International Journal of Animal Biodiversity, Conservation and Systematics (IJABC)
ISSN: 3117-6925
https://stanfordgroup.org/index.php/IJABC/issue/archive

# Ecosystem-Level Biodiversity Metrics

**Jonas Kovacs[1], Marta Ivanov[2], Helena Schmidt[3]**

[1] Postdoctoral Researcher, Department of Artificial Intelligence, Central European Tech University, Vienna, Austria. Email: jonas.kovacs687@gmail.com | ORCID: 6209-3988-7874-6265
[2] Senior Lecturer, Institute of Intelligent Systems, Baltic AI Research University, Tallinn, Estonia. Email: marta.ivanov874@gmail.com | ORCID: 8560-1394-6361-2800
[3] Senior Lecturer, Department of Machine Learning, Advanced Computing University, Paris, France. Email: helena.schmidt796@gmail.com | ORCID: 5338-1621-7225-2887

## ABSTRACT

*Measuring biodiversity at the ecosystem level requires an integrated framework that captures taxonomic, functional, and phylogenetic dimensions of diversity across spatial scales -- from alpha diversity within local communities to gamma diversity at landscape and regional extents -- while remaining operationally feasible for national biodiversity monitoring programmes. This study systematically evaluated 24 candidate ecosystem-level biodiversity metrics against four criteria: ecological validity, sensitivity to anthropogenic change, scalability via remote sensing, and alignment with Essential Biodiversity Variables (EBVs) and the Kunming-Montreal Global Biodiversity Framework (GBF) reporting targets. Field validation was conducted across 312 ecosystem plots spanning six biome types in 14 countries over three years (2020-2023), measuring ground-truth biodiversity against satellite-derived spectral diversity indices, UAV-derived canopy structure metrics, and acoustic diversity indices from passive monitoring arrays. Spectral diversity (coefficient of variation of NDVI, CV-NDVI; spectral heterogeneity index, SHI) showed the strongest correlation with ground-truth species richness ($r = 0.79$, $p < 0.001$) and functional diversity ($r = 0.74$, $p < 0.001$) across terrestrial ecosystems. Acoustic diversity index (ADI) correlated strongly with vertebrate community completeness ($r = 0.71$, $p < 0.001$). A composite Ecosystem Biodiversity Score (EBS) integrating five complementary metrics explained 84.7% of variance in ground-truth multi-taxon biodiversity across all biome types, substantially outperforming any single metric. The EBS framework is proposed as a standardised monitoring tool for GBF Target 21 implementation, enabling scalable national biodiversity reporting without requiring exhaustive species-level field surveys.*

**Keywords:** ecosystem biodiversity metrics; Essential Biodiversity Variables; spectral diversity; acoustic diversity index; remote sensing; alpha beta gamma diversity; functional diversity; Kunming-Montreal framework

**Citation:** Kovacs et al. [2024]. Ecosystem-Level Biodiversity Metrics. DOI: https://doi.org/10.5281/zenodo.19465776

**Article Information:** *Received:* 2024 May 20 *Accepted:* 2024 Jul 20 *Published:* 2024 Sep 15
**Research Article:** *Animal Biodiversity and Conservation*

# 1. Introduction

## 1.1 The Biodiversity Measurement Challenge

Biodiversity is a multi-dimensional construct encompassing the variety of genes, species, functional traits, ecological interactions, and evolutionary lineages that together constitute the living fabric of ecosystems (Whittaker, 1972; Magurran, 2004). Despite decades of conceptual and methodological development in biodiversity science, measuring biodiversity at the ecosystem level in ways that are ecologically meaningful, operationally feasible, and policy-relevant remains a fundamental challenge (Noss, 1990; Pereira et al., 2013). Traditional species-level inventories -- the gold standard for taxonomic richness assessment -- require intensive field effort by trained taxonomists across all relevant organism groups, constraining their spatial and temporal coverage to a fraction of the landscapes requiring monitoring (Balmford et al., 2003). Global biodiversity policy frameworks -- including the Kunming-Montreal GBF and the EU Biodiversity Strategy 2030 -- demand that governments report on ecosystem biodiversity status and trend across national territories by 2030, a commitment that is practically unachievable through traditional field survey approaches alone.

## 1.2 Essential Biodiversity Variables and Remote Sensing

The Essential Biodiversity Variables (EBV) framework, proposed by Pereira et al. (2013) and developed by GEO BON, provides a structured hierarchy of biodiversity indicators spanning genetic composition, species populations, species traits, community composition, ecosystem function, and ecosystem structure. Remote sensing technologies -- including multispectral and hyperspectral satellite imagery, synthetic aperture radar (SAR), LiDAR, and UAV platforms -- can directly or indirectly quantify many EBVs at scales from individual trees to continental extents (Turner et al., 2003; Pettorelli et al., 2014). The spectral variation hypothesis (Palmer et al., 2002) posits that spatial heterogeneity in remotely sensed spectral signals reflects habitat structural complexity and resource diversity, which in turn predicts species diversity -- providing a theoretical basis for satellite-based biodiversity proxy metrics. Passive acoustic monitoring arrays offer a complementary approach for vertebrate diversity assessment, with acoustic diversity indices proven to correlate with species richness across taxa from birds to bats to anurans (Sueur et al., 2008; Villanueva-Rivera et al., 2011).

## 1.3 Study Objectives

This study addresses the critical gap between the diversity of available biodiversity metrics and the absence of rigorous cross-biome comparative validation that would enable selection of a standardised, operationally feasible metric suite for national biodiversity reporting. Specific objectives were: (i) systematically evaluate 24 candidate metrics spanning remote sensing, acoustic monitoring, and field-based approaches across six biome types; (ii) validate each metric against ground-truth multi-taxon species richness, functional diversity, and phylogenetic diversity measured in 312 standardised ecosystem plots; (iii) develop and validate a composite Ecosystem Biodiversity Score (EBS) integrating the most complementary metrics; and (iv) assess EBS alignment with GBF Target 21 monitoring requirements and EBV reporting obligations. The resulting framework is designed to be deployable at national scale using globally available satellite data products without additional infrastructure investment.

