Local data matters: Improving biodiversity risk and impact assessment through a data quality focus

Robert M Goodsell[1,2*], Emma Granqvist[1*#], Christophe Christiaen[3], Fredrik Ronquist[1].

1. Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Box 50007, SE-104 05 Stockholm, Sweden

2. School of Applied Sciences, University of the West of England, Bristol, BS16 1QY, United Kingdom.

3. Smith School of Enterprise and the Environment, University of Oxford, School of Geography and the Environment, South Parks Road, Oxford, OX1 3QY, United Kingdom

*** These authors contributed equally**

 # Corresponding author: emma.granqvist@nrm.se

ORCID

Robert M Goodsell	0000-0002-3349-1876
Emma Granqvist	0000-0002-1513-1674
Christophe Christiaen	0000-0002-6030-5071
Fredrik Ronquist	0000-0002-3929-251X

Abstract

Widespread degradation of nature has increased pressure on corporations and financial institutions to assess and mitigate their biodiversity impact, however, collecting relevant local data can be costly. The increasing availability of biodiversity and Earth observation (EO) data suggest that impact can instead be assessed cost-effectively through extrapolation using existing data and sophisticated modeling. Indeed, a review of the datasets and tools currently used by corporations and financial institutions shows that extrapolation to local sites from global datasets, or using only proxies, is the dominant approach. Here, we test the reliability of such assessments by combining high resolution earth observation time series data with extensive biodiversity data from recent environmental DNA (eDNA) surveys of two countries with widely varying conditions, Sweden and Madagascar. We use machine learning in combination with high-quality biodiversity data to predict five essential biodiversity variables (EBVs) for local sites, using cross-validation to test prediction accuracy. The results show that reasonably accurate EBV predictions can be obtained for sites with some local data, but performance declines considerably when modelling summary measures at new sites. Moreover, the quality of predictions, both within sites and at new sites, is dependent on the EBV and country. To address the concerns over the reliability of model-based EBV assessments, we propose a biodiversity data hierarchy framework, which can be used by organizations to track stepwise improvements in the data sources underpinning their biodiversity impact assessments.

Keywords:
Biodiversity impact assessment and reporting; eDNA; Earth observation; Machine learning; Sustainable finance; Essential biodiversity variables;

Highlights:

- Current biodiversity impact assessment and reporting is heavily leaning on freely available global biodiversity datasets such as the IUCN Red List and the World Database on Key Biodiversity Areas.
- Increasing numbers of large high-quality datasets on biodiversity are being made accessible, for example environmental DNA data and Earth Observation data, strengthening the power of predictive modelling.
- Nevertheless, analyses of example datasets show that predictive modelling without any local data at key sites of interest decreases performance significantly, increasing the risk of uninformative or misleading biodiversity assessments.
- We propose a data quality hierarchy framework, which can be used to highlight these issues and successively improve the reliability of biodiversity risk and impact assessments.

Introduction

The drastic degradation of nature is widely appreciated to be driven in part by large multi-national corporations and institutions, causing economic and financial risks for organisations that depend on and/or impact nature (1). As a consequence, the field of sustainable finance has been rapidly gaining momentum and with it increasing pressure from both financial regulators (1,2) and investors for corporations to measure and mitigate their impacts on biodiversity (3,4).

Organisations need data on the state of biodiversity in areas where they interact with it, to measure and report on nature-related financial risks, opportunities and impacts to inform investors, regulators or the general public. To do so, organisations rely on third party, often global, biodiversity datasets and tools (5), due to their ease of use, (global) coverage, limited input data requirements, free availability or other reasons. It is for the same reasons that these tools are often promoted or signposted by industry initiatives such as the Taskforce for Nature-related Disclosures (TNFD) (6) or the EU's Business @ Biodiversity Platform (7), but they face several limitations when it comes to understanding and reporting on biodiversity impact. Many of these tools present an indirect biodiversity impact assessment based on pressures and practices rather than direct biodiversity observations (8,9). For example, land-use change or conventional agricultural practices are well-documented drivers of biodiversity loss. These pressure data points rely on aggregated sector, global or country level averages and offer limited guidance to inform risk mitigation or investment decisions (10). Where tools incorporate actual biodiversity or species data, these cover only a small set of species, excluding species that contribute most to ecosystem functioning such as invertebrates, fungi and microorganisms (10,11). Similarly many tools focus on only a single summary metric, which provides an over-simplified view of complex natural systems. Finally, biodiversity data is highly geographically and taxonomically biased, which limits the utility of global models for assessments (12).

Fortunately, technological advances are rapidly expanding the scale at which it is possible to collect additional biodiversity data. For example, eDNA allows rapid and cost-effective assessment of previously difficult to work with biodiversity components, such as the invertebrate fauna (13). Similar advances in acoustic monitoring, camera traps, image analysis and collection (e.g. through the use of unmanned aerial vehicles) are also accelerating the rate of data collection (14–16). In recent years, a large number of high-quality biodiversity datasets have been published and made accessible, for example via the Global Biodiversity Information Facility (GBIF) (www.gbif.org). However, there is still a lack of unbiased biodiversity data (17) at the scales necessary to assess changes driven by institutional activity, such as land development, agriculture, or natural resource extraction (18).

Through the rapid adoption and use of remote sensing technology, earth observation (EO) data containing environmental information is also increasingly more detailed and accessible (9,19). There is now an abundance of high spatial resolution EO data containing climatic and habitat cover variables, often covering continental or global scales. One key example of delivering such data is Copernicus, the open access EO part of the European Union's space program (20). Using these high resolution environmental data combined with high-quality biodiversity datasets, it is possible to use the statistical relationships between environmental data and biodiversity observations to estimate key biodiversity variables at un-surveyed or low-surveyed locations.

Due to its inherent complexity and diversity, predictive modelling of biodiversity dynamics can be a time consuming, and challenging task. Typically, models predict the abundance or distribution of species individually, however even modern computational and statistical models can find this challenging when predicting the incidence of thousands of species. An alternative approach is therefore to derive summary measures of biodiversity, for example 'Essential Biodiversity Variables' (EBV's) (21) , and then predict these to new locations. This approach is an attractive proposition for practical use-cases, as it can simplify the statistical task, and allow large-scale assessment of EBV patterns. However, many implementations are difficult to validate due to limited or no raw biodiversity data available at the locations of interest (e.g. see (55)).

Here we present results showing a significant reliance on global datasets and tools over local or national datasets, used to prioritise interventions and resources. In a recent review, we outlined the potential of using eDNA collection for biodiversity impact assessment, and also noted the complementarity of this type of data and Earth Observation data (8). Here, we combine these two data sources and use flexible machine learning algorithms to predict summary measures of biodiversity in an EBV framework. We demonstrate the difficulty in estimating generalisable patterns of EBV's despite using state of the art biodiversity data and high resolution environmental covariates. The results clearly show that predicting EBV's to new locations, without local data, considerably reduces predictive performance. This highlights a risk in the current approach to biodiversity risk and impact assessments by both real economy corporates and financial institutions. Although model derived estimates and data products can and will be useful for biodiversity impact assessment and reporting, increased pressure must be placed on the collection and analysis of local biodiversity data. We end with a set of recommendations for improving biodiversity assessments by using a biodiversity data hierarchy framework, and tracking stepwise improvements with a data quality focus.

Methods

Analysing current use of biodiversity datasets and tools

Given the nascent and rapidly evolving nature of nature-related financial disclosures, we performed a qualitative content analysis of a sample of TNFD-aligned corporate disclosures to understand the current use of biodiversity datasets by corporations and financial institutions. The sample included 84 publicly available reports, 57 from real economy corporations and 27 from financial institutions, listed on the Taskforce on Nature-related Financial Disclosures (TNFD) website as TNFD report examples (49). From this list, we only analysed reports that were available in English.

