CONSERVATION

# Needed: Better metrics, rigorously tested

## Predictive models of biodiversity change are required to inform conservation policy decisions

By Ben Collen[1*] and Emily Nicholson[2]

Biodiversity is measured in many different ways, because no single measure adequately captures nature's many forms and functions. Over the past decade, numerous metrics for biodiversity—including species abundance, extinction risk, distribution, genetic variability, species turnover, and trait diversity—have been used to create indicators to track how biodiversity has changed (1–3). These indicators have made it clear that biodiversity loss, however it is measured, is showing little sign of abatement (1,4) and that humans must respond to safeguard the provision of natural services on which we all rely (5,6). But which metrics provide the most informative indicators under which circumstances? And how can the growing list of indicators of biodiversity change best serve conservation policy decisions?

### Alignment of target and indicator

If we are interested in the outcome of a prospering economy, we measure its performance over time using metrics such as cost of goods, income, and employment numbers. Those metrics are then used to produce indicators such as GDP and RPI, which indicate how the economy is performing. Similarly, metrics like species abundance are used to create indicators of the health of biodiversity.

For an indicator to help achieve a particular conservation target, target and indicator need to be closely aligned (4). There is little point in measuring progress toward a target with an indicator based on a metric that is only loosely related to the desired outcome. For example, ensuring that protected areas maintain their biodiversity is a fundamental goal of conservation. Targets are frequently centered on the extent of area under protection in a given country or region (e.g. 17% of land should be under protection by 2020; 6). Here, the implicit assumption is that the greater the area protected, the more biodiversity will prosper. However, this ignores factors that influence the effectiveness of the protected area: governance, funding, the type of species within them. Protected areas differ greatly in the protection they afford species, but management effectiveness indicators are currently only available for a fraction (<5%) of protected areas (4). Using just one metric as an indicator may not achieve the desired outcome.

Ideally, the chain between metric, indicator, and policy should start with specific targets. In fisheries management, targets are often very explicit, typically relating to the sustainability of fish stocks; this helps to guide fisheries policy, management interventions, and detecting fishing impacts on marine biodiversity (7). Targets can vary widely in scale; metrics such as change in total biomass, catch, and mean trophic level are used to evaluate sustainability targets for whole ecosystem management (8), whereas changes in metrics such as recruitment and abundance are used to construct indicators under alternative scenarios for management of specific fish stocks (9).

In contrast, global biodiversity targets such as those agreed to in the Convention on Biological Diversity (CBD) (6) tend to be less specific. As a result, alignment between metric, indicator, and target can be poor. For example, one CBD target calls for pollution to be reduced to levels that are not detrimental to ecosystem function or biodiversity. This laudable target does not detail important features: Which pollutants, ecosystem functions, or particular aspects of biodiversity should be addressed? The distinction is important because many functions will trade off with one another, and prioritizing some aspects of biodiversity will be at the cost of others. Efforts to measure progress toward this target, hold polluting countries and industries to account, and diagnose which interventions work best are made more difficult.

The outcomes of global biodiversity targets are perhaps inevitably less focused than those in specified circumstances such as fisheries management. However, with greater demand and scrutiny placed on biodiversity indicators (4) through targets such as CBD, how can they better support conservation efforts? One way forward is improved prediction.

### Projecting forward

Effective conservation policy decisions require an explicit understanding of the links between desired outcomes of conservation, how those outcomes can be measured, and the proposed actions needed to achieve them (10). One way to accomplish this is to project forward the impacts of a prospective policy. In doing so, both the impact of the policy and the ability of indicators to detect change in biodiversity can be measured (see the first figure). By assessing alternative policies against a suite of metrics, the best combination of metrics and indicators for evaluating policy impacts can be identified.

In a recent study, Kelly *et al.* showed

1 Centre for Biodiversity & Environment Research, University College London, Gower Street, London WC1E 6BT, United Kingdom
2 Centre of Excellence for Environmental Decisions, School of Botany, The University of Melbourne, Australia
* b.collen@ucl.ac.uk

how indicators can provide a link between broad conservation targets and local scale implementation (11). The authors applied abundance metrics to decisions on wildfire management in Australia. The results showed that optimizing fire management using an indicator based on geometric mean abundance of the community resulted in improved biodiversity outcomes compared with the conventional wisdom of maximizing the prevalence of different fire management regimes. This approach demonstrates that clearly defined management goals are necessary to maximize biodiversity in fire-prone ecosystems.

In fisheries science, substantial emphasis has been placed on ensuring that indicators respond in predictable ways to particular interventions or pressures, enabling decision makers to tease apart the impacts of different drivers of change (12). In-depth knowledge of such indicator behavior is currently lacking from analyses of most global biodiversity indicators. Put simply, any useful indicator must be able to pick up the impact of a management intervention.