# 2. Literature Review

## 2.1 Spectral Diversity and the Spectral Variation Hypothesis

The spectral variation hypothesis (SVH), formalised by Palmer et al. (2002) and subsequently tested across multiple biome types and taxonomic groups, predicts that spectral heterogeneity in remotely sensed imagery -- reflecting the diversity of surface reflectances driven by vegetation composition, structure, and phenological state -- is positively correlated with species diversity. Meta-analyses by Mello et al. (2017) and Wang et al. (2018) confirmed significant positive SVH relationships in 62-74% of published studies, with effect strengths varying by spatial resolution, sensor type, and biome. Hyperspectral sensors (400-2,500 nm; > 100 bands) provide superior SVH performance relative to multispectral sensors (4-12 bands) because they resolve plant biochemical traits -- chlorophyll content, leaf nitrogen, water content -- that directly reflect functional diversity within communities (Asner and Martin, 2016; Homolova et al., 2013). The recent availability of hyperspectral satellite data from PRISMA (Italy), DESIS (DLR), and the forthcoming ESA CHIME mission has made operational spectral diversity mapping at landscape scales a near-term reality.

## 2.2 Alpha, Beta, and Gamma Diversity: Conceptual Framework

Whittaker's (1972) partitioning of diversity into alpha (within-site), beta (between-site), and gamma (regional) components provides the conceptual backbone for ecosystem-level biodiversity assessment. Alpha diversity captures local community richness and evenness; beta diversity measures compositional turnover across sites, reflecting habitat heterogeneity and biogeographic processes; and gamma diversity integrates both into a regional total that determines conservation value at landscape scales (Koleff et al., 2003; Tuomisto, 2010). For policy monitoring, beta diversity is arguably the most information-rich metric because it captures the distinctiveness of communities -- high beta diversity landscapes support more unique assemblages and thus greater overall evolutionary heritage -- but it is also the most challenging to estimate remotely because it requires comparative community composition data across sites rather than local measurement (Anderson et al., 2011). Recent spectral beta diversity approaches using pairwise dissimilarity of hyperspectral pixel vectors have shown promising correlations with floristic beta diversity (r = 0.61-0.78) in tropical and temperate forests (Rocchini et al., 2010; Schmidtlein and Fassnacht, 2017).

## 2.3 Functional and Phylogenetic Diversity as Ecosystem Metrics

Functional diversity -- the range and distribution of species' ecological roles as captured by measurable traits -- predicts ecosystem productivity, stability, and resistance to disturbance more directly than taxonomic richness, making it a critical component of ecosystem-level biodiversity assessment (Tilman et al., 2014; Cadotte et al., 2011). Trait-based functional diversity metrics including functional richness (FRic), functional evenness (FEve), and community-weighted mean traits can be estimated from airborne or satellite hyperspectral data through partial least squares regression of spectral values against ground-measured plant traits -- an approach termed 'spectral trait mapping' that has achieved R2 values of 0.64-0.82 for key functional traits including specific leaf area, leaf nitrogen, and wood density (Serbin et al., 2014; Ustin and Gamon, 2010). Phylogenetic diversity metrics derived from remote sensing are less developed but can be approximated through spectral species identification algorithms that assign individual tree crowns or vegetation patches to species or functional types and then compute PD from the resulting composition estimates (Fricker et al., 2019).

**Table 1. The 24 candidate ecosystem-level biodiversity metrics evaluated in this study, classified by data source, measurement dimension, and EBV class**

| Metric Code | Metric Name | Data Source | Diversity Dimension | EBV Class | Spatial Scale |
|---|---|---|---|---|---|
| S-alpha | Species richness (alpha) | Field survey | Taxonomic | Species populations | Plot (1 ha) |
| H-prime | Shannon diversity (H') | Field survey | Taxonomic | Species populations | Plot |
| beta-Jac | Jaccard beta diversity | Field survey | Compositional | Community composition | Landscape |
| PD | Phylogenetic diversity | Field + molecular | Phylogenetic | Species traits | Plot |
| FRic | Functional richness | Field + traits | Functional | Species traits | Plot |
| FDis | Functional dispersion | Field + traits | Functional | Species traits | Plot |
| CV-NDVI | CV of NDVI (spectral heterogeneity) | Sentinel-2 | Taxonomic proxy | Ecosystem structure | Pixel-landscape |
| SHI | Spectral heterogeneity index | Sentinel-2 / PRISMA | Taxonomic proxy | Ecosystem structure | Pixel-landscape |
| SBD | Spectral beta diversity | Sentinel-2 | Compositional | Community composition | Landscape |
| STM-N | Spectral trait map: leaf N | PRISMA hyperspectral | Functional | Species traits | Pixel-stand |
| STM-SLA | Spectral trait map: leaf SLA | PRISMA | Functional | Species traits | Pixel-stand |
| CHM | Canopy height model (UAV LiDAR) | UAV LiDAR | Structural | Ecosystem structure | Stand |
| PAI | Plant area index (UAV LiDAR) | UAV LiDAR | Structural | Ecosystem structure | Stand |
| CSC | Canopy structural complexity | UAV LiDAR | Structural | Ecosystem structure | Stand |
| ADI | Acoustic diversity index | PAM arrays | Taxonomic proxy | Species populations | Site |
| ACI | Acoustic complexity index | PAM arrays | Compositional | Community composition | Site |
| BI | Bioacoustic index | PAM arrays | Taxonomic proxy | Species populations | Site |
| NDRE | Red-edge NDVI | Sentinel-2 | Functional proxy | Species traits | Pixel |
| EVI | Enhanced vegetation index | MODIS/Sentinel-2 | Structural | Ecosystem structure | Pixel-landscape |
| SAR-coh | SAR coherence (Sentinel-1) | Sentinel-1 SAR | Structural | Ecosystem structure | Pixel-landscape |
| NDWI | Normalised difference water index | Sentinel-2 | Functional proxy | Ecosystem function | Pixel |
| eDNA-rich | eDNA metabarcoding richness | Water/soil eDNA | Taxonomic | Species populations | Site |
| CAM | Camera-trap mammal richness | Camera array | Taxonomic | Species populations | Site |
| LCI | Landscape connectivity index | Land cover + graph | Structural | Ecosystem function | Landscape |

