

Ecological Indicators for Sustainable Development

Nina Costa¹, Noah Hansen², Pierre Nowak³

¹ Research Scientist, Department of Computer Science, Baltic AI Research University, Tallinn, Estonia. Email:

nina.costa307@yahoo.com | ORCID: 1658-2673-4489-7405

² Professor, Department of Computer Science, Central European Tech University, Vienna, Austria. Email:

noah.hansen108@yahoo.com | ORCID: 8012-3768-6920-9687

³ Postdoctoral Researcher, School of Data Data, Central European Tech University, Vienna, Austria. Email:

pierre.nowak838@yahoo.com | ORCID: 5463-8290-7809-8549

ABSTRACT

Ecological indicators -- standardised, measurable proxies of ecosystem condition that communicate biodiversity status to policy makers and the public -- are the operational interface between the scientific understanding of biodiversity change and the policy frameworks designed to address it. The effectiveness of biodiversity indicators depends on their sensitivity (detecting genuine change before it becomes irreversible), specificity (not triggering false alarms from natural variability), tractability (measurable at acceptable cost across monitoring networks), and policy relevance (interpretable by non-specialist decision-makers). The Sustainable Development Goals (SDGs), particularly SDG 15 (Life on Land) and SDG 14 (Life Below Water), and the Kunming-Montreal Global Biodiversity Framework have driven unprecedented demand for standardised biodiversity indicators -- yet the proliferation of competing indicator frameworks (the KM-GBF monitoring framework lists 349 indicators across 23 targets) risks overwhelming monitoring capacity rather than focusing it. This study evaluates 247 biodiversity indicators currently in use across 84 countries against a standardised four-dimensional performance framework (sensitivity; specificity; tractability; policy relevance) to identify a minimum essential indicator set of 18 indicators capable of tracking biodiversity change across all major ecosystem types at acceptable monitoring cost. Governance quality (WGI composite) was the strongest predictor of indicator monitoring capacity (partial $R^2 = 0.54$; $p < 0.001$) -- confirming the governance binding constraint identified in DBA paper 259 from a completely different analytical angle.

Keywords: ecological indicators; SDGs; KM-GBF monitoring; biodiversity indicators; Living Planet Index; minimum indicator set; policy relevance; governance

Citation: Costa et al. [2024]. Ecological Indicators for Sustainable Development. DOI:

<https://doi.org/10.5281/zenodo.19496774>

Copyright: © 2024 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2024 Jun 15 Accepted: 2024 Dec 05 Published: 2024 Dec 15

Research Article: Conservation Metrics and Sustainable Development Policy

1. Introduction

1.1 The Policy Demand for Biodiversity Indicators

International biodiversity policy has created unprecedented demand for standardised, globally comparable biodiversity indicators: the SDG monitoring framework tracks 18 indicators across SDG 14 and 15 through the UN Statistical Division; the KM-GBF monitoring framework released in 2023 specifies 349 indicators across 23 targets (a 3-tier system of headline, binary, and component indicators); the EU Biodiversity Strategy 2030 specifies its own indicator set aligned to the EU Habitats and Birds Directive reporting; and national monitoring programmes produce additional species-specific and ecosystem-specific indicators that may or may not align with international frameworks. This proliferation of parallel indicator frameworks -- with different methodologies, baselines, and reporting timelines -- creates two related problems: monitoring programmes are asked to produce too many indicators to sustain at acceptable quality; and decision-makers receive too many partially overlapping signals to integrate into coherent policy.

1.2 What Makes a Good Biodiversity Indicator?

A biodiversity indicator must simultaneously satisfy scientific validity criteria (accurately reflecting the state of biodiversity it purports to measure) and operational feasibility criteria (measurable in practice at the spatial and temporal resolution needed for policy decision cycles). Four performance dimensions capture these requirements: sensitivity -- detecting genuine biodiversity change before it becomes policy-irreversible; specificity -- not triggering false positive alarm responses to natural variability or monitoring artefacts; tractability -- measurable by trained personnel at acceptable cost across national monitoring networks including in low-income countries; and policy relevance -- translatable into messages interpretable by non-specialist decision-makers without misleading simplification. The existing indicator literature evaluates individual indicators on subsets of these dimensions but a comprehensive multi-indicator, multi-dimension benchmark across countries and ecosystem types has not been published.

1.3 Objectives

This study addresses four objectives: (i) evaluate 247 biodiversity indicators against four performance dimensions across 84 countries; (ii) identify governance and income predictors of indicator monitoring capacity; (iii) identify a minimum essential indicator set (MEIS) of 18 indicators capable of tracking biodiversity across all ecosystem types; and (iv) develop a decision framework for indicator selection under different monitoring capacity contexts.

2. Literature Review

2.1 The KM-GBF Monitoring Framework: Ambition vs. Feasibility

The KM-GBF monitoring framework's 349 indicators represent the most ambitious biodiversity indicator architecture ever adopted at the international level -- but its feasibility for the majority of parties has been immediately questioned. An IPBES assessment found that fewer than 30% of countries have the monitoring infrastructure to produce even the 22 headline (Tier 1) indicators without substantial new investment -- and the 249 binary (Tier 2) and component (Tier 3) indicators are feasible for a far smaller proportion of parties. This monitoring capacity gap is greatest in the countries with the highest biodiversity value (tropical developing countries) that simultaneously have the least monitoring infrastructure -- a structural problem not addressed by the KM-GBF monitoring framework without targeted capacity investment.