**Rigorous testing**

Two aspects that affect indicator performance are the design of the indicator and the quality of the data that underpin it. An indicator may perform badly because data available to calculate it are difficult to obtain or are biased geographically or taxonomically, rendering it unrepresentative of the components of biodiversity that the indicator is designed to reflect (10). In other cases, excellent data may be available but the metric may be a poor proxy for the aspects of biodiversity of interest. For example, the rate of forest loss is often cited as a chief driver of decline of wildlife populations (5,6), but the impact of bushmeat hunting, also recognized to be a key driver, is rarely measured. Thus, biodiversity declines in forested areas can continue undetected even when the rate of forest loss is slowed or halted.

Different metrics can give varying impressions of conservation success. For example, if a wildlife population collapses leaving a much smaller population surviving in only one region, tracking abundance yields a picture of decline whereas tracking extinction risk suggests recovery (see the second figure). One problem lies in adapting metrics designed for another purpose: Extinction risk assessments provide an instantaneous snapshot, and were not designed to evaluate change through time (10). Another problem is expectation: The risk to the species has diminished towards the end of the example because the population is now stable, albeit at a much lower level than before.

Several fisheries indicators have been subjected to rigorous evaluation of whether or not they reliably predict changes in marine ecosystems (7, 8), but few biodiversity indicators have been tested in this way. There are exceptions; tests of indicator performance have shown that data biases can give a misleading impression of policy impacts (7,10). Other recent studies favorably related the Living Planet Index's underlying metric (geometric mean abundance) to models of species viability from ecological theory (13), and explored its mathematical properties (14), demonstrating it is fit for purpose to measure trends in species extinction risk.

More extensive stress testing of biodiversity indicators would enhance knowledge of how biodiversity is changing, show whether the existing indicators can measure that change, and help identify the most appropriate policies to counteract biodiversity declines. Predictive modelling will help ensure that biodiversity indicators are capable of supporting conservation policy decisions. For example, sampling model systems, mimicking the way data are collected in biodiversity monitoring programmes, can be used to calculate indicators, and to provide a completeness that is not available in real-world monitoring data. This framework—referred to as management strategy evaluation—has been used to test indicators for fisheries management (7, 8) and has also been applied to the evaluation of other biodiversity indicators (10), showing that while some indicators perform well, others need rethinking.

**Toward a meaningful set of metrics**

If the right information to guide conservation policy cannot be gleaned from existing metrics, then gathering global-level data for an array of new metrics would be a costly endeavor. But the inconsistent delivery measures of biodiversity change means that a set of agreed metrics of biodiversity is urgently required (2). Striking the right balance between expanding existing datasets and developing new, more appropriately designed monitoring programs and metrics will be vital if measures of biodiversity change are to robustly support conservation decisions.

In doing so, conservation science must be rigorous. Testing the modeled performance of alternative management actions prior to implementation should be the gold standard for conservation decision making. Indicators of change must also be subjected to rigorous performance tests. Such evaluation was mentioned in the selection of indicators of the CBD 2010 target, with all indicators "identified for immediate testing." Yet, with few exceptions, the indicators remain largely unevaluated in their capacity to report meaningfully on conservation targets and the means of achieving them; this remains a critical task for predictive conservation science if it is to influence conservation progress.

**REFERENCES AND NOTES**

1. S. H. M. Butchart *et al*., *Science* **328**, 1164 (2010).
2. H. M. Pereira *et al*., Science **339**, 277 (2013).
3. B. Collen *et al*., *Conserv. Biol*. **23**, 317 (2009).
4. D. P. Tittensor *et al*., *Science* **xxx**, xxxx (2014).
5. G. M. Mace et al., *Biodiversity: Millennium Ecosystem Assessment: current state and trends: findings of the condition and trends working group. Ecosystems and Human well-being* (Island Press, 2005).
6. Secretariat of the Convention on Biological Diversity, Global Biodiversity Outlook 3 (2010).
7. T. A. Branch *et al., Nature* **468**, 431 (2010).

8. E. Fulton, A. Smith, A. Punt, *ICES J. Mar. Sci.* **62**, 540 (2005).
9. F. J. Mueter, N. A. Bond, J. N. Ianelli, A. B. Hollowed, ICES J. Mar. Sci. **68**, 1284 (2011).
10. E. Nicholson *et al*., *PLoS One* **7**, e41128 (2012).
11. L. T. Kelly, A. F. Bennett, M. F. Clarke, M. A. McCarthy, *Conserv. Biol.* 10.1111/cobi.12384 (2014).
12. J. S. Link *et al., ICES J. Mar. Sci*. **67**, 787 (2009).
13. M. McCarthy *et al*., *Conserv. Biol*. **00**, 1 (2014).
14. S. T. Buckland A. E. Magurran, J. B. Illian, S. E. Newson, *Ecosphere* **2**, 1 (2011).



**Acknowledgements:** EN is supported by a Centenary Research Fellowship.


**The power of prediction.** Predictive modeling of the impacts of alternative policies on biodiversity (A, B and C *vs*. business-as-usual, BAU). Assessment of alternative policies demonstrates their potential contribution to meeting biodiversity targets.

**Recovery or decline?** Two metrics lead to different conclusions following the regional extirpation of a hypothetical species to a low but stable population size.