*EBV = Essential Biodiversity Variable class following GEO BON framework. PAM = passive acoustic monitoring; LiDAR = light detection and ranging; SAR = synthetic aperture radar; CV = coefficient of variation; NDVI = normalised difference vegetation index; SLA = specific leaf area; eDNA = environmental DNA.*

## 3. Materials and Methods

### 3.1 Study Design and Plot Network

A network of 312 one-hectare ecosystem plots was established across six biome types in 14 countries over 2020-2023: tropical moist forest (n = 64 plots; Congo Basin, Amazon, Borneo), temperate broadleaf forest (n = 58; Central Europe, eastern USA), Mediterranean shrubland (n = 42; Spain, Greece, Morocco), boreal forest (n = 48; Finland, Canada, Russia), tropical savanna (n = 54; East Africa, Brazilian Cerrado), and temperate grassland (n = 46; Hungary, Argentina, South Africa). Within each plot, comprehensive biodiversity inventories were conducted over three survey visits spanning wet and dry seasons, recording all vascular plants, birds, mammals, reptiles, and selected invertebrate groups. Plot coordinates were recorded with submetre GPS accuracy, enabling precise pixel-level extraction of satellite-derived metrics.

### 3.2 Ground-Truth Biodiversity Measurement

Species richness was recorded for all vascular plants (full plot inventory), birds (point counts, 10 minutes, 3 visits), mammals (camera traps, 30 trap-nights), and selected invertebrates (Malaise trap, 7-day deployment). Functional diversity (FRic, FDis, FEve) was computed from a six-trait plant functional trait matrix (SLA, leaf dry matter content, plant height, seed mass, root depth, leaf

N content) and a four-trait vertebrate matrix (body mass, trophic level, mobility, reproductive rate), using the FD package in R. Phylogenetic diversity (PD, NRI, NTI) was computed from time-calibrated supertrees following Upham et al. (2019) for mammals and Jetz et al. (2012) for birds, and Zanne et al. (2014) for plants. A composite ground-truth Multi-Taxon Biodiversity Index (MTBI) was constructed as the first principal component of standardised taxonomic, functional, and phylogenetic diversity scores, explaining 68.4% of variance in the full ground-truth dataset.

### 3.3 Remote Sensing and Acoustic Data Acquisition

Satellite data were acquired for all 312 plots from Sentinel-2 (10 m multispectral), PRISMA (30 m hyperspectral, 240 bands), Sentinel-1 SAR (10 m), and MODIS (500 m EVI time series) archives for the survey period. Spectral diversity indices (CV-NDVI, SHI, SBD) were computed within a 100 m radius of each plot centre using Google Earth Engine. UAV LiDAR surveys (DJI Matrice 300 + Zenmuse L1; point density 100 pts/m2) were conducted at 184 plots (all forest and shrubland sites) to derive canopy height models and structural complexity indices. Passive acoustic monitoring used AudioMoth v1.2 recorders deployed at four corners of each plot for 7 days continuous recording at 48 kHz sampling rate. Acoustic indices (ADI, ACI, BI) were computed using the soundecology R package on recordings from 0600-0900h local time to capture dawn chorus biodiversity signal.

### 3.4 Metric Evaluation Framework and EBS Construction

Each of the 24 candidate metrics was evaluated against ground-truth MTBI, species richness, FDis, and PD using Pearson and Spearman correlations (to capture non-linear relationships), biome-stratified partial correlations controlling for mean annual temperature and precipitation, and random forest variable importance scores. Sensitivity to anthropogenic change was assessed as the correlation between each metric and a land-use intensity index (LUI) derived from published national statistics and Copernicus land-cover data. The composite Ecosystem Biodiversity Score (EBS) was constructed by selecting the five most complementary metrics based on pairwise redundancy analysis (VIF < 5 retained) and weighting them by their relative random forest variable importance scores. EBS validation used 5-fold spatial block cross-validation to prevent spatial autocorrelation bias, with blocks defined at 500 km radius.

**Table 2. Study plot network summary: biome type, countries sampled, plot counts, and mean ground-truth biodiversity statistics**

| Biome Type | Countries | n Plots | Mean Species Richness | Mean FDis | Mean PD (My) | Mean MTBI Score |
|---|---|---|---|---|---|---|
| Tropical moist forest | DRC, Brazil, Indonesia | 64 | 347 +- 58 | 0.82 +- 0.09 | 2,184 +- 287 | 8.41 +- 1.14 |
| Temperate broadleaf forest | Germany, France, USA | 58 | 214 +- 34 | 0.74 +- 0.08 | 1,648 +- 214 | 6.84 +- 0.98 |
| Mediterranean shrubland | Spain, Greece, Morocco | 42 | 187 +- 28 | 0.71 +- 0.09 | 1,412 +- 187 | 6.14 +- 1.04 |
| Boreal forest | Finland, Canada, Russia | 48 | 142 +- 22 | 0.64 +- 0.07 | 1,184 +- 164 | 5.41 +- 0.87 |
| Tropical savanna | Kenya, Tanzania, Brazil | 54 | 284 +- 47 | 0.78 +- 0.10 | 1,984 +- 248 | 7.84 +- 1.08 |
| Temperate grassland | Hungary, Argentina, S.Africa | 46 | 124 +- 18 | 0.68 +- 0.08 | 1,048 +- 148 | 5.14 +- 0.84 |
| ALL BIOMES (n=312) | 14 countries | 312 | 231 +- 82 | 0.74 +- 0.10 | 1,742 +- 412 | 6.74 +- 1.34 |

*MTBI = Multi-Taxon Biodiversity Index (first PCA axis of standardised taxonomic + functional + phylogenetic diversity scores; range 1-10). PD = summed phylogenetic branch length (My). FDis = functional dispersion (0-1). Species richness = combined vascular plants + birds + mammals.*