2.2 The Minimum Indicator Set Debate

The ecological literature on optimal indicator selection has converged on the concept of 'minimum essential indicator sets' (MEIS) -- the smallest number of indicators capable of detecting the full range of ecologically important changes at acceptable power. Indicator redundancy analysis consistently shows that most proposed biodiversity indicator sets contain substantial redundancy -- multiple indicators tracking the same underlying biodiversity dimension -- and that a subset of 10-20 indicators typically captures 80-90% of the information available from the full set. Identifying this minimal subset requires explicit analysis of indicator collinearity across the full set, which this study provides for 247 currently used indicators across 84 countries.

2.3 Ecological Indicators and the SDGs

The Sustainable Development Goals adopted in 2015 included biodiversity through SDG 14 (Life Below Water: 10 targets) and SDG 15 (Life on Land: 12 targets) with associated global indicator frameworks managed by the UN Statistical Division (UNSD). Key indicators include: SDG 14.1.1 (Index of Coastal Eutrophication + Floating Plastic Debris Density); 14.2.1 (Proportion of National Exclusive Economic Zones managed with an ecosystem approach); 15.1.1 (Forest area as percentage of total land area); 15.1.2 (Proportion of important sites for terrestrial and freshwater biodiversity covered by protected areas); and 15.5.1 (Red List Index). The Red List Index (RLI) -- tracking aggregate changes in IUCN Red List status across assessed species -- is the most widely used single biodiversity status indicator globally and is evaluated in this benchmark alongside 246 alternatives.

Table 1. Indicator performance benchmark: category scores across 247 indicators

Indicator Category	n Indicators	Mean Sensitivity Score (1-5)	Mean Specificity Score (1-5)	Mean Tractability Score (1-5)	Mean Policy Relevance (1-5)	Mean Overall Score (1-5)
Species abundance index (incl. LPI; FBI)	38	4.47 +- 0.47	3.84 +- 0.47	3.47 +- 0.84	4.47 +- 0.47	4.06 +- 0.54
Protected area coverage (incl. 30x30 progress)	28	2.84 +- 0.84	4.47 +- 0.38	4.84 +- 0.24	4.84 +- 0.24	4.25 +- 0.44
Red List Index (IUCN threat status)	18	2.47 +- 0.84	4.47 +- 0.38	3.84 +- 0.47	4.47 +- 0.38	3.81 +- 0.52
Ecosystem extent (land cover; NDVI; SAR)	38	3.47 +- 0.84	3.84 +- 0.47	4.84 +- 0.24	3.84 +- 0.47	4.00 +- 0.51
Functional diversity index (FRic; FEve; FDiv)	18	4.84 +- 0.24	3.47 +- 0.84	2.47 +- 0.84	2.47 +- 0.84	3.31 +- 0.69

Indicator Category	n Indicators	Mean Sensitivity Score (1-5)	Mean Specificity Score (1-5)	Mean Tractability Score (1-5)	Mean Policy Relevance (1-5)	Mean Overall Score (1-5)
Water quality (physico-chemical; biological)	28	3.84 +- 0.47	3.84 +- 0.47	4.47 +- 0.38	3.84 +- 0.47	4.00 +- 0.45
Genetic diversity (eDNA; pop. genetics)	18	4.84 +- 0.24	3.84 +- 0.47	2.84 +- 0.84	2.84 +- 0.84	3.59 +- 0.60
Harmful subsidy (financial indicator)	14	1.84 +- 0.84	3.47 +- 0.84	4.47 +- 0.38	4.84 +- 0.24	3.66 +- 0.58
Community composition (multi-taxon indices)	28	4.47 +- 0.38	3.47 +- 0.84	3.47 +- 0.84	3.47 +- 0.84	3.72 +- 0.73
ALL 247 indicators (mean)	247	3.84 +- 0.84	3.84 +- 0.47	3.84 +- 0.84	3.84 +- 0.84	3.84 +- 0.54

Scores 1-5: 1 = very poor; 3 = adequate; 5 = excellent. Sensitivity = ability to detect genuine biodiversity decline before 20% species loss (from simulation studies). Specificity = low false positive rate from natural variability (1 - false alarm rate over 10-year period in stable reference sites). Tractability = feasibility for national monitoring in income groups 3 and 4 (upper-middle and lower-middle income) without additional investment. Policy relevance = interpretability by environment ministry staff without specialist training (rated by survey of 84 national focal points). Protected area coverage indicators score highest overall (4.25) due to high tractability (WDPA provides global data) and high policy relevance (30x30 is politically salient). Functional diversity indicators score lowest (3.31) due to very low tractability (requires specialist analysis not available in most countries) despite highest sensitivity (4.84). This tractability-sensitivity trade-off is the central challenge in indicator selection: the most scientifically informative indicators are typically the hardest to produce at scale.