We used NotebookLM to identify datasets and tools for assessing nature-related financial risks mentioned in each report, using consistent prompts. NotebookLM is an online tool for synthesizing documents using artificial intelligence, based on uploaded documents only, which provides direct references to the source text for each of its findings. The list of datasets and tools was then manually validated and use cases and local datasets were categorised. Categories were developed inductively through iterative reading and interpreting of extracted information. An overview of the full sample of reports, results, and categories is available in Supplementary Materials part 1.

Biodiversity data

To calculate our summary measures, we use data from the Insect Biome Atlas project, which is one of the largest invertebrate surveys to date (22). The survey consisted of large-scale malaise trap inventories across the extents of Sweden and Madagascar. Malaise traps were placed at 198 locations in Sweden, and 50 in Madagascar, and emptied at frequent intervals (monthly – weekly, depending on the season). This survey campaign yielded over 4500 samples from Sweden and 2000 from Madagascar, samples were stored in ethanol and processed using a state-of-the-art molecular identification pipeline (22,23). This data represents the gold standard in invertebrate diversity surveys, yielding over 30,000 operational taxonomic units (OTU's), for Sweden and over 70,000 for Madagascar.

EBV's

Natural systems are inherently complex and any approach to assess impact must measure multiple dimensions of biodiversity. Given this, we use the GEO BON Essential Biodiversity Variable (EBV) framework (21) to guide our derivation of parameters suitable to assess the state of multiple axes of invertebrate diversity using the IBA data. These EBVs aim to capture the key processes that shape natural systems at multiple levels of organisation [11], and provide an attractive framework for impact reporting as they capture the multiple levels required to effectively conserve nature.

From our molecular survey data, we derive five summary EBV's that describe key aspects of biodiversity spanning three of the GEOBON EBV classes (TABLE 1).

EBV name	**GEOBON class**	**Summary description**
Species richness	Community	The total OTU richness of a site
Local contribution to beta diversity (LCBD)	Community	A metric quantifying the uniqueness of a site in terms of community composition
Functional dispersion	Traits / Community	A metric quantifying the functional diversity of a community

Functional evenness	Traits / Community	A metric quantifying the evenness of functional traits of a community
Genetic diversity	Genetic	A metric quantifying the genetic diversity across all organisms at a site

TABLE 1: Essential biodiversity variables (EBVs) that were derived from the invertebrate data.

Species richness

We derive a measure of species richness, i.e operational taxonomic unit (OTU) richness, to represent a measure of taxonomic diversity for each individual sample. This measure provides an overall summary of the species diversity at a site. We derive this measure by simply taking the sum of unique OTU's within each sample.

Local contributions to beta diversity

Although richness is a useful summary, it is often an inadequate measure to capture the full complexity of biological diversity (24,25). Ecological communities often display emergent properties, and an important component of diversity is the composition of organisms at each site. Conserving unique species or communities is an essential aspect of protecting diversity, which might be neglected by species richness measures, whilst species richness (alpha diversity) describes the overall species diversity within a site, *beta* diversity metrics aim to describe the differences in composition between sites. We therefore derive a measure of beta diversity using *Local Contributions to Beta Diversity* (LCBD) which aims to capture the ecological uniqueness at each site (26). We calculate LCBD using dissimilarity matrices describing the species replacement differences (i.e. species turnover) (27) to enable our measure of LCBD to indicate sites that may have high conservation value in terms of rare or unique species. To provide comparative measures of LCBD between samples, we excluded weeks where the number of sites surveyed were lower than 40. This ensured survey effort remained relatively even across samples, and allowed a better comparison of community uniqueness across sites.

Functional dispersal and evenness

The diversity of functional traits within communities influence ecosystem functions and is therefore central to conservation efforts (28–30). To describe the trait diversity of our communities we derive two metrics of functional diversity. i) *Functional dispersal* which describes the total extent of the trait space i.e. the total trait diversity displayed across species within a site-level community. ii) *Functional evenness* which describes the evenness of the distribution of species-level traits within a site-level community. Consideration of both of these metrics is important, as alone they may provide an incomplete picture of functional diversity. For example a site with *high dispersal* might indicate a superficially (functionally) diverse community, but in the context of *low evenness* this community may actually be composed of only a few disparate clusters of species, with species within each cluster sharing very similar traits. On the other hand, *high dispersal* with *high evenness* might indicate a diverse community, with all organisms displaying differentiation in their functional traits.

We use trait data collated from multiple sources to describe three axes of functional diversity. These include feeding niche and habitat designations (derived from (31)), which relate to the feeding ecology of the organisms, and body size data derived from (32), which is a general proxy for many life-history related traits (33). As our trait data consist of categorical (feeding niche and habitat categories) and continuous measurements (body size) we transform our trait data to standardise measurements across data types by calculating the Gower distance between OTU's based on trait data measurements (34). We then use the Gower distance measurements in a principal co-ordinates analysis (PCoA), to project measurements in a n-dimensional trait space, the axes of this PCoA analysis are then used to characterise each site in terms of it's functional composition. This provides a quantitative representation of the difference between trait composition within our samples (35). We assessed the quality of functional trait spaces using the approach outlined in (30). We then use the first three dimensions of the PCoA axes (describing 88% and 86% of trait variation in Sweden and Madagascar respectively) as our trait measurements, and use these with the respective *evenness* and *dispersal* functions within the *fundiversity* R package (36).

Genetic diversity

To summarise the effective genetic diversity present at each site, we derive measure of the total sample-level genetic variability using a Shannon index of the frequency of amplicon sequence variants (ASV's) from our molecular data:

$$H = -\sum p\_i * \log(p\_i)$$

Where *p_i* represents the proportion of ASV's belonging to species *i*. We then take the sample wise mean of this measure to summarise the total genetic variability of a sample, so that our EBV represents a measure of the average species-level genetic variability across all species present within a sample. To account for sequencing and amplification bias that might inflate ASV frequency, we calibrate species level ASV counts based on the frequency of ASV's detected from spike-in sequences..

Feature	Timescale	Spatial resolution	country	source	References
Maximum temperature	Daily (weekly average)	0.1° x 0.1° (~ 9 km)	Both	ERA5	(37)
Minimum temperature	Daily (weekly average)	0.1° x 0.1° (~ 9 km)	Both	ERA5	(37)
Precipitatio n	Daily (weekly average)	0.1° x 0.1° (~ 9 km)	Both	ERA5	(37)
Leaf area index (high)	Daily (weekly average)	0.1° x 0.1° (~ 9 km)	Both	ERA5	(37)
Leaf area index (low)	Daily (weekly average)	0.1° x 0.1° (~ 9 km)	Both	ERA5	(37)
Photoperiod	Daily (weekly average)	NA	Both	Derived from time of the year and trap location	NA
Seasonal components (x2)	Weekly	NA	Both	Harmonic functions of the week of the year	NA
Unknown forest	static	100m	Both	Copernicus	(38)
Evergreen broadleaf forest	static	100m	Madaga scar only	Copernicus	(38)

Deciduous broadleaf forest	static	100m	Both	Copernicus	(38)
Evergreen needle-leaf forest	static	100m	Sweden only	Copernicus	(38)
Mixed forest	static	100m	Sweden only	Copernicus	(38)
Unknown forest	static	100m	Both	Copernicus	(38)
Crop cover	static	100m	Both	Copernicus	(38)
Shrub cover	static	100m	Both	Copernicus	(38)
Moss cover	static	100m	Sweden only	Copernicus	(38)
Grass cover	static	100m	Both	Copernicus	(38)
Water cover	static	100m	Sweden only	Copernicus	(38)
Urban cover	static	100m	Sweden only	Copernicus	(38)
Sampling period	Weekly	NA	Both	Derived from sample time between observations	NA

TABLE 2: Environmental covariates that were derived from the earth observation data, along with resolution and source information.