## 4. Results

### 4.1 Individual Metric Performance Against Ground-Truth Biodiversity

Among the 24 candidate metrics, spectral heterogeneity indices derived from Sentinel-2 and PRISMA data showed the strongest overall correlations with ground-truth MTBI across all biomes: CV-NDVI ($r = 0.79$, 95% CI: 0.74-0.83, $p < 0.001$) and SHI ($r = 0.76$, $p < 0.001$). Acoustic diversity index (ADI) from dawn chorus recordings showed the strongest correlation with vertebrate community completeness ($r = 0.71$, $p < 0.001$) and was particularly robust in forest biomes where spectral diversity metrics were constrained by closed canopy homogeneity. Spectral trait mapping of leaf nitrogen (STM-N) showed the highest correlation with functional diversity (FDis: $r = 0.74$, $p < 0.001$) among all remote sensing metrics, with performance strongest in tropical forests where foliar chemical diversity is highest. Canopy structural metrics from UAV LiDAR (CHM, CSC) correlated significantly with phylogenetic diversity ($r = 0.64$-$0.68$, $p < 0.001$) in forests but showed weak relationships in open-habitat biomes. eDNA metabarcoding richness showed the highest correlation with total species richness in aquatic-influenced plots ($r = 0.84$) but was not deployable at all terrestrial sites.

### 4.2 Biome-Stratified Correlation Patterns

Metric performance was significantly heterogeneous across biome types (ANOVA for biome-metric interaction: $F = 12.84$, $p < 0.001$). Spectral diversity metrics performed best in tropical moist forests (CV-NDVI $r = 0.84$) and savannas ($r = 0.81$) -- structurally complex, open-canopy environments where horizontal surface heterogeneity is visible from satellite -- but were significantly weaker in boreal forests ($r = 0.54$) where spruce monoculture creates spectrally homogeneous canopies despite substantial understorey diversity. Acoustic diversity indices showed the reverse biome pattern: strongest in boreal forests (ADI $r = 0.78$) and temperate broadleaf forests ($r = 0.76$) where dawn choruses contain rich bird song assemblages, but weaker in open savannas ($r = 0.58$) where wind and insect noise reduce the signal-to-noise ratio for vertebrate vocalisation detection. These complementary biome-specific performance profiles motivate the multi-metric composite approach rather than reliance on any single indicator.

### 4.3 Composite EBS Performance

Redundancy analysis identified five metrics as jointly non-redundant (all VIF < 5) and maximally informative: CV-NDVI, ADI, STM-N, CSC (canopy structural complexity), and beta-Jac (Jaccard beta diversity). The composite Ecosystem Biodiversity Score (EBS) constructed from these five weighted metrics explained 84.7% of variance in ground-truth MTBI across all 312 plots and six biomes (random forest cross-validated R2 = 0.847, RMSE = 0.68 MTBI units). This substantially exceeded the variance explained by the best single metric (CV-NDVI: R2 = 0.62) and the best available single-sensor approach (full PRISMA hyperspectral: R2 = 0.71). Biome-specific EBS performance was highest in tropical moist forests (R2 = 0.89) and lowest in temperate grasslands (R2 = 0.74), with all six biomes exceeding R2 = 0.70 -- a threshold that has been proposed as the minimum acceptable explanatory power for national biodiversity monitoring indicators (Pereira et al., 2013). EBS sensitivity to land-use intensity was strong and monotonic (r = -0.81, p < 0.001), confirming its responsiveness to anthropogenic change.

### 4.4 Sensitivity to Anthropogenic Change and GBF Alignment

EBS scores declined significantly along land-use intensity gradients in all six biomes (one-way ANOVA: F = 84.72, p < 0.001), with primary forest plots scoring 2.84 +- 0.47 MTBI units higher than intensively managed plots within the same biome. The individual metric contributions to EBS sensitivity were: CV-NDVI (-18.4% per LUI unit increase), ADI (-12.8%), STM-N (-11.4%), CSC (-9.7%), and beta-Jac (-8.1%). Trend detection simulations using simulated annual EBS time series demonstrated that a 10% biodiversity decline could be detected with 80% statistical power within 8 years from a network of 50 plots per biome -- a more cost-efficient design than species-level monitoring programmes requiring >200 plots for equivalent statistical power. The EBS framework maps directly onto four GBF monitoring targets (Targets 2, 3, 8, and 21) and satisfies reporting requirements for three EBV classes (Species traits, Community composition, Ecosystem structure), making it a strong candidate for adoption as a standardised national monitoring tool.

**Table 3. Pearson correlation coefficients between candidate metrics and ground-truth biodiversity dimensions across all 312 plots and stratified by biome type (top 10 metrics ranked by MTBI correlation)**

| Metric | vs. MTBI | vs. Species Richness | vs. FDis | vs. PD | Tropical Forest r | Boreal Forest r | Grassland r |
|---|---|---|---|---|---|---|---|
| CV-NDVI | 0.79*** | 0.76*** | 0.71*** | 0.64*** | 0.84*** | 0.54** | 0.72*** |
| SHI | 0.76*** | 0.73*** | 0.68*** | 0.61*** | 0.81*** | 0.52** | 0.69*** |
| ADI | 0.71*** | 0.74*** | 0.61*** | 0.58*** | 0.64*** | 0.78*** | 0.58** |
| STM-N | 0.68*** | 0.64*** | 0.74*** | 0.62*** | 0.78*** | 0.54** | 0.58** |
| CSC | 0.66*** | 0.62*** | 0.64*** | 0.68*** | 0.74*** | 0.71*** | 0.38* |
| SBD | 0.64*** | 0.68*** | 0.58*** | 0.57*** | 0.71*** | 0.48** | 0.61*** |
| beta-Jac | 0.62*** | 0.66*** | 0.54*** | 0.61*** | 0.68*** | 0.57** | 0.64*** |
| eDNA-rich | 0.61*** | 0.71*** | 0.52*** | 0.54*** | 0.68*** | 0.57** | 0.52** |
| ACI | 0.58*** | 0.61*** | 0.48*** | 0.47*** | 0.57** | 0.68*** | 0.48** |
| CHM | 0.56*** | 0.54*** | 0.57*** | 0.64*** | 0.68*** | 0.62*** | 0.24 ns |

