data. Quantitatively and qualitatively assessing data-derived decision biases and the typologies of their impacts on people and communities is an important first step to effectively mitigating the potential negative impacts of these disparities.

MORE BIODIVERSITY DATA AND BETTER MODELS MIGHT NOT SOLVE PROBLEMS

The past decade has marked a shift away from labor- and resource-intensive specimen collection and field surveys and toward a new generation of decentralized monitoring tools. Participatory science programs, artificial intelligence–supported sensors, and eDNA promise to substantially increase the number of records per research dollar and person-hour. More automated digitization of natural history collections around the world is increasing the capacity to understand long-term changes in ecosystems.

These technologies and their resultant data streams will undoubtedly provide critical information to fill gaps in our knowledge about global biodiversity and inform more robust global policy strategies. But as biodiversity data become easier and cheaper to collect, will sampling become widespread enough that biases are an artifact of the past, buried under the massive amounts of new information?

Although new monitoring technologies continue collecting information about global biodiversity and its degradation at finer

GRAPHIC: M. HERSHER/*SCIENCE*

resolutions and with a broader scope, this increasing amount of information has yet to yield more representative data coverage of biodiversity distributions. Instead, new waves of biodiversity data have entrenched the long-known overrepresentation of certain regions, taxa, and time periods in global biodiversity data repositories (*12*).

Regardless of the volume or velocity of data collection, where, when, how, and by whom species are observed will always be shaped by social, political, and economic processes (*13*).

Collecting perfectly uniform global biodiversity data isn't the only possible solution for addressing the gaps and disparities in existing data. Ecologists and statisticians have worked extensively on methods for bias correction of existing biodiversity data to infer how species distributions and populations vary in time and space despite imperfect data (*4*).

Nonrandom sampling effort can be addressed in two ways: One is by assuming that unobserved variation in sampling (e.g., geopolitical conflict–associated differences in sampling effort) is not confounded with the natural process of interest (e.g., biodiversity distributions and their change); another is by "correcting" the bias in the process of interest with data preprocessing or model-based inference. In the case of social drivers of biodiversity sampling at continental and global scales, neither of these technical fixes is likely adequate to remove biases.

The first option—assuming that the sampling process in question is not related to the natural process of interest—is baseless in most cases. As outlined above, the drivers of data collection are often deeply intertwined with the natural processes that scientists often seek to assess.

The second option—correcting for biases—is only as effective as the capacity of quantifiable variables to explain the biases in the data. In ecology, bias correction tends to focus on bioclimatic conditions, latitudinal disparities, and simple accessibility metrics (e.g., population density, proximity to roads) (*4*), meaning that the other social infrastructures underlying these data likely remain reflected in ecological insights (e.g., species distribution models, metrics of community change) and downstream decisions (e.g., conservation priorities). Archiving and digitizing human societies' darkest hours—from war to colonialism to systemic racism—may allow researchers to start to disentangle the past, present, and future signatures of humans on both biodiversity and the data capturing its distribution and change. Characterizing unintended sociopolitical patterns in data is an important step toward developing analytical methodologies that more accurately reflect true biodiversity patterns.

Although careful statistical models can help identify and control for data disparities that can be quantified, they are not a panacea. Quantitatively correcting socially determined bias across spatial and temporal scales from the top down would require a near-complete census of these multiscale and interacting biases—an infeasible trap. Even when such quantification reveals statistically clear associations, conducting inference on the multidirectional and interacting causal mechanisms that link social infrastructure, monitoring, and biodiversity is impossible without a deeper understanding of those systems than global synthesis data can provide.

"Moving beyond quick technical fixes will require connecting strategically to community-based partners..."

Further, some human drivers of observational (and ecological) processes are not digitizable or easily reduced into quantitative metrics. Although it might be possible to investigate the impact of past residential segregation policies in the United States because there is geospatial information on its history, dimensions that cannot be reduced to polygons on a map, such as human preferences, scientific funding patterns, and industrial priorities, may continue to be reflected in downstream data products and decisions. "Datafication" can thus create another layer of bias: between the social, political, and cultural dimensions of data that are easy to digitize and those that are not (*14*).

DATA AS SOCIAL INFRASTRUCTURE: BIODIVERSITY MONITORING FOR EQUITABLE DECISIONS

The realization that more data or better models will never fully solve systemic bias does not mean there are no solutions. It means there are no shortcuts—no getting around the need for local engagement, context-specific knowledge, and case-by-case considerations when using this data. Investments in future monitoring should not only prioritize new technologies that ease the collection of massive amounts of biodiversity data, but also ensure that those data include information about the local context and social infrastructures.

Moving beyond quick technical fixes will require connecting strategically to community-based partners and leveraging expertise in social, ethical, cultural, and political processes underlying data infrastructures and their histories. Community-based monitoring and information systems (CBMIS) provide a compelling framework for locally engaged monitoring and are highlighted in the GBF as one means of filling data gaps (*2*). Established networks of CBMIS are already operational in several countries and have proven effective at contributing to national and global-scale monitoring of ecosystems (*15*). Initiatives such as the International Forestry Resources and Institutions (IFRI) collect information on institutional and social variables, alongside ecosystem data, through a network of locally led Collaborating Research Centers to understand the interrelationships among social and ecological processes and outcomes in forest systems around the world. There is no technocratic solution for incorporating all relevant information about ecosystems and their social contexts into formal frameworks for assessing biodiversity or devising policy strategies at global scales. However, complementing global frameworks and synthesis databases with decentralized knowledge collected as part of CBMIS (and programs like IFRI) might help expose and ameliorate data disparities that underpin biodiversity monitoring and mitigate the implications of these disparities on the distributional equity of downstream conservation decision-making.

The success of the GBF, and the meaningfulness of its proposed indicators, requires that policy-makers and scientists resist technocratic shortcuts and instead assess the equity implications of data disparities, support local knowledge generation, and work toward governance systems and monitoring frameworks that engage with biodiversity data as social infrastructure. ■

REFERENCES AND NOTES

1. A. Gonzalez, M. C. Londoño, *Science* **378**, 1147 (2022).
2. Convention on Biological Diversity, Kunming-Montreal Global Biodiversity Framework (2022).
3. A. C. Hughes *et al.*, *Ecography* **44**, 1259 (2021).
4. D. I. Warton, I. W. Renner, D. Ramp, *PLOS ONE* **8**, e79168 (2013).
5. A. Zizka *et al.*, *J. Biogeogr.* **48**, 2715 (2021).
6. D. Ellis-Soto, M. Chapman, D. H. Locke, *Nat. Hum. Behav.* **7**, 1869 (2023).
7. B. Lenzner *et al.*, *Nat. Ecol. Evol.* **6**, 1723 (2022).
8. K. M. Gaynor *et al.*, *Front. Ecol. Environ.* **14**, 533 (2016).
9. C. J. Schell *et al.*, *Science* **369**, eaay4497 (2020).
10. C. Scoville, *Theory Soc.* **48**, 1 (2019).
11. S. T. Garnett *et al.*, *Nat. Sustain.* **1**, 369 (2018).
12. B. H. Daru, J. Rodriguez, *Nat. Ecol. Evol.* **7**, 816 (2023).
13. J. E. Goldstein, H. O. Faxon, *Environ. Plan. E Nat. Space* **5**, 39 (2022).
14. M. Chapman *et al.*, *One Earth* **4**, 790 (2021).
15. M. Ferrari *et al.*, *Biodiversity (Nepean)* **2**, 57 (2015).

10.1126/science.adh8874