Environmental data

To predict our EBV's we compile a list of environmental covariates known to correlate with invertebrate biodiversity, including climatic and landcover variables. We derive these variables from high-resolution earth observation data from several sources (TABLE 2). Included in these variables we include two 'seasonal terms' to account for unobserved seasonal patterns not captured by our environmental data. These seasonal terms are calculated as harmonic sinusoidal functions of our time index (i..e the week of the year):

$$S1 = sin(2\pi t/T)$$

$$S2 = cos(2\pi t/T),$$

Where *t* indexes the specific week of the year, and *T* indexes the total size of the time periods (i.e. 52). This produces two seasonal covariate terms that we include as predictors in our model. We also derive the photoperiod as a function of time of the year and latitude of the sample site. We also include sampling period as a function of the time between samples to account for variation in sampling effort due to differing exposure times of traps across seasons.

Assessing spatial and temporal trends in EBV data.

As a major component of the predictive performance and utility of models are the presence of correlations between environmental variation and the response variable (i.e. EBV's), we make broad assessments of the presence of temporal and spatial signals in the raw EBV data. To assess this we fit generalised additive models in the R package mgcv (39). We fit models to each EBV containing spatial and temporal terms that aim to capture broad-scale environmental variability but include no specific environmental covariates:

$$\mathbf{y = s(week) + s(lat,lon),}$$

where *s()* represents smooth terms for the week of the year, and an interactive term between latitude and longitude respectively. Where appropriate (i.e. for species richness models) we include an offset term to account for unequal sampling effort between summer (weekly samples - April - September), and winter (monthly samples). We assess the degree of spatial and temporal trends by plotting model fits and visually evaluating the trends.

Model fitting

To assess the statistical relationships between EBV's and our environmental data we fit gradient boosted regression trees implemented in the XGboost R package (40).

These models represent flexible approaches to prediction tasks and have demonstrated good predictive performance in similar biodiversity applications (41). To optimise predictive performance we optimise hyperparameters by performing a simple grid-search across possible combinations. As the learning rate parameter (eta) often influences model performance more significantly than other parameters, we split this process into two stages. First we perform a grid search across a larger number of parameter values for eta, whilst holding all other parameter values constant. Once we have found an adequate value for eta, we then perform a grid search over the remaining hyperparameter values. The values for each grid-search stage are detailed in tables S4 & S5, and final hyperparameter values for each final model are listed in table S6.

Model validation

To assess model performance within sites and at new sites we conduct two cross-validation exercises, see Figure 1 for an illustration. To assess how well our models predict EBV values within our sampled sites we perform a stratified test-train split, i.e. we remove 20% of observations from each individual site, and attempt to predict these missing data from the model fit. This approach removes a total of 20% of all observations from our dataset simultaneously. This exercise provides us with a baseline value on which we can judge out-of-site prediction but also serves to assess how models may perform with the inclusion of local data.

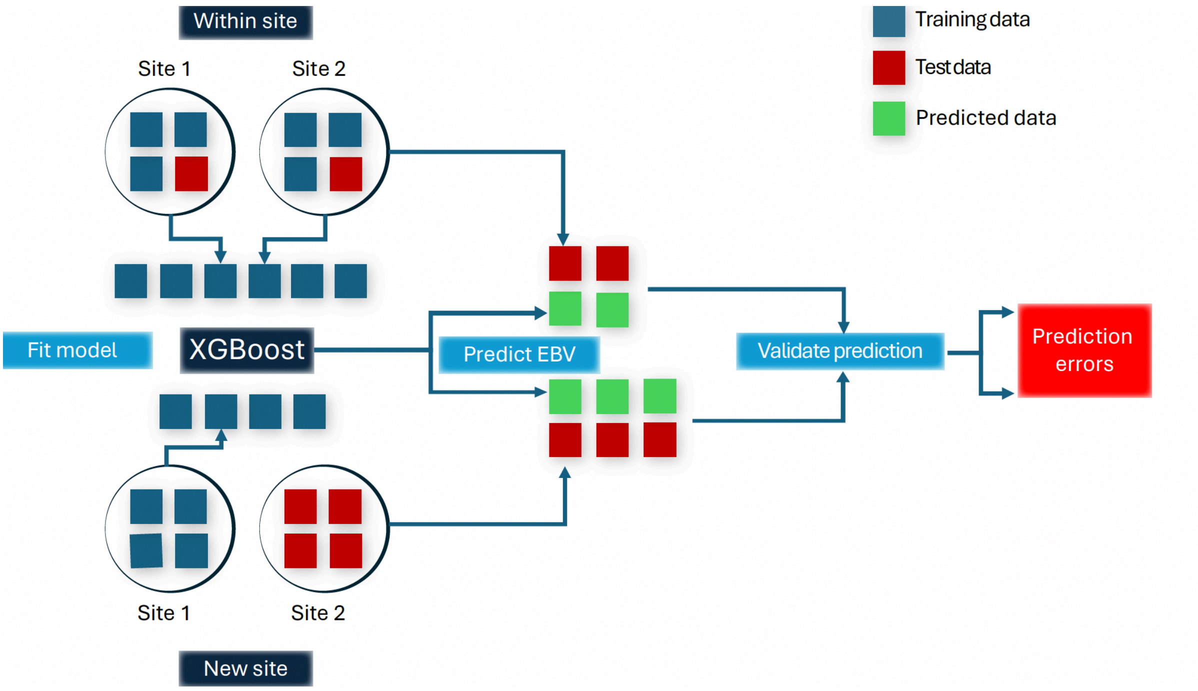

Figure 1: Predicting EBVs from the trained model within sites (top panel) and to new sites (bottom panel). The model has been trained on biodiversity data from the Insect Biome Atlas in Sweden and Madagascar, and environmental data from Copernicus/ERA5.

To assess out-of-sample prediction we perform a leave-one-site-out cross validation exercise, where each individual site is sequentially left out of the training data. In each iteration only a single site's worth of data is removed from the model, and models are fit to the remaining data. Missing site-level test data are predicted to assess accuracy. This approach predicts values at entirely new sites and emulates the mapping of EBV's across large spatial scales. We use the mean absolute error (MAE) between observed and predicted values to assess predictive accuracy,

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y_i}| \quad (1)$$

Where y_i represents the observed value of a given EBV at observation i, and $\widehat{y_i}$ represents the model predicted value. n represents the total number of observations in the sample. For each CV exercise we calculate MAE for each site-level held-out sample, providing a measure of the predictive performance for that given site.

Assessing utility of the predictions:

To assess how predicted measures of EBV's perform in decision making, we compare site rankings (i.e. the rank of each site in terms of its absolute value of a given EBV), across observed and predicted values. Site rankings are a common tool in impact assessment and ecological surveys to determine the attribution of funds, plan protections or interventions (42–45). Changes in rank between observed and predicted site-level EBV values provide a more tractable metric to assess the utility of predicted measures of EBV's for decision making. We categorise sites into four equal sized quantiles (i.e. priority-tiers) based on their ranking, then assess the change in group (quantile), between observed and predicted values. Site-ranking followed by classification into priority tiers (e.g. 'high', 'medium', 'low', or 'priority' vs 'non-priority') is commonly used in conservation planning and resource allocation and aligns well with practice where funding is assigned based on broad categories of status (46–48).

Results

Current use of biodiversity datasets and tools for risk assessment purposes

Across our sample we see that global tools for screening nature-related impacts and dependencies such as ENCORE, are widely used by both real economy corporates (67%) and financial institutions (93%). Other widely used global tools focus either on water, such as WRI Aqueduct (31%) or WWF Water Risk Filter (12%) or focus on biodiversity, such as IBAT (24%) or WWF Biodiversity Risk Filter (21%).
Specific data layers which are used to represent the state of biodiversity include the IUCN Red List (13%), Global Forest Watch (10%), World Database on Key Biodiversity Areas (7%), Biodiversity Intactness Index (6%), RESOLVE Ecoregions and Biomes (2%) and Ecologically or Biologically significant Marine Areas (2%).