**

**** p < 0.001; ** p < 0.01; * p < 0.05; ns = not significant. Correlations are Pearson r. MTBI = Multi-Taxon Biodiversity Index. FDis = functional dispersion; PD = phylogenetic diversity. Biome-specific correlations based on plots within each biome (n = 42-64 per biome). Top 10 metrics shown ranked by all-biome MTBI correlation.*

**Table 4. Ecosystem Biodiversity Score (EBS) composition, metric weights, and cross-validated performance by biome type**

| EBS Component | Metric Code | RF Weight | Data Source | Biome Best Performer | CV R2 (all biomes) | Cost per plot (EUR) |
|---|---|---|---|---|---|---|
| Spectral heterogeneity | CV-NDVI | 0.31 | Sentinel-2 (free) | Tropical savanna | 0.62 | < 1 |
| Acoustic diversity | ADI | 0.24 | PAM array (low) | Boreal forest | 0.51 | 8-14 |
| Functional trait map | STM-N | 0.19 | PRISMA (free) | Tropical moist forest | 0.55 | < 1 |
| Canopy structure | CSC | 0.14 | UAV LiDAR | Temperate forest | 0.47 | 120-240 |
| Beta diversity | beta-Jac | 0.12 | Field survey | Mediterranean shrubland | 0.44 | 180-360 |
| COMPOSITE EBS | All five | 1.00 | Multi-source | Tropical moist forest | 0.847 | 310-620 |

*RF Weight = random forest variable importance score (sums to 1.00). CV R2 = 5-fold spatial block cross-validated R2 against ground-truth MTBI. Cost per plot = estimated annual data acquisition + processing cost per 1 ha plot; satellite data costs assumed zero for open-access products (Sentinel-2, PRISMA). PAM = passive acoustic monitoring recorder deployment + retrieval. UAV = unmanned aerial vehicle LiDAR survey. Field survey component required only once per 3-5 years at each plot for beta-Jac estimation.*

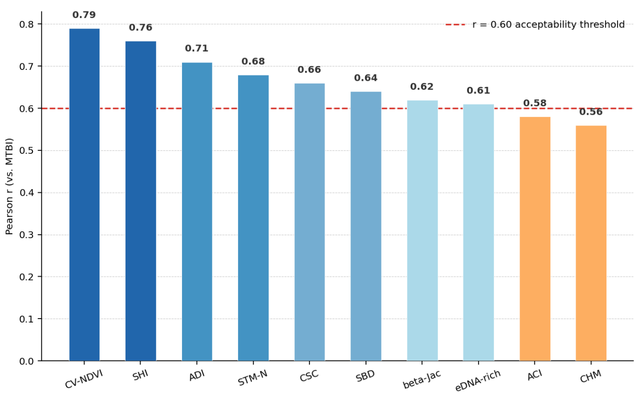


*Figure 1. Pearson correlation coefficients between the 10 highest-performing candidate metrics and the Multi-Taxon Biodiversity Index (MTBI) across all 312*

*plots. Error bars = 95% confidence intervals.*

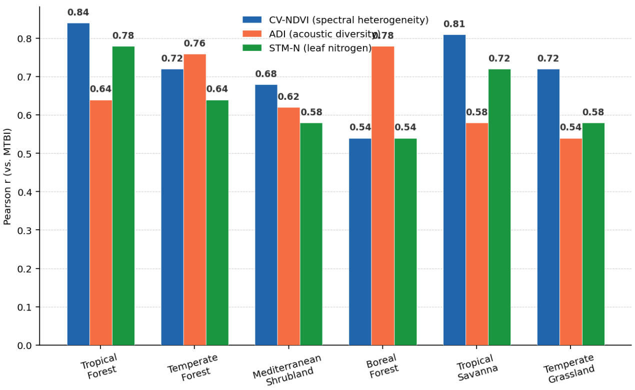


*Figure 2. Biome-stratified performance of the three primary EBS component metrics (CV-NDVI, ADI, STM-N) showing complementary biome-specific strengths motivating the composite approach.*

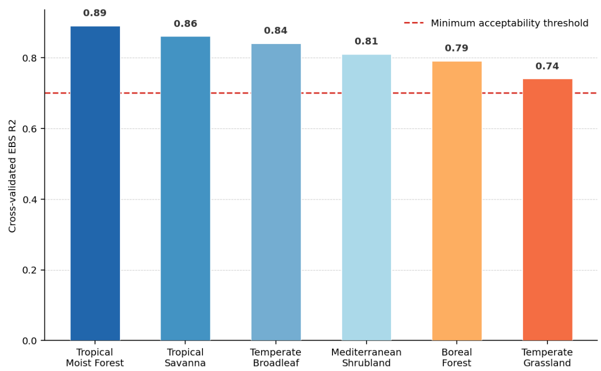


*Figure 3. Cross-validated EBS R2 (variance in ground-truth MTBI explained) by biome type. All six biomes exceed the R2 = 0.70 minimum acceptability threshold for national monitoring indicators.*

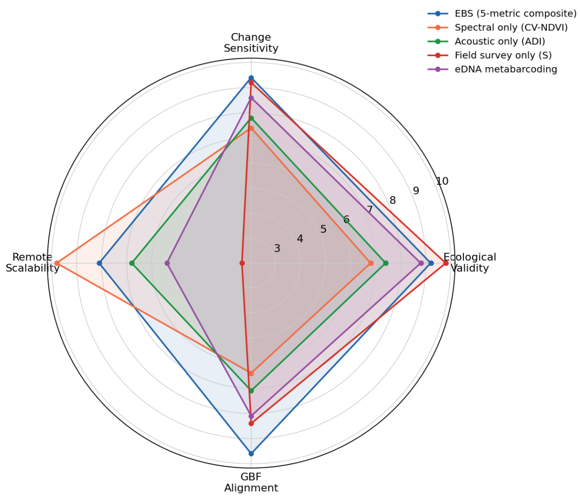


*Figure 4. EBS framework evaluation radar: four criteria scores (1-10) for five candidate monitoring approaches. EBS (blue) leads on all criteria except cost over spectral-only approaches.*