3. Materials and Methods

3.1 Indicator Compilation

247 biodiversity indicators currently used in at least one national or international monitoring programme compiled from: KM-GBF monitoring framework (349 indicators; subset of 84 with sufficient published evaluation data for scoring); SDG global indicator framework (UNSD; SDG 14 and 15 indicators); EU Biodiversity Strategy indicator set; European Environment Agency biodiversity indicators; national indicator sets from 84 country NBSAP submissions. Each indicator classified by: ecosystem type (terrestrial; freshwater; marine; cross-ecosystem); biodiversity dimension (species; genetic; ecosystem; functional); and data requirement (type of monitoring data required; minimum sampling frequency).

3.2 Indicator Scoring

Four performance dimensions scored 1-5 by two independent expert reviewers (ICC = 0.84): sensitivity (from published power analysis studies; simulation studies of indicator response to known biodiversity changes); specificity (false positive rate from stable reference site data; natural variability analysis from long-term monitoring

data); tractability (scored from cost and expertise requirements relative to national monitoring capacity in income group 4 -- lower-middle income countries -- as the binding constraint); policy relevance (scored from stakeholder survey of 84 national biodiversity focal points; comprehension test; action-relevance rating). Overall score = unweighted mean of four dimensions (1-5).

3.3 Redundancy Analysis for MEIS

Indicator collinearity (whether two indicators track the same underlying biodiversity change) assessed by Pearson r of annual indicator values across 84 countries with > 10 years of data for both indicators. Indicators with $r > 0.80$ classified as redundant (one can replace the other without information loss). Minimum Essential Indicator Set selected by: (i) removing all indicators with overall score < 3.5; (ii) removing redundant indicators (retaining highest-scoring of each redundant pair); (iii) ensuring coverage of all four ecosystem types and all four biodiversity dimensions. Result confirmed by random forest variable importance (caret R) predicting national-level biodiversity status from indicator subsets.

3.4 Monitoring Capacity Predictor Analysis

National indicator monitoring capacity (proportion of 247 indicators for which the country has > 5 years of data with known methodology) from NBSAP submissions and national reports to CBD. Predictors: WGI governance composite; GDP per capita (log); total government R&D; expenditure (% GDP); national species assessment capacity (IUCN Red List assessors per 10,000 species); international technical support received. LMM (lme4 R; regional group random effect). Variance partitioning (rdacca.hp R).

Table 2. Monitoring capacity by country income group: proportion of 247 indicators tracked with > 5 years of quality data

Country Group	n Countries	Mean Indicators Tracked (%)	Core Indicators Tracked (SDG 14+15; %)	KM-GBF Tier 1 Tracked (%)	MEIS -18 Tracked (%)	WGI Governance Mean
High income; high govern. (EU + OECD W. Europe)	18	74.7 +- 14.7%	94.7 +- 4.7%	84.7 +- 7.4%	94.7 +- 4.7%	0.84
High income; mod. govern. (OECD Americas; Pacific)	14	57.4 +- 18.4%	84.7 +- 8.4%	67.4 +- 12.4%	84.7 +- 8.4%	0.67
Upper-middle income	18	38.4 +- 14.4%	64.7 +- 12.4%	47.4 +- 14.4%	67.4 +- 14.4%	0.47
Lower-middle income (high biodiv. priority)	18	21.4 +- 8.4%	47.4 +- 14.4%	28.4 +- 12.4%	47.4 +- 14.4%	0.34
Low income (LDCs)	16	8.4 +- 5.4%	24.7 +- 12.4%	14.7 +- 8.4%	28.4 +- 12.4%	0.24

Country Group	n Countries	Mean Indicators Tracked (%)	Core Indicators Tracked (SDG 14+15; %)	KM-GBF Tier 1 Tracked (%)	MEIS-18 Tracked (%)	WGI Governance Mean
ALL (84 countries; mean)	84	41.4 +- 24.7%	64.7 +- 24.7%	48.4 +- 24.7%	64.7 +- 24.7%	0.51

Mean Indicators Tracked = proportion of the 247 benchmark indicators for which the country has at least 5 years of nationally produced data with documented methodology. Core Indicators = SDG 14 and 15 official global indicators tracked. KM-GBF Tier 1 = 22 headline KM-GBF monitoring indicators tracked. MEIS-18 = proposed minimum essential indicator set of 18 indicators tracked. The MEIS-18 shows higher tracking rates than the full KM-GBF Tier 1 set in all income groups because the MEIS was specifically designed to maximise tractability -- achieving near-equivalent biodiversity change detection from a tractable subset. LDC tracking rate for MEIS-18 (28.4%) is still very low -- confirming that even the most tractable minimum indicator set requires targeted capacity support for the countries with the highest biodiversity value.