In general we find that financial institutions primarily use global datasets and tools for assessing nature-related risks and impacts across their portfolios, with only 26% of financial institutions mentioning the use of any national, local or internal datasets. Where these are mentioned, it is typically for banks or insurers that have national portfolios or specialist institutions invested in physical assets (e.g., renewable energy) rather than financial assets (e.g, equities or bonds). Nearly all corporates analysed (95%) use at least one global tool or dataset to assess nature-related risks and impacts, but they are more likely to complement this with internal, local or national data (61%). This includes data which has been collected previously as part of (baseline) environmental impact assessments (19%), as part of ongoing (operational) monitoring processes (35%) or that was collected specifically for the purpose of the assessment exercise through internal, supplier or community surveys and workshops (19%). Additionally, third party data layers, typically from state, national or regional specialist sources are used to complement internal and global datasets for the assessment (23%).

The application of these tools and datasets primarily serve a general screening purpose and creation of a heatmap of potential impacts and dependencies across companies' operations or supply chains (75%) or across financial institutions' portfolios (85%). Sometimes this is where the analysis ends, but often the results are used to identify sites, suppliers or portfolio companies, as well as individual environmental issues, pressures or ecosystem services that are prioritised for further analysis. Financial institutions also use the results of these assessments to identify priority companies or topics for their stewardship and engagement strategies (41%). Other use cases include due diligence of suppliers (4%), quantification for reporting purposes (4%), scenario analysis (4%), tracking own performance against targets (2%) or benchmarks (2%), supporting clients (2%) or integrating within investment decisions (1%).

The preference of global over national, local or internal information is not surprising. Many of the reports analysed are prepared by companies with global operations or global portfolios. The use of national datasets is more prevalent in entities with a strong national footprint, although even then global screening tools are used. While the fact that financial institutions hardly use internal data is to be expected, we find that only 49% of corporates mention the use of internal datasets to inform a company wide risk assessment. This would mean that for half of the sample, third party national and global datasets and tools are used at face value, without tailoring the outputs based on internal, operational insights.

Many of these organisations are still in the early stages of understanding and assessing nature-related (financial) risks. For many organisations these are the first TNFD aligned reports that they have produced and they often mention ambitions to improve data collection and granularity or scope of analysis in the future. Our analysis itself is also limited by what companies actually disclose in terms of methodology, data sources and usage of results which may underestimate the level of sophistication and capabilities within organisations. However it does show that companies and financial institutions today rely primarily on global, generalised tools and datasets to understand issues which are inherently local and context specific, even when more granular or context specific information is available. While these are primarily used in the initial stages of a risk assessment and management process, they do inform where scarce (financial) resources and attention are prioritised to mitigate risks and impacts.

Quality of model-based assessments

Temporal variation. In our analysis of the quality of model-based assessments, we found that EBV's exhibited a range of temporal trends between Sweden (Figure 2A), and Madagascar (Figure 2B). In Sweden, most trends correlated with seasonal shifts in temperature, with EBV values changing between the summer and winter months. However, EBV's displayed large differences between the degree of seasonal variability. Species richness and LCBD, showing large temporal trends, functional trait dispersal and evenness showed shifts of similar degree, whilst genetic diversity displayed lower temporal variability. EBV's in Madagascar displayed trends that correlated with changes between the wet and dry season (Figure 2B), however all EBV's displayed relatively little temporal change. Temporal trends in Species richness, LCBD, and functional metrics were small, whilst genetic diversity remained more or less the same over the year.

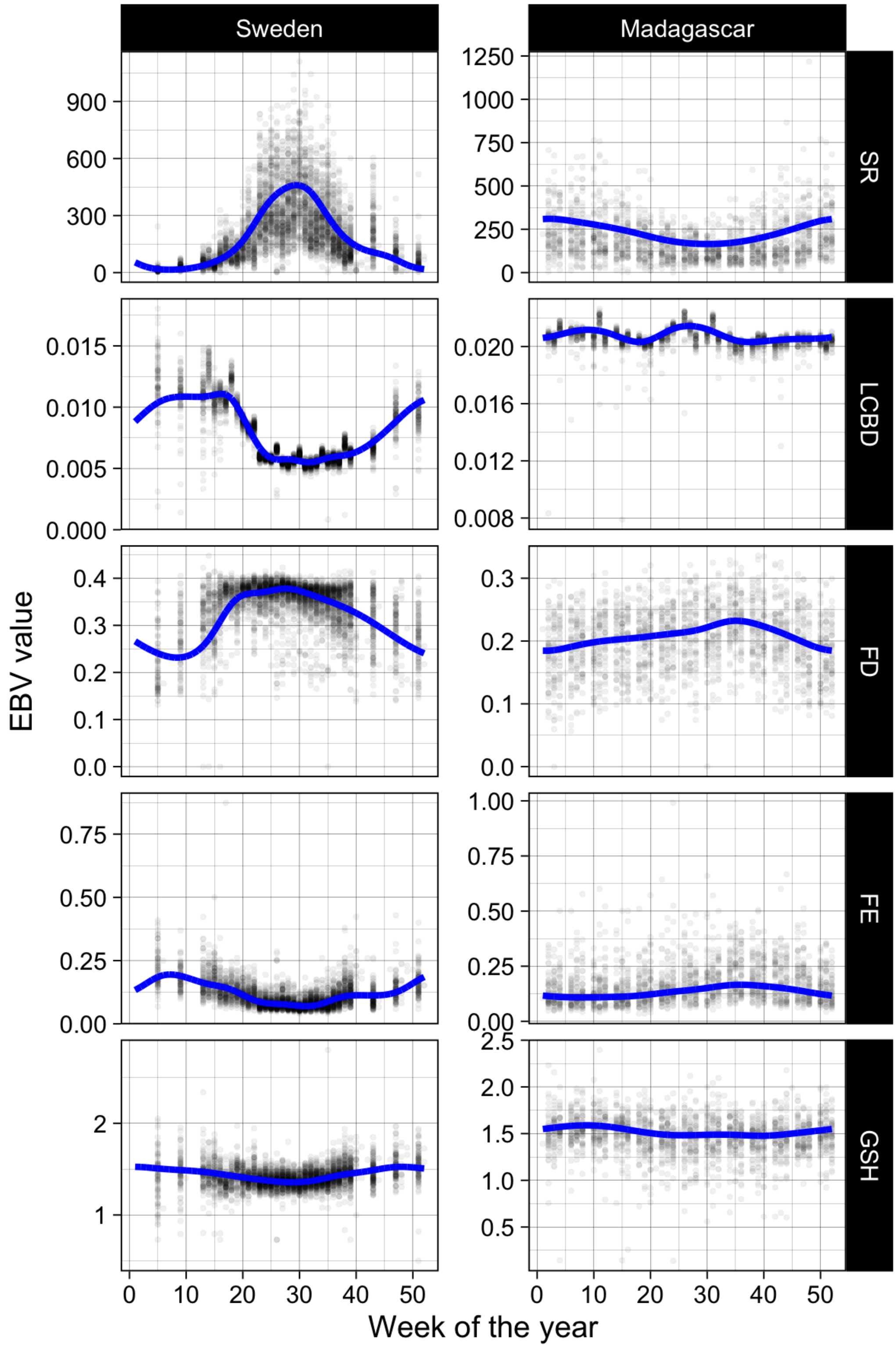

Figure 2. Estimated temporal trends for each EBV for each country. Results are presented for species richness (SR),Local contributions to beta diversity (LCBD), functional dispersal (FD), functional evenness (FE) and genetic diversity (GSH) for Sweden and Madagascar..

Spatial variation. Similar to temporal trends, EBV's exhibited different degrees of spatial variation across both countries. In Sweden (Figure 3A), species richness displayed a latitudinal gradient with higher concentrations in the south of the country, and a similar but less pronounced pattern was displayed for functional dispersal. Functional evenness displayed the inverse pattern, with lower values in the south of the country. Genetic diversity and LCBD demonstrated very little spatial patterning.