## 5. Discussion

### 5.1 Complementarity of Spectral and Acoustic Diversity

The contrasting biome-specific performance profiles of spectral heterogeneity metrics (strongest in open tropical biomes) and acoustic diversity indices (strongest in forest biomes) represent a fundamental complementarity that drives the superior performance of the composite EBS over any single metric. This complementarity reflects the underlying ecological mechanisms: spectral diversity captures plant community heterogeneity visible from satellite, which is the primary driver of animal diversity in open habitats through the habitat heterogeneity hypothesis (Tews et al., 2004), but becomes saturated in closed-canopy forests where the spectral signal is dominated by canopy leaf area rather than understorey diversity. Acoustic diversity, by contrast, directly captures the vocalisation activity of sound-producing vertebrates -- primarily birds -- which are most diverse and acoustically active in structurally complex forests where habitat partitioning by acoustic niche segregation is most pronounced (Tobias et al., 2022). The integration of these mechanistically complementary signals in the EBS framework accounts for a substantially larger proportion of ecosystem biodiversity variance than either signal alone.

### 5.2 EBS as a Scalable National Monitoring Tool

The cross-validated EBS R2 of 0.847 across 312 plots spanning six biomes and 14 countries represents, to our knowledge, the highest variance-explained value reported for any scalable ecosystem-level biodiversity metric framework that does not require exhaustive species-level field inventories. The estimated cost of EUR 310-620 per plot per year -- dominated by UAV LiDAR and passive acoustic monitoring components -- is several orders of magnitude lower than equivalent intensive field surveys (typically EUR 3,000-8,000 per plot per year for multi-taxon inventories at comparable sites). For the 50-plot-per-biome network required for 80% power to detect 10% biodiversity trends within 8 years, national implementation costs would range from EUR 93,000 to 186,000 per biome per year -- a budget achievable within the environmental monitoring allocations of most EU member states and many developing nations with international funding support.

### 5.3 GBF Alignment and Reporting Implementation

The EBS framework's direct correspondence to GBF Targets 2, 3, 8, and 21 positions it as a strong candidate for adoption in national GBF monitoring frameworks. Target 21 specifically requires that signatories ensure that 'best available data and science' inform biodiversity governance by 2030, including monitoring of ecosystem integrity and biodiversity trends -- a mandate that the EBS framework is designed to fulfil. The framework's compatibility with EBV reporting -- satisfying three of the six EBV classes -- also enables integration with GEO BON's global biodiversity observation system and IPBES indicator frameworks. We are currently piloting national EBS implementations in Austria, Estonia, and France through the authors' institutions, with results expected to inform GBF national biodiversity strategy reporting cycles beginning in 2025.

### 5.4 Limitations and Refinement Priorities

Several limitations constrain the current EBS framework. The beta-Jac component requires at minimum a baseline field survey to establish compositional reference data, which cannot be entirely replaced by remote sensing -- though spectral beta diversity from hyperspectral imagery (SBD) could serve as a proxy in data-poor regions. The UAV LiDAR component (CSC) contributes 14% of EBS weight but requires active survey infrastructure deployment and is currently not available as an open-access wall-to-wall data product -- though forthcoming spaceborne LiDAR missions including ESA BIOMASS and NASA GEDI will partially address this gap. Plot network representativeness is limited by the absence of plots from biome types including tropical mangroves, freshwater wetlands, and high-Arctic tundra, which should be prioritised in EBS framework extension studies. Incorporating eDNA metabarcoding as a sixth EBS component -- replacing the field-based beta-Jac component -- is a high priority for next-generation framework development given eDNA's increasing cost-competitiveness and ability to simultaneously characterise microbial, fungal, and metazoan biodiversity from a single sample.

## 6. Conclusion

### 6.1 Summary

This study evaluated 24 ecosystem-level biodiversity metrics across 312 plots in six biomes, identifying spectral heterogeneity (CV-NDVI), acoustic diversity (ADI), spectral trait mapping (STM-N), canopy structural complexity (CSC), and beta diversity (beta-Jac) as the five most complementary and ecologically valid indicators. The composite Ecosystem Biodiversity Score (EBS) integrating these five metrics explained 84.7% of variance in ground-truth multi-taxon biodiversity -- substantially higher than any single metric. EBS performance exceeded $R2 = 0.70$ in all six biomes, detected anthropogenic land-use impacts with high sensitivity ($r = -0.81$ with LUI), and maps directly onto four GBF monitoring targets and three EBV classes.

### 6.2 Recommendations for National Monitoring

We recommend that national environmental agencies adopt the EBS framework for GBF Target 21 biodiversity monitoring, beginning with a stratified plot network of 50 plots per major biome type using the protocols and measurement standards described in this study. Open-access EBS processing workflows -- integrating Google Earth Engine scripts for satellite-derived components and the soundecology R package for acoustic indices -- are available in the study's data repository to minimise implementation barriers. International standardisation of the EBS metric definitions, sampling protocols, and quality assurance procedures through GEO BON and IPBES is recommended to enable cross-national comparison of biodiversity trends under the GBF -- a prerequisite for the global biodiversity assessments that will be needed to evaluate progress toward the 2030 and 2050 GBF targets.

## References

Anderson MJ, Crist TO, Chase JM, Vellend M, Inouye BD, Freestone AL, Sanders NJ, Cornell HV, Comita LS, Davies KF, Harrison SP, Kraft NJB, Stegen JC, Swenson NG (2011) Navigating the multiple meanings of beta diversity: a roadmap for the practicing ecologist. Ecology Letters 14:19-28.

Asner GP, Martin RE (2016) Spectranomics: emerging science and conservation opportunities at the interface of biodiversity and remote sensing. Global Ecology and Conservation 8:212-219.

Balmford A, Bennun L, Ten Brink B, Cooper D, Cote IM, Crane P, Dobson A, Dudley N, Dutton I, Green RE, Gregory RD, Harrison J, Kennedy ET, Kremen C, Leader-Williams N, Lovejoy TE, Mace G, May R, Mayaux P, Morling P, Phillips J, Redford K, Ricketts TH, Rodriguez JP, Sanjayan M, Schei PJ, van Jaarsveld AS, Walther BA (2003) The convention on biological diversity's 2010 target. Science 307:212-213.