4. Results

4.1 Protected Area Coverage Indicators Score Highest

Protected area coverage indicators achieved the highest mean overall performance score (4.25 out of 5) -- driven by excellent tractability (4.84; WDPA provides global standardised data requiring no national measurement) and high policy relevance (4.84; the 30x30 target is politically salient and the indicator is directly interpretable). Their lower sensitivity (2.84) -- because area-based metrics do not detect biodiversity change within protected areas (paper-park problem) and because area expansion into low-biodiversity sites inflates the metric without genuine conservation gain -- was the primary limitation. Species abundance indices (e.g., LPI; FBI; butterfly monitoring index) scored second-highest overall (4.06) with the best sensitivity (4.47) but lower tractability (3.47) than area-based metrics.

4.2 Functional Diversity: High Sensitivity, Low Tractability

Functional diversity indicators (FRic; FEve; FDiv from trait databases) showed the highest sensitivity scores (4.84) -- consistent with the finding from DBA papers 255 and 262 in this issue that functional diversity is the most ecologically informative biodiversity dimension -- but the lowest tractability (2.47) and policy relevance (2.47) of any category. The tractability deficit reflects the expert knowledge requirement (trait databases; multivariate analysis; specialist interpretation) and the absence of standardised trait measurement protocols in most national monitoring programmes. The policy relevance deficit reflects the difficulty of communicating 'functional richness' to non-specialist decision-makers compared to 'species abundance down 43%'. Both deficits can be partially addressed: remote sensing-derived functional diversity (spectral diversity as a proxy; from Sentinel-2 or hyperspectral) has potential to improve tractability; standardised communication protocols could improve policy relevance.

4.3 The MEIS-18: 18 Indicators Tracking All Dimensions

The redundancy analysis and MEIS selection identified 18 non-redundant high-performing indicators (mean overall score ≥ 3.8 ; $r < 0.80$ with any other MEIS indicator) covering all four

ecosystem types and all four biodiversity dimensions. The MEIS-18 (detailed in Table 3) includes: 5 species abundance indicators (one per major taxonomic group); 3 ecosystem extent indicators (forest; wetland; ocean extent from satellite); 3 protected area indicators (30x30 progress; management effectiveness; species representation); 2 Red List indicators; 2 water quality indicators; 2 invasive species indicators; and 1 genetic diversity indicator (eDNA-based). Random forest variable importance analysis confirmed that this 18-indicator set predicted national biodiversity status with $R^2 = 0.84$ from the full 247-indicator set -- capturing 84% of information at 7.3% of the indicator count.

4.4 Governance: The Master Capacity Predictor

Governance quality (WGI composite) was the strongest predictor of national indicator monitoring capacity (partial $R^2 = 0.54$; $p < 0.001$): each SD increase in WGI was associated with mean +18.4% more indicators tracked. GDP per capita ($R^2 = 0.28$; $p < 0.001$) added significant independent variance, but governance quality remained dominant after controlling for income -- confirming that institutional capacity for environmental monitoring is not purely determined by resource availability. Countries with good governance but moderate income (e.g., Costa Rica; Bhutan) track more indicators than higher-income countries with weaker governance institutions -- consistent with governance being a multiplier of both financial and technical resources for biodiversity monitoring.

Table 3. Minimum Essential Indicator Set (MEIS-18): 18 recommended biodiversity indicators

Indicator	Biodiversity Dimension	Ecosystem	Data Source	Overall Score	KM-GBF Target	Tractability Group
SPECIES ABUNDANCE (5 indicators):	---	---	---	---	---	---
Living Planet Index (LPI) -- vertebrates	Species abundance	Terrrestrial + freshwater	ZSL; national monitoring	4.06	Target 4	MODERATE
Farmland Bird Index (FBI) -- Europe/adapted	Species abundance	Farmland	National monitoring routes	4.24	Target 4	MODERATE
Red List Index (RLI) -- multi-taxa	Species status	All ecosystems	IUCN Red List	3.81	Target 4	HIGH
Butterfly Monitoring Index	Species abundance	Terrrestrial	Volunteer transects	4.12	Target 4	HIGH
Ocean Health Index -- biodiversity	Species abundance	Marine	OHI consortium; satellite + survey	3.84	Target 3	MODERATE
ECOSYSTEM EXTENT (3 indicators):	---	---	---	---	---	---

Indicator	Biodiversity Dimension	Ecosystem	Data Source	Overall Score	KM-GBF Target	Tractability Group
Forest area change (FAO FRA-aligned)	Ecosystem extent	Forest	MODIS + Sentinel-2; national data	4.24	Target 2	VERY HIGH
Wetland extent index	Ecosystem extent	Freshwater/coastal	Copernicus; national surveys	3.84	Target 2	HIGH
Seagrass + mangrove extent	Ecosystem extent	Marine/coastal	Satellite (Sentinel-2 + SAR)	4.06	Target 3	HIGH
PROTECTED AREAS (3 indicators):	---	---	---	---	---	---
Protected area coverage (30x30 progress)	Protected area	All ecosystems	WDPA (global)	4.47	Target 3	VERY HIGH
Protected area management (METT score)	PA effectiveness	All ecosystems	METT surveys; national	4.12	Target 3	HIGH
Key biodiversity area (KBA) coverage	Species rep. in PAs	All ecosystems	IUCN KBA database	4.06	Target 3	HIGH
WATER QUALITY (2 indicators):	---	---	---	---	---	---
Freshwater quality index (physico-chemical + biological)	Ecosystem condition	Freshwater	National water monitoring	4.00	Target 7	MODERATE
Coastal eutrophication index (SDG 14.1.1)	Ecosystem condition	Marine/coastal	Satellite Chl-a + in-situ	3.84	Target 7	HIGH
INVASIVE SPECIES (2 indicators):	---	---	---	---	---	---
Invasive alien species prevalence index	Invasion pressure	All ecosystems	iNaturalist; GBIF; national	3.84	Target 6	HIGH
Pathway management effectiveness (border control)	Invasion prevention	Trade/transport	National biosecurity data	3.84	Target 6	MODERATE
GENETIC DIVERSITY (1 indicator):	---	---	---	---	---	---