In Madagascar (Figure 3B), species richness displayed high degrees of spatial pattern, with concentrations in the north east of the country. Functional evenness displayed the inverse of this pattern, with high values concentrated to the south west of the country. Unlike in Sweden, genetic diversity did display spatial variation, with the south of the island having higher concentrations. LCBD did not show any significant spatial variation.

Predictive performance. When predicting EBV values at new sites, prediction error increased for all EBV's in both countries, however the degree of this increase was both country and EBV dependent (Figure 4). Prediction error for species richness at new sites increased considerably for models trained on both Swedish and Malagasy data (MAE increased by 62% and 68% respectively). Genetic diversity (GSH) prediction error also increased by similar degrees in both countries, 25% in Sweden, and 20% in Madagascar. However, the increase in prediction error for functional evenness (FE) was considerably larger in Sweden (48%), than for Madagascar (36%). Inversely, for functional dispersal (FD) the increase in prediction error was larger in Madagascar (32%), than for Sweden (19%). Predictive errors for community uniqueness (LCBD) in Sweden displayed were relatively high (18%), but were negligible for Madagascar (1%).

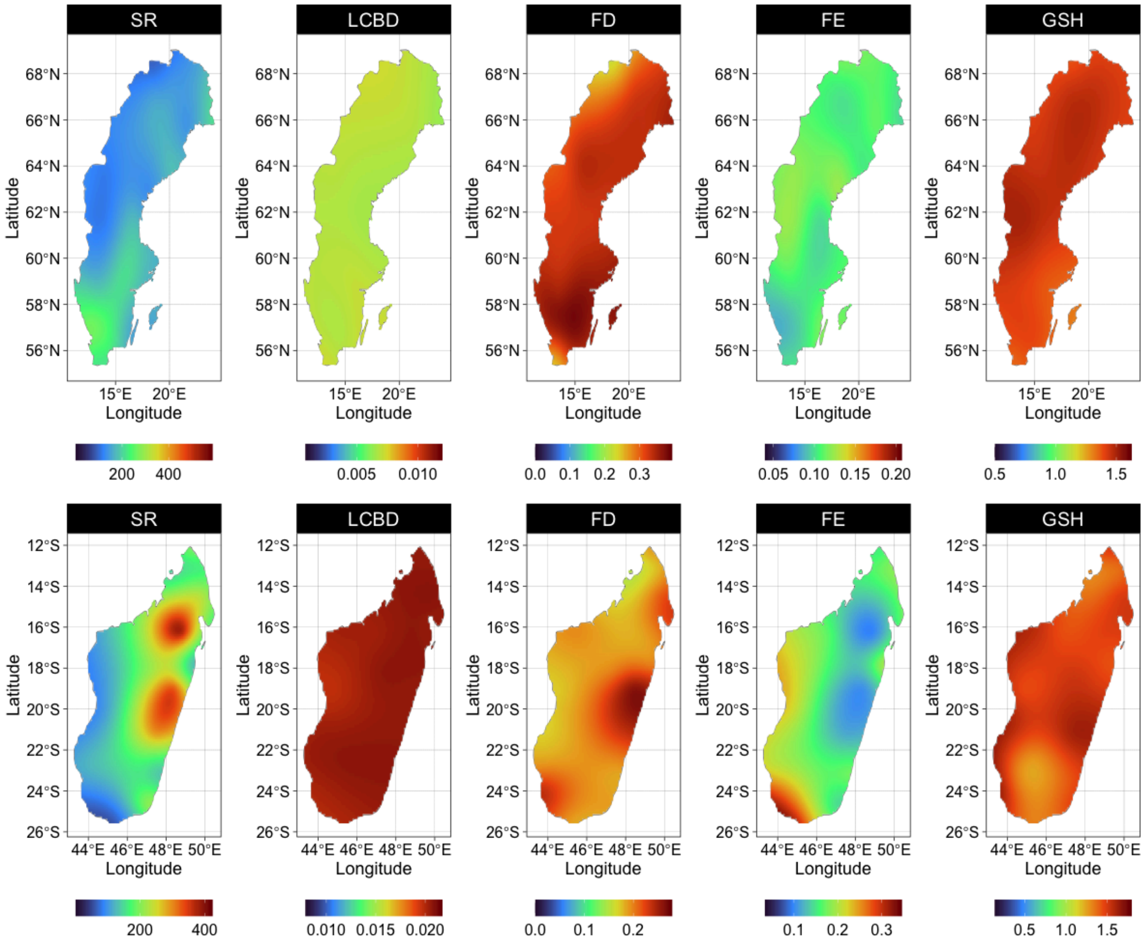

Figure 3. Estimated spatial trends for each EBV for each country. Sweden top row, Madagascar bottom row.

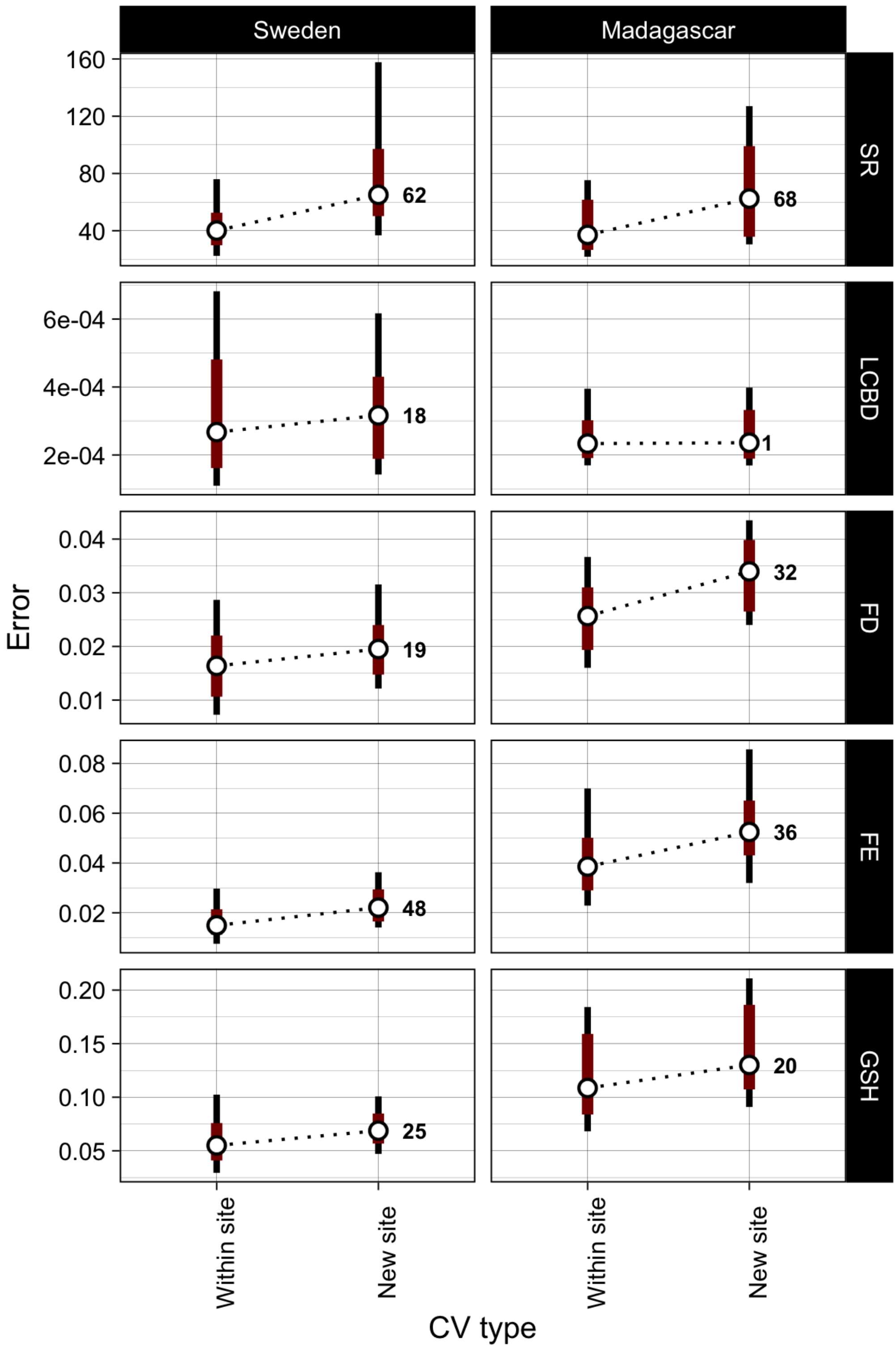

Figure 4. Mean absolute errors (MAE) from the cross validation exercise for the stratified test train split (within site), and leave one site out (new site) methods. Results are presented for species richness (SR),Local contributions to beta diversity (LCBD), functional dispersal (FD), functional evenness (FE) and genetic diversity (GSH) for Sweden and Madagascar. Numbers next to each point represent the mean absolute percentage error change between within site and new site validation sets. Points represent the median MAE value, and thick and thin error bars are 50% and 80% quantile intervals respectively.