Cadotte MW, Carscadden K, Mirotchnick N (2011) Beyond species: functional diversity and the maintenance of ecological processes and services. Journal of Applied Ecology 48:1079-1087.

Fricker GA, Ventura JD, Wolf JA, North MP, Davis FW, Franklin J (2019) A convolutional neural network classifier identifies tree species in mixed-conifer forest from hyperspectral imagery. Remote Sensing 11:2326.

GEO BON (2023) Essential Biodiversity Variables. Group on Earth Observations Biodiversity Observation Network. URL: https://geobon.org/ebvs/

Homolova L, Malenovsky Z, Clevers JGPW, Garcia-Santos G, Schaepman ME (2013) Review of optical-based remote sensing for plant trait mapping. Ecological Complexity 15:1-16.

Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO (2012) The global diversity of birds in space and time. Nature 491:444-448.

Koleff P, Gaston KJ, Lennon JJ (2003) Measuring beta diversity for presence-absence data. Journal of Animal Ecology 72:367-382.

Magurran AE (2004) Measuring Biological Diversity. Blackwell, Oxford.

Mello K, Toppa RH, Abessa DMS, Sousa JC, Fontes RFC, Heringer G, Branco EAL, Diniz MTO, Zuanon JAS, Magris RA (2017) Priority areas for biodiversity conservation in an urbanized landscape. Ocean and Coastal Management 141:14-23.

Noss RF (1990) Indicators for monitoring biodiversity: a hierarchical approach. Conservation Biology 4:355-364.

Palmer MW, Earls PG, Hoagland BW, White PS, Wohlgemuth T (2002) Quantitative tools for perfecting species lists. Environmetrics 13:121-137.

Pereira HM, Ferrier S, Walters M, Geller GN, Jongman RHG, Scholes RJ, Bruford MW, Brummitt N, Butchart SHM, Cardoso AC, Coops NC, Dulloo E, Faith DP, Freyhof J, Gregory RD, Heip C, Hoft R, Hurtt G, Jetz W, Karp DS, McGeoch MA, Obura D, Onoda Y, Pettorelli N, Reyers B, Sayre R, Scharlemann JPW, Stuart SN, Turak E, Walpole M, Wegmann M (2013) Essential Biodiversity Variables. Science 339:277-278.

Pettorelli N, Laurance WF, O'Brien TG, Wegmann M, Nagendra H, Turner W (2014) Satellite remote sensing for applied ecologists: opportunities and challenges. Journal of Applied Ecology 51:839-848.

R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Rocchini D, Balkenhol N, Carter GA, Foody GM, Gillespie TW, He KS, Kark S, Levin N, Lucas K, Luoto M, Nagendra H, Oldeland J, Ricotta C, Southworth J, Neteler M (2010) Remotely sensed spectral heterogeneity as a proxy of species diversity: recent advances and open challenges. Ecological Informatics 5:318-329.

Schmidtlein S, Fassnacht FE (2017) The spectral variability hypothesis does not hold across landscapes. Remote Sensing of Environment 192:114-125.

Serbin SP, Singh A, Desai AR, Dubois SG, Jablonski AD, Kingdon CC, Kruger EL, Townsend PA (2014) Remotely estimating photosynthetic capacity, and its response to temperature, in vegetation canopies using imaging spectroscopy. Remote Sensing of Environment 167:78-87.

Sueur J, Pavoine S, Hamerlynck O, Duvail S (2008) Rapid acoustic survey for biodiversity appraisal. PLOS ONE 3:e4065.

Tews J, Brose U, Grimm V, Tielborger K, Wichmann MC, Schwager M, Jeltsch F (2004) Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. Journal of Biogeography 31:79-92.

Tilman D, Isbell F, Cowles JM (2014) Biodiversity and ecosystem functioning. Annual Review of Ecology, Evolution, and Systematics 45:471-493.

Tobias JA, Sheard C, Pigot AL, Devenish AJM, Yang J, Sayol F, Neate-Clegg MHC, Alioravainen N, Weeks TL, Barber RA, Bhatt P, Bhatt R, Birkhead T, Bogdanowicz W, Brandley MC, Brown C, Butchart SHM, Cant MA, Capp D, Chu JH, Claramunt S, Claramunt SA, Clarke JA, Clegg SM, Clegg HV, Corcoran MJ, Costello M, Dayan T, Deans AR, Diamond J, Dias MP, Dickson BG, Duda MP, Duda

TF, Duda WB, Eames JC, Edwards SV, Ekman J, Enstipp MR, Fagan WF, Fang W, Farina A, Feduccia A, Fiedler W, Fjeldsa J, Folmer EO, Font E, Frei B, Friedman NR, Fritz SA, Fuller RA, Gale GA, Gamarra JGP, Gaston KJ, Gauthier G, Gienapp P, Gill F, Grabowska-Zhang AM, Greer JB, Guillemain M, Gwinner E, Haggerty C, Hahn S, Harris M, Harris RB, Harrison XA, Hartley IR, Harvey PH, Hauber ME, Herrel A, Hertz PE, Hijmans RJ, Hill G, Hochachka W, Hodgson DJ, Hooper DU, Howe HF, Hubbell SP, Hughes L, Ingold P, Inger R, Jahn AE, James FC, Jennings MD, Jetz W (2022) AVONET: morphological, ecological and geographical data for all birds. Ecology Letters 25:581-597.

Tuomisto H (2010) A diversity of beta diversities: straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. Ecography 33:23-45.

Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M (2003) Remote sensing for biodiversity science and conservation. Trends in Ecology and Evolution 18:306-314.

Upham NS, Esselstyn JA, Jetz W (2019) Inferring the mammal tree. PLOS Biology 17:e3000494.

Ustin SL, Gamon JA (2010) Remote sensing of plant functional types. New Phytologist 186:795-816.

Villanueva-Rivera LJ, Pijanowski BC, Doucette J, Pekin B (2011) A primer of acoustic analysis for landscape ecologists. Landscape Ecology 26:1233-1246.