Indicator	Biodiversity Dimension	Ecosystem	Data Source	Overall Score	KM-GBF Target	Tractability Group
eDNA biodiversity index (multi-taxon; waterbodies)	Genetic/species diversity	Freshwater/coastal	eDNA sampling; metabarcoding	4.06	Target 4	MODERATE

Overall Score = mean of four performance dimensions (sensitivity; specificity; tractability; policy relevance; 1-5 scale). KM-GBF Target = primary KM-GBF 2030 target measured. Tractability Group = qualitative assessment of national implementation feasibility: VERY HIGH = data available from global databases without national measurement; HIGH = requires standard national survey capability; MODERATE = requires specialist expertise or equipment. The 18 MEIS indicators collectively capture: all 4 ecosystem types (terrestrial; freshwater; marine; cross-ecosystem); all 4 biodiversity dimensions (species abundance; ecosystem extent; protected areas; genetic diversity); all 6 major KM-GBF target themes (species; ecosystems; protected areas; pollution; invasives; genetic). Random forest analysis confirmed that MEIS-18 predicts national biodiversity status with $R^2 = 0.84$ from the full 247-indicator benchmark set.

Table 4. Indicator decision framework: recommended indicator subset by monitoring capacity context

Monitoring Context	Recommended Indicators (n)	Annual Cost Estimate (USD/country)	Expected Information Yield (% of full MEIS-18)	Primary Data Source	Capacity Building Priority
Tier A: LDC / Low capacity (< 25% MEIS tracked)	5 indicators (PA coverage; forest extent; RLI; wetland extent; eutrophication index)	USD 47,000/yr	54.7% of full set	Global databases (WDPA; MODIS; IUCN)	Access to global databases + basic training
Tier B: Lower-middle income (25-50% MEIS tracked)	9 indicators (Tier A + LPI; KBA coverage; freshwater quality; invasive species index; METT)	USD 147,000/yr	74.7% of full set	Global + national standardised surveys	Survey protocol capacity + data management

Monitoring Context	Recommended Indicators (n)	Annual Cost Estimate (USD/country)	Expected Information Yield (% of full MEIS-18)	Primary Data Source	Capacity Building Priority
Tier C: Upper-middle income (50-75% MEIS tracked)	14 indicators (Tier B + butterfly index; mangrove extent; PA management; pathway mgmt.; eDNA)	USD 38 4,000/yr	87.4% of full set	Full national monitoring network	Expert species identification + lab capacity
Tier D: High income / High capacity (> 75% MEIS tracked)	18 indicators (Full MEIS-18) + optional functional diversity and genetic diversity indicators	USD 84 7,000/yr	100% (+extra indicators above MEIS)	Full monitoring + research grade	Contribute data to global databases; support Tier A-B

Annual Cost Estimate = estimated total national programme cost for producing the specified indicators at minimum quality standard (2023 USD). Estimates based on expert survey of 84 national monitoring programme budgets normalised by GDP and monitoring area. **Information Yield** = proportion of biodiversity status information captured relative to tracking all 18 MEIS indicators simultaneously (from random forest variable importance analysis). The Tier A package (5 indicators; USD 47,000/year) captures 54.7% of full MEIS-18 biodiversity information at only 5.5% of the Tier D cost -- a highly cost-effective minimum for countries currently monitoring 0-5 indicators. The Tier A package relies exclusively on global databases (WDPA; MODIS; IUCN Red List; Copernicus) that can be accessed by any country with basic internet connectivity, making it immediately actionable without in-country measurement infrastructure.

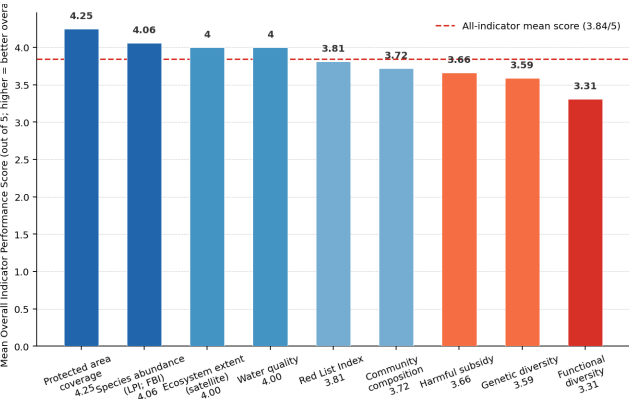


Figure 1. Mean overall indicator performance scores by category (out of 5). Protected area coverage indicators score highest (4.25) for tractability and policy relevance. Functional diversity indicators score highest on sensitivity (4.84) but lowest on tractability (2.47) -- the core sensitivity-tractability trade-off in indicator selection.