Change of ranking. Group rank change for both within and new site predictive exercises was considerable across all EBV's and both countries (Figure 5), but again was both EBV and location dependent. In both countries group ranks for species richness predictions were relatively well preserved for within-site predictive tasks with 78% of all sites remaining in their original group rank for Sweden and Madagascar respectively. However, this dropped to 35% and 44% when predictions were made to new sites, with 25% and 15% of all sites changing at least two group rank categories (e.g. moving from the group containing the highest richness sites, to the group containing the second lowest richness sites). Similar trends were seen for Functional dispersal and functional evenness in Madagascar, with 73% and 87% of sites being placed in the correct category for within site prediction, yet these fall to 38% and 44% when predictions are made to new sites. However most predictions (78% for both EBV's) remain within one category change from their original, and only 22% for over two rank changes for dispersal and evenness. Sweden displayed similar patterns for functional metrics, the within site group rank change for functional metrics were only 68% and 73% for dispersal and evenness respectively, however proportional decrease in miss classifying priority tiers was more varied, with functional dispersal having 56% of sites retain their original category, but eveness decreased to only 31% of sites.

LCBD was relatively consistent between countries in terms of group rank change, with Sweden displaying very little change in group rankings between within-site and new-site predictions, decreasing from 72% to 68%. In Madagascar, 82% of sites retained their original group rank for within-site predictions, but this decreased to 69% for new-site predictions. Finally, within-site group rank was reasonably well retained for genetic diversity in Madagascar at 69% , but decreased to 31% for new sites. Conversely Sweden, had group rank retention of 64% and 42% for within and new site predictions.

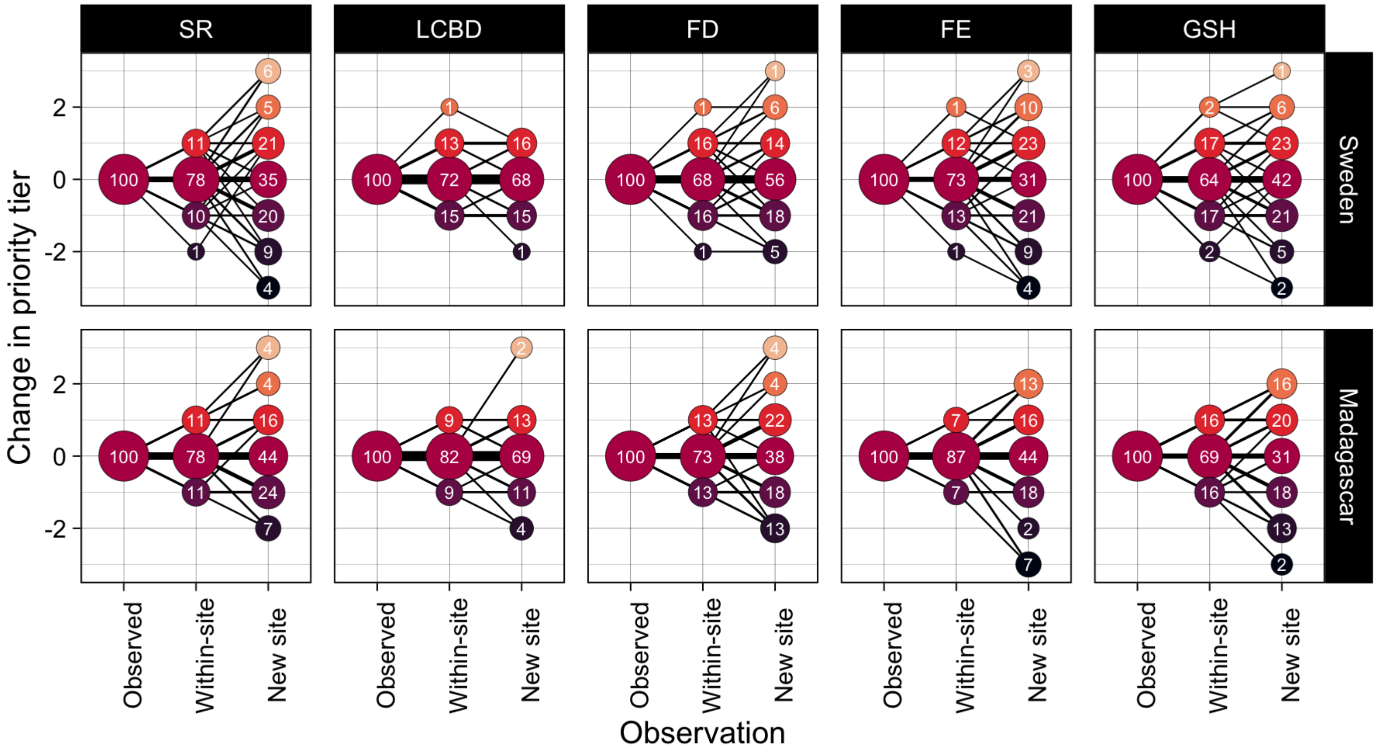

Figure 5. The change in priority tier between observed, within-site, and new site predictions for each EBV in Sweden (A), and Madagascar (B). Figures show the percentage of sites that change ranking groups (i.e. the quartile assigned to the observed EBV value) for within-site and new site prediction compared to their original ranking group. The Y axis illustrates the directional change in priority tier ranking from its observed tier, i.e. a value of -1 illustrates that sites moved down a single tier rank compared to their original observed category. The size of the points is proportional to the total number of sites in that category, and the percentage of sites in each is displayed on each point. Line thickness illustrates the number of sites that changed priority tiers between observed, within site predicted values, and new site predicted values.

Discussion

Results' implications for biodiversity risk and impact assessments

Using the summarise and predict approach (50,51), and the growing amount of large high-quality biodiversity datasets and relevant environmental variables, EBVs can be assessed on a large scale with a computationally feasible approach. This would address a key obstacle for biodiversity risk and impact assessments: the data deficiency. However, the demonstrated degree of variation in the individual estimated EBVs, as well as the predictive errors, show high dependency on both geographical location and seasonality. Especially at sites without any local biodiversity data the predictive performance drops significantly, illustrated by the observed change of rank of individual sites. These results show the difficulty of basing a reliable biodiversity risk and impact assessment on new site predictions, for example when choosing one site over another for development by site rankings. At the core is the crucial context-dependency of biodiversity measurements, making large-scale extrapolations to new sites challenging even using high-quality datasets from other sites.

Implications for corporates and financial institutions

These results have implications for companies and investors relying on global extrapolated biodiversity datasets for site-level analyses. And it is at the site-level that these institutions are assessing and reporting their nature-related financial risks, in line with the TNFD's Locate, Evaluate, Assess and Prepare approach. Hence low predictability of summary values for sites without locally collected biodiversity information will be an issue for corporates as very little of the biodiversity data available to researchers, NGOs and tool developers, is available for sites operated by corporates. For instance only 0.3% of the total occurrence records on the GBIF platform has been reported by corporates (52). Ultimately this leads to highly different or inappropriate results which are used to identify priority sites or companies and inform resource allocation decisions for further nature action interventions.