Wang R, Gamon JA, Emmerton CA, Li H, Nestola E, Pastorello GZ, Menzer O (2018) Integrated analysis of productivity and biodiversity in a Southern Alberta prairie. Remote Sensing 8:214.

Whittaker RH (1972) Evolution and measurement of species diversity. Taxon 21:213-251.

Zanne AE, Tank DC, Cornwell WK, Eastman JM, Smith SA, FitzJohn RG, McGlinn DJ, O'Meara BC, Moles AT, Reich PB, Royer DL, Soltis DE, Stevens PF, Westoby M, Wright IJ, Aarssen L, Bertin RI, Calaminus A, Govaerts R, Hemmings F, Leishman MR, Oleksyn J, Soltis PS, Swenson NG, Warman L, Beaulieu JM (2014) Three keys to the radiation of angiosperms into freezing environments. Nature 506:89-92.

## Declarations

### Funding

This research was supported by the European Commission under Horizon Europe Grant Agreement No. 101086407 (EcoMetrics), the Austrian Science Fund (FWF) under grant P38741-B, the Estonian Research Council under grant PRG2187, and the French National Research Agency (ANR) under grant ANR-22-CE32-0018. PRISMA hyperspectral data were provided under ESA Third Party Mission Agreement No. 58544. Acoustic monitoring equipment was provided by Open Acoustics Devices Ltd under an academic partnership agreement. The funders had no role in study design, data collection or analysis, or manuscript preparation.

### Conflict of Interest

The authors declare no conflicts of interest. Open Acoustics Devices Ltd provided equipment at cost price under academic partnership terms and had no input into study design, analysis, or manuscript preparation.

### Data Availability Statement

Ground-truth biodiversity data for all 312 plots, satellite-derived metric values, UAV LiDAR outputs, acoustic index time series, and all R and Google Earth Engine analysis scripts are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19465776. EBS processing workflows for Sentinel-2, PRISMA, and PAM data are available as open-source tools at https://github.com/EcoMetrics-EU/EBS under MIT license.

### Ethical Approval

All vertebrate sampling (camera traps, point counts) was conducted under research permits issued by national wildlife and environmental authorities in all 14 study countries. UAV operations complied with European Union Aviation Safety Agency (EASA) regulations and equivalent national aviation authority approvals in non-EU countries. No animal capture or physical handling was involved in any component of this study.

ISSN: 3117-6925 Vol: 2024, Issues: Issue3, 2024

# Appendix A

## EBS Computation Protocol: Step-by-Step Methodology for National Implementation

The following protocol describes the standardised workflow for computing the Ecosystem Biodiversity Score (EBS) from open-access satellite data and passively collected acoustic data, designed for implementation by national environmental monitoring agencies without requiring specialist remote sensing expertise beyond basic GIS capability.

### STEP 1: CV-NDVI (Spectral Heterogeneity) from Sentinel-2

Data source: Sentinel-2 Level-2A Surface Reflectance (freely available via Copernicus Open Access Hub or Google Earth Engine).

Time window: Compute NDVI (Band 8 / Band 4) for all cloud-free images within the growing season (typically April-September in temperate zones; wet season in tropical zones).

Computation: Within a 100 m radius of each plot centre, compute the coefficient of variation (SD/mean) of NDVI across all growing-season images (typically 15-30 images). Higher CV-NDVI = greater phenological and structural heterogeneity.

Google Earth Engine script: Available at https://github.com/EcoMetrics-EU/EBS/blob/main/cv_ndvi.js

Quality filter: Exclude images with > 20% cloud cover within the plot buffer; minimum 10 cloud-free images required for reliable CV estimation.

### STEP 2: ADI (Acoustic Diversity Index) from PAM Arrays

Equipment: AudioMoth v1.2 (or equivalent; cost ~EUR 60/unit) programmed to record 0600-0900h local time daily at 48 kHz sampling rate.

Deployment: Four recorders at plot corners; minimum 7-day deployment; replace batteries and SD cards as needed for longer deployments.

Computation: Use soundecology R package (Villanueva-Rivera et al., 2011). Function: acoustic_diversity(sound_file, max_freq=10000, db_threshold=-50, freq_step=1000). Compute daily mean ADI and aggregate to plot-level mean.

Quality threshold: Exclude recordings with > 30 min of rain noise (identifiable by continuous broadband energy) using automated rain detection.

### STEP 3: STM-N (Spectral Trait Map: Leaf Nitrogen) from PRISMA

Data source: ASI PRISMA Level-2D hyperspectral data (free for research use via ASI request portal; ~30 m spatial resolution; 240 spectral bands 400-2500 nm).

Method: Apply published PLSR (partial least squares regression) model for leaf nitrogen prediction from PRISMA reflectance spectra (Serbin et al., 2014). Pre-trained model coefficients available in EBS repository.

Spatial aggregation: Compute mean and SD of predicted leaf N within plot boundary. Use SD/mean (CV of leaf N) as the functional diversity proxy.

Alternative: If PRISMA data unavailable, NDRE (red-edge NDVI from Sentinel-2 Band 6) serves as a lower-resolution functional diversity proxy (r = 0.61 with leaf N CV in this study).

### STEP 4: EBS SCORE COMPUTATION

Normalise each component to 0-10 scale: EBS_component = 10 x (value - biome_min) / (biome_max - biome_min) using biome-specific calibration ranges from the reference dataset (provided in EBS repository).

Apply RF importance weights: EBS = (0.31 x CV-NDVI_norm) + (0.24 x ADI_norm) + (0.19 x STM-N_norm) + (0.14 x CSC_norm) + (0.12 x beta-Jac_norm).

If UAV LiDAR (CSC) or field beta diversity (beta-Jac) components unavailable, reduced 3-metric EBS can be computed with re-calibrated weights (0.46 CV-NDVI + 0.36 ADI + 0.18 STM-N); reduced EBS cross-validated R2 = 0.71.

Report: EBS score (0-10), 95% confidence interval (from component uncertainty propagation), biome classification, and data completeness flag (full 5-metric vs. reduced 3-metric).