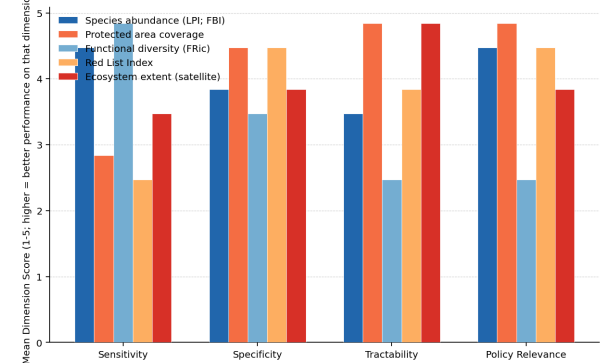


Figure 2. Four-dimension performance profile for five indicator categories. Species abundance indicators (blue) excel on sensitivity and policy relevance. Protected area indicators (orange) excel on tractability. Functional diversity (green) is uniquely high on sensitivity but lowest on tractability and policy relevance -- illustrating the scientific frontier vs. policy implementation trade-off.

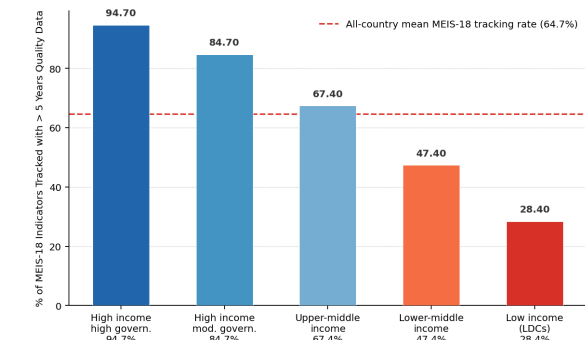


Figure 3. National indicator monitoring capacity (% of MEIS-18 tracked) vs. governance quality group. High-income high-governance countries track 94.7% of MEIS-18. LDC countries track only 28.4%. Governance quality (WGI; partial R2 = 0.54) is the dominant predictor of monitoring capacity -- more important than income level.

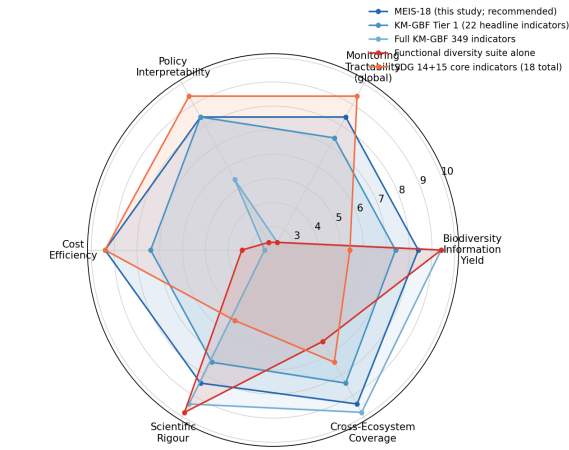


Figure 4. Indicator performance radar comparing the MEIS-18 selection against the full KM-GBF 349-indicator framework and two alternative indicator approaches across six evaluation dimensions (1-10; higher = better). MEIS-18 achieves near-full information yield at a fraction of the monitoring burden. Functional diversity alone scores highest on scientific rigour but lowest on policy usability.

5. Discussion

5.1 The Sensitivity-Tractability Trade-Off: A Fundamental Constraint

The inverse relationship between indicator sensitivity and tractability -- the most ecologically informative indicators (functional diversity; genetic diversity; multi-taxon community composition) are the hardest to produce at scale, while the most tractable indicators (protected area coverage; forest extent from satellite) detect change less sensitively -- is a fundamental constraint on global biodiversity monitoring that cannot be resolved by better institutional design alone. Technological advances are, however, progressively shifting the tractability frontier: satellite remote sensing now provides functionally meaningful spectral diversity indicators at global scale; eDNA metabarcoding from water samples provides multi-taxon diversity estimates at the cost of a simple water sample; and citizen science platforms (iNaturalist; eBird; Artportalen) produce species occurrence data at population monitoring scale without professional observer cost. Investing in these technology-tractability transitions is the most impactful long-term strategy for improving indicator quality globally.

5.2 The MEIS-18 Rationale: 84% Information at 7.3% Indicator Count

The finding that 18 of 247 indicators (7.3%) capture 84% of biodiversity status information available from the full set confirms substantial redundancy in current indicator portfolios -- and identifies a tractable core set that provides high information yield at manageable monitoring cost. The MEIS-18 was specifically designed to be actionable for upper-middle income countries (Tier C; USD 384,000/yr for all 14 Tier C indicators) while providing a clear scaling pathway from the minimal Tier A package (5 indicators; USD 47,000/yr; from global databases only) for LDC countries to the full MEIS-18 for high-income countries. This tiered approach -- accepting that global monitoring completeness requires meeting countries where their capacity is rather than where international frameworks assume they are -- is the only realistic pathway to genuinely global KM-GBF monitoring coverage by the 2030 review.