The risk of misuse or inappropriate use of the (global) biodiversity datasets and tools we've identified being used by businesses and financial institutions, is increasingly being documented by the scientific community (9,10,12,53). One way to mitigate this could be to advocate for datasets and tools to be accompanied by some evaluation of the predictive error or model uncertainty, which some already provide. However such supply side measures might not be sufficient to avoid misuse as documented by (53): "*While many tool providers caveat that their products are intended only for sector-level screening purposes, companies may nonetheless use them to 'tick the boxes' for due diligence on nature-related risks. In practice, firms may refer to these tool results to assert there were no issues, even though the tools were never designed to deliver site-specific risk assessments"*.

One alternative way to communicate this issue is through the concept of data quality. In 2022 the Partnership for Carbon Accounting Financials launched a global accounting and reporting standard for financed emissions by financial institutions. This standard emphasizes the notion of data quality and the variations of data quality from different types and sources of reported and modelled greenhouse gas emissions of portfolio companies. It introduces a 'data hierarchy' and a 'data quality score' and argues for the use of the best available data point at the time of analysis or reporting; the calculation and reporting of the aggregate data quality score from input data points; and the constant improvement of the data quality score (54). Similar principles could be applied to incentivise primary biodiversity data collection and reporting to reduce uncertainties and contribute to the improvement of (global) summary biodiversity measures.

Moving towards higher data quality

Using the concepts of data hierarchy and data quality scores, we propose a first draft of a biodiversity data hierarchy in Table 3. The goal would be to track improvements in corporate biodiversity risk and impact assessment through such a framework, similar to the existing one for greenhouse gas emissions. Table 3 lists different states and types of biodiversity information, and we call the "processing pathways" the process of moving between the boxes. For example, our present analysis has taken direct biodiversity data (Box a), into EBVs (Box A) and predicted those summary measures within sites (Box B) and to new sites (Box C). An even more accurate approach, but computationally and modelling-wise more demanding, would be to go from biodiversity data (a) to predicted biodiversity data (b), and then calculate predicted summary measures (B). Current common practice in biodiversity risk and impact assessments is often a mix of proxies and some small quantity of measured data. For example the GLOBIO model (55) uses a combination of direct biodiversity data and proxies to produce prediction summaries (aA + dD). ENCORE, also widely used, does not use any measured biodiversity data directly, but rather works with proxies and pressures (dD). The problem with these commonly used pathways is that they lose the context dependency of any actual biodiversity data used, and the ratio of relevant and informative biodiversity data versus proxy data is often very low. When used to assess specific sites of interest to a corporation, these issues could easily lead to an inflated impression of well-grounded results and mask uninformative assessments.

What would a high-quality pathway look like, while also being computationally feasible for corporations? Our suggestion is that this would be the pathway aAB, or aABC while making sure there are limited new sites where predictions are made without any local data. Using a proxy-based pathway such as dD to assess biodiversity may also be both realistic and suitable for specific purposes, but most likely at larger scales than for individual companies. This could be for example when investors are screening large portfolios of assets, or for supply-chain analyses that have known and trusted proxy variables for biodiversity impact. Crucially, the uncertainty of predictions, or model outcomes, should accompany results. Using a data hierarchy and data quality scores should be regarded as an essential complement to biodiversity risk and impact assessments and reporting, both clarifying what level of data quality has gone into the results but also tracking improvements over time.

	Raw data	Summary
Direct data	**a. Biodiversity data** example: occurrence data in GBIF	**A. Data summary measures** example: our EBV calculations
Predicted data in existing sites	**b. Predicted biodiversity data** example: species distribution models, **within sampled sites**	**B. Predicted summary measures** example: our predicted EBVs, **within sampled sites**
Predicted data in new sites	**c. Predicted biodiversity data** example: species distribution models, **to new sites**	**C. Predicted summary measures** example: our predicted EBVs, **to new sites**
Proxy data	**d. Proxies for biodiversity data** example: land use change, pressures	**D. Summary measures based on proxies**

TABLE 3. A first draft of a biodiversity data hierarchy. The table lists different states and types of biodiversity information. The *Data* column and the *Summary* column are states, representing raw data versus summarized data. The rows are types of biodiversity information, divided into *Direct data*, which is measured data, two types of *Predicted data,* and *Proxy data*. The process of moving between boxes is called processing pathways.

Practical recommendations for data collection at sites of interest

As noted in the Introduction, modern monitoring methods have made collection of local biodiversity data feasible and affordable. Our results show that sites completely without data will have predictions with high levels of uncertainty, even when combining high-quality global biodiversity datasets and environmental data. Even a small amount of local data is likely to improve model performance significantly. Concrete advice for corporations implementing data collection at sites of interest include noting the local context, designing appropriate sampling strategies and investigating potential reference datasets. There is material available guiding the private sector on best practices of data collection and sharing (see e.g. (56)). In an ideal situation any data collected by individual institutions and corporations would be collected according to global data standards and subsequently uploaded to platforms such as GBIF (18,52). This would then contribute to the pool of open reference data and create a positive spiral of data growth.

Outlook

Machine learning and statistical modelling has potential to contribute with real insights and utility at large scales for corporate biodiversity risk and impact assessments. However, the assessments will only be as good as the underlying training data for the models, and given the context-specific nature of biodiversity variables local data remains key. The landscape of biodiversity risk and impact reporting is still developing, and therefore there is currently a window of opportunity for corporations to highlight the stepwise improvements the field is making. Increased data collection by corporations, even at low levels, at their own sites of impact would put their risk and impact assessments on a high-quality data path for the future. We propose to track these incremental improvements using a data hierarchy and data quality scores, similar to what is done for greenhouse gas emissions, as outlined in Table 3. Developing and improving on such a framework is, in our opinion, a promising direction of work both for the academic and business community.

Acknowledgments

This research was conducted under the Finance to Revive Biodiversity (FinBio) program, financed by Mistra – the Swedish Foundation for Strategic Environmental Research (DIA 2020/10). Part of the computations were carried out on the Dardel computing resource, which is part of the National Academic Infrastructure for Supercomputing in Sweden (NAISS), and is financially supported by the Swedish Research Council.

CRediT authorship contribution statement

Conceptualization: E.G., R.G., C.C., F.R.

Formal analysis: R.G., E.G.

Investigation: C.C.

Writing - Original Draft: E.G., C.C., R.G.

Writing - Review & Editing: E.G., R.G., F.R.

Funding acquisition: F.R.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data and code availability

The underlying data and code is available at https://github.com/ronquistlab/finbio-ebv including instructions on how to download the raw data from the original sources.

References

1. Dasgupta P. The economics of biodiversity: the Dasgupta review: full report. Updated: 18 February 2021. London: HM Treasury; 2021. 610 p.

2. TNFD. Taskforce on Nature-related Financial Disclosures (TNFD) Recommendations on Nature-related Financial Disclosures [Internet]. 2023 Sep. Available from: https://tnfd.global/wp-content/uploads/2023/08/Recommendations_of_the_Taskforce_on_Nature-related_Financial_Disclosures_September_2023.pdf?v=1695118661

3. Thomas H, Chung YF, Maron M, Rhodes JR, Simmonds JS, Ward MS, et al. Achieving “nature positive” requires net gain legislation. Science. 2024 Oct 25;386(6720):383–5.

4. Hering D, Schürings C, Wenskus F, Blackstock K, Borja A, Birk S, et al. Securing success for the Nature Restoration Law. Science. 2023 Dec 15;382(6676):1248–50.