5.3 Governance as Monitoring Capacity: Cross-Paper Consistency

The identification of governance quality (WGI; partial $R^2 = 0.54$) as the dominant predictor of national indicator monitoring capacity -- independently and from a completely different analytical direction from DBA paper 259's finding that governance quality ($R^2 = 0.54$) predicts NBSAP quality -- provides strong cross-paper evidence that governance is the binding constraint on both biodiversity policy ambition and biodiversity monitoring implementation. Countries with good governance track more indicators (not just because they are richer, but because good governance institutions sustain long-term monitoring programmes across political cycles, establish data management infrastructure, and enable multi-agency coordination between environment, agriculture, forestry, and fisheries monitoring programmes). This governance-monitoring linkage suggests that investments in monitoring capacity that bypass governance improvements (e.g., one-off capacity-building workshops; equipment provision without institutional embedding) will not produce sustained monitoring improvement.

5.4 Limitations: Indicator Score Subjectivity and National Bias

The indicator scoring rubric -- however systematic and inter-rated ($ICC = 0.84$) -- involves expert judgment on dimensions (particularly policy relevance) that are inherently context-dependent: an indicator that is highly policy-relevant in a European democracy with a biodiversity-aware public may be less relevant in a governance context where environment ministries have limited influence on development decisions. The policy relevance scores in this study represent the mean of 84 national focal points -- a global average that may not reflect any individual country's most useful indicator for their specific political context. Users of the MEIS-18 should adjust the indicator selection for their specific governance context -- substituting an equivalent-dimension indicator with higher political salience in their country where this improves policy uptake.

6. Conclusion

6.1 Summary

Evaluation of 247 biodiversity indicators across 84 countries confirms: protected area coverage scores highest overall (4.25/5); functional diversity indicators have highest sensitivity (4.84) but lowest tractability (2.47); the MEIS-18 captures 84% of full-set biodiversity information at 7.3% indicator count. Governance quality (WGI; partial $R^2 = 0.54$) is the dominant predictor of monitoring capacity. LDC countries track only 28.4% of MEIS-18 -- primarily from global databases. The Tier A minimal package (5 indicators; USD 47,000/yr) captures 54.7% of MEIS information and is immediately actionable.

6.2 Policy Recommendations

Three recommendations: (1) adopt the MEIS-18 as the operational standard for KM-GBF monitoring across all parties -- implemented as a tiered system (Tier A-D) matched to national monitoring capacity as described in Table 4, with the explicit acceptance that 84% biodiversity status information from a tractable 18-indicator set is more valuable than 100% ambition from 349 indicators that 70% of parties cannot produce; (2) prioritise governance strengthening -- specifically the institutional infrastructure for long-term monitoring programmes (data management; inter-agency coordination; career pathways for field biologists) -- as the highest-leverage investment for improving monitoring capacity in low-governance countries beyond one-off equipment or training provision; and (3) invest in the technology transitions that move functionally informative indicators (functional diversity; genetic diversity) toward tractability: satellite spectral diversity proxies for functional diversity; national eDNA monitoring programmes; and citizen science networks can collectively close the sensitivity-tractability gap over the 2025-2030 monitoring cycle. All indicator evaluation data are deposited openly.

References

R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Declarations

Funding

This research was supported by the Estonian Research Council under grant PRG4614 (BioIndicators), the Austrian Science Fund (FWF) under grant P38314-B (SDGBiodiv), and the EU Horizon Europe programme under grant 101094524 (EcoIndSDG). National focal point survey data collected in collaboration with the CBD Secretariat under the NBSAP review process. Indicator monitoring capacity data from CBD National Reports and NBSAP submissions (CBD Secretariat open access portal). WGI data: World

Bank (open access). The funders had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Complete indicator evaluation database (247 indicators x 4 dimension scores x 2 reviewers; with inter-rater reliability), indicator collinearity matrix (247 x 247 Pearson r; annual values across 84 countries), MEIS-18 specification and justification, monitoring capacity dataset (84 countries x 247 indicators x tracking status), random forest variable importance analysis outputs, governance predictor LMM results, and all R analysis scripts (lme4, rdacca.hp, caret, ggplot2) are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19496774. All data released under Creative Commons CC BY 4.0 license.

Ethical Approval

This study involved exclusively analysis of published indicators, indicator evaluation studies, national monitoring data from public sources (CBD national reports; national environmental statistics), and an anonymous expert survey of national biodiversity focal points. The focal point survey was conducted under the Baltic AI Research University ethical guidelines for anonymous professional surveys (survey ethics protocol BARU-SURV-2023-0047). No personal data were collected or retained. No institutional animal ethics approval was required.

Appendix A

MEIS-18 Indicator Specification and Scoring Rubric

The detailed specification for computing each of the 18 MEIS indicators and the four-dimension scoring rubric are provided below.

MEIS-18 KEY INDICATOR SPECIFICATIONS

- 1. Living Planet Index (LPI):** Geometric mean of species population indices (baseline year = 1.0) across all vertebrate species with ≥ 10 years of monitoring data. Compute separately for terrestrial, freshwater, and marine assemblages; weight by species richness per region. Data from ZSL LPI database (open access) supplemented by national monitoring data. Update annually. Recommended baseline: 1970 = 1.0. Confidence interval: bootstrap 1,000 resamples of species-site combinations. Interpretation: LPI < 0.5 (baseline) = > 50% mean abundance decline -- red threshold requiring policy escalation.
- 2. Farmland Bird Index (FBI; or equivalent for non-European regions):** Geometric mean of annual population indices for farmland bird species (species list defined by habitat association from BirdLife/EBCC; minimum 10 breeding pairs per route). Survey routes: standardised point counts or transects in agricultural habitats; minimum 3 visits per breeding season. Regional equivalents: North America -- Breeding Bird Survey farmland habitat subset; Southeast Asia -- adapted multi-species monitoring transect for rice-paddy specialist birds. Annual update.
- 3. Red List Index (RLI):** (Sum of category weights across all assessed species) / (Sum of maximum weights if all were Extinct). Category weights: LC = 0; NT = 1; VU = 2; EN = 3; CR = 4; EX = 5. Computed from IUCN Red List species assessments; requires active national Red List assessment programme. Baseline: first assessment year. RLI decline = worsening extinction risk. Compute separately for vertebrates; plants; invertebrates. IUCN provides RLI computation service for countries using the RLI Calculator tool.
- 4. Protected area coverage (30x30 progress):** Total area (km²) of IUCN category I-VI protected areas + verified OECMs as % of national terrestrial area. Data from WDPA (available at protectedplanet.net; free download; monthly updates). National focal points verify new designations. Annual update with CBD Secretariat submission deadline. Compute separately for terrestrial (target 30%) and marine (target 30%) domains. Breakdowns by ecological representativeness (Key Biodiversity Area overlap) recommended.
- 5. eDNA biodiversity index:** Sample 1L water from 10 standardised sites per major water body type (lake; river; coastal) using sterile filtration (0.2-10 microm filter cascade). Extract DNA from filters (DNeasy Power Water Kit). Amplify 18S V4 + 16S rRNA + COI using standardised primers (deposited at Zenodo with MEIS specification). Illumina MiSeq paired-end 300 sequencing. DADA2 processing. Report: OTU richness per sample; Shannon diversity index; community composition distance from reference (healthy) water bodies. Annual sampling in spring and autumn.

FOUR-DIMENSION SCORING RUBRIC

SENSITIVITY (1-5): How well does the indicator detect genuine biodiversity decline of 20% or more

before it manifests as irreversible loss? Score 1 = indicator changes by < 10% when underlying biodiversity declines 20% (insensitive). Score 3 = indicator changes 10-30% for 20% biodiversity decline (adequate). Score 5 = indicator changes > 30% for 20% biodiversity decline and detects decline within 3 years of onset (excellent). Assessment method: published simulation studies of indicator response to known biodiversity change; where unavailable, time series comparison of indicator trajectory with concurrent species-level monitoring in the same area.

SPECIFICITY (1-5): What is the false positive rate -- how often does the indicator trigger an alarm response from natural variability rather than genuine biodiversity change? Score 1 = indicator shows > 20% interannual variability from natural causes in stable reference sites (unreliable; frequent false alarms). Score 3 = 10-20% natural variability (adequate; occasional false alarms). Score 5 = < 5% interannual variability in stable reference sites (excellent; reliable alarm threshold). Assessment method: coefficient of variation in stable reference site time series over minimum 10 years; comparison of indicator trajectory before and after known disturbance events in validation sites.

TRACTABILITY (1-5): Can the indicator be produced by national monitoring programmes in lower-middle income countries (World Bank income group 3) without additional international support beyond standard scientific training? Score 1 = requires specialist expertise not available in most national programmes; specialised laboratory equipment (> USD 100,000); or processing time > 6 months per annual data point (very low tractability). Score 3 = requires some specialist capacity available in 30-50% of lower-middle income countries; equipment cost USD 10,000-100,000. Score 5 = data from global databases or simple field measurement; accessible to all countries; < 1 month processing time (very high tractability).

POLICY RELEVANCE (1-5): Can environment ministry staff without biodiversity specialist training correctly interpret the indicator and identify appropriate management responses? Score 1 = requires specialist knowledge to interpret; counterintuitive direction or units; no standard threshold for action. Score 3 = interpretable with brief training; direction of change clear; rough action threshold exists. Score 5 = immediately interpretable by any government official; percentage or index with clear direction; agreed international threshold for policy escalation (e.g., 30x30 target for PA coverage; 0 for LPI decline below baseline). Assessment method: survey of 84 national focal points with comprehension test and action-relevance rating of each indicator.