5. Finance for Biodiversity Foundation. Guide on biodiversity measurement approaches [Internet]. 2024 Feb. Available from: https://www.financeforbiodiversity.org/wp-content/uploads/Finance-for-Biodiversity_Guide-on-biodiversity-measurement-approaches_3rd-edition-1.pdf

6. Tools Catalogue – TNFD [Internet]. [cited 2025 Nov 11]. Available from: https://tnfd.global/assessment-guidance/tools-catalogue/

7. J. Lammerant, M. Starkey, A. De Horde, A.M. Bor, K. Driesen, G. Vanderheyden. Assessment of Biodiversity Measurement Approaches for Businesses and Financial Institutions. EU Business @ Biodiversity Platform; 2022. Report No.: Update report 4.

8. Granqvist E, Goodsell RM, Töpel M, Ronquist F. The transformative potential of eDNA-based biodiversity impact assessment. Curr Opin Environ Sustain. 2025 Apr;73:101517.

9. Katic PG, Cerretelli S, Haggar J, Santika T, Walsh C. Mainstreaming biodiversity in business decisions: Taking stock of tools and gaps. Biol Conserv. 2023 Jan 1;277:109831.

10. Kashyap S, Abela CM, Blum V, Crona B. Business and finance on a path towards meaningful biodiversity reporting? Curr Opin Environ Sustain. 2025 Dec 1;77:101588.

11. Granqvist E, Goodsell RM, Töpel M, Ronquist F. The transformative potential of eDNA-based biodiversity impact assessment. Curr Opin Environ Sustain. 2025 Apr;73:101517.

12. Chapman M, Goldstein BR, Schell CJ, Brashares JS, Carter NH, Ellis-Soto D, et al. Biodiversity monitoring for a just planetary future. Science. 2024 Jan 5;383(6678):34–6.

13. Iwaszkiewicz-Eggebrecht E, Łukasik P, Buczek M, Deng J, Hartop EA, Havnås H, et al. FAVIS: Fast and versatile protocol for non-destructive metabarcoding of bulk insect samples. PloS One. 2023;18(7):e0286272.

14. Desjonquères C, Villén-Pérez S, De Marco P, Márquez R, Beltrán JF, Llusia D. Acoustic species distribution models (aSDMs): A framework to forecast shifts in calling behaviour under climate change. Methods Ecol Evol. 2022;13(10):2275–88.

15. Oliver RY, Iannarilli F, Ahumada J, Fegraus E, Flores N, Kays R, et al. Camera trapping expands the view into global biodiversity and its change. Philos Trans R Soc Lond B Biol Sci. 2023 Jul 17;378(1881):20220232.

16. Lyu X, Li X, Dang D, Dou H, Wang K, Lou A. Unmanned Aerial Vehicle (UAV) Remote Sensing in Grassland Ecosystem Monitoring: A Systematic Review. Remote Sens. 2022 Feb 23;14(5):1096.

17. Troudet J, Grandcolas P, Blin A, Vignes-Lebbe R, Legendre F. Taxonomic bias in biodiversity data and societal preferences. Sci Rep. 2017 Aug 22;7(1):9132.

18. Hasan F, Nyström J, Andersson C, Da Silva A, Högström A, Granqvist E, et al. The business case for investing in biodiversity data [Internet]. 2025 [cited 2025 Nov 11]. Available from: https://ecoevorxiv.org/repository/view/8492/

19. Miller BL, Lombardo S, Rosenthal A, O'Shea T, Luers A, Lavista-Ferres JM, et al. Corporate Biodiversity Reporting Can Be Scaled With AI and Earth Observation—But Will Miss the Point Without Guidance From Conservation Scientists. Conserv Lett. 2025;18(5):e13153.

20. Apicella L, De Martino M, Quarati A. Copernicus User Uptake: From Data to Applications. ISPRS Int J Geo-Inf. 2022 Feb 9;11(2):121.

21. Pereira HM, Ferrier S, Walters M, Geller GN, Jongman RHG, Scholes RJ, et al. Essential Biodiversity Variables. Science. 2013 Jan 18;339(6117):277–8.

22. Miraldo A, Sundh J, Iwaszkiewicz-Eggebrecht E, Buczek M, Goodsell R, Johansson H, et al. Data of the Insect Biome Atlas: a metabarcoding survey of the terrestrial arthropods of Sweden and Madagascar. Sci Data. 2025 May 21;12(1):835.

23. Sundh J, Granqvist E, Iwaszkiewicz-Eggebrecht E, Manoharan L, Van Dijk LJA, Goodsell R, et al. HAPP: High-Accuracy Pipeline for Processing deep

metabarcoding data [Internet]. 2024 [cited 2025 Sep 2]. Available from: http://biorxiv.org/lookup/doi/10.1101/2024.12.20.629441

24. Hillebrand H, Blasius B, Borer ET, Chase JM, Downing JA, Eriksson BK, et al. Biodiversity change is uncoupled from species richness trends: Consequences for conservation and monitoring. Cadotte M, editor. J Appl Ecol. 2018 Jan;55(1):169–84.

25. Fleishman E, Noss R, Noon B. Utility and limitations of species richness metrics for conservation planning. Ecol Indic. 2006 Aug;6(3):543–53.

26. Legendre P, De Cáceres M. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. Morlon H, editor. Ecol Lett. 2013 Aug;16(8):951–63.

27. Legendre P. Interpreting the replacement and richness difference components of beta diversity. Glob Ecol Biogeogr. 2014 Nov;23(11):1324–34.

28. Gross N, Bagousse-Pinguet YL, Liancourt P, Berdugo M, Gotelli NJ, Maestre FT. Functional trait diversity maximizes ecosystem multifunctionality. Nat Ecol Evol. 2017 Apr 18;1(5):0132.

29. Gagic V, Bartomeus I, Jonsson T, Taylor A, Winqvist C, Fischer C, et al. Functional identity and diversity of animals predict ecosystem functioning better than species-based indices. Proc R Soc B Biol Sci. 2015 Feb 22;282(1801):20142620.

30. Laureto LMO, Cianciaruso MV, Samia DSM. Functional diversity: an overview of its history and applicability. Nat Conserv. 2015 Jul;13(2):112–6.

31. Ronquist F, Forshage M, Häggqvist S, Karlsson D, Hovmöller R, Bergsten J, et al. Completing Linnaeus's inventory of the Swedish insect fauna: Only 5,000 species left? Cerretti P, editor. PLOS ONE. 2020 Mar 4;15(3):e0228561.

32. Rainford JL, Hofreiter M, Mayhew PJ. Phylogenetic analyses suggest that diversification and body size evolution are independent in insects. BMC Evol Biol. 2016 Dec;16(1):8.

33. Moretti M, Dias ATC, De Bello F, Altermatt F, Chown SL, Azcárate FM, et al. Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. Fox C, editor. Funct Ecol. 2017 Mar;31(3):558–67.

34. Gower JC. Some distance properties of latent root and vector methods used in multivariate analysis. Biom 53 325–338. 1966;

35. Maire E, Grenouillet G, Brosse S, Villéger S. How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. Glob Ecol Biogeogr. 2015 Jun;24(6):728–40.

36. Grenié M, Gruson H. fundiversity: a modular R package to compute functional diversity indices. Ecography. 2023 Mar;2023(3):e06585.

37. C3S. ERA5 hourly data on single levels from 1940 to present [Internet]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS); 2018 [cited

2025 Dec 4]. Available from:
https://cds.climate.copernicus.eu/doi/10.24381/cds.adbb2d47

38. Buchhorn M, Smets B, Bertels L, Roo BD, Lesiv M, Tsendbazar NE, et al. Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2018: Globe [Internet]. Zenodo; 2020 [cited 2025 Nov 23]. Available from: https://zenodo.org/record/3518037

39. Wood SN. Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models. J R Stat Soc Ser B Stat Methodol. 2011 Jan 1;73(1):3–36.

40. Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mitchell, Ignacio Cano, Tianyi Zhou,, Mu Li, Junyuan Xie, Min Lin, Yifeng Geng, Yutian Li and Jiaming Yuan. xgboost: Extreme Gradient Boosting [Internet]. 2023. Available from: https://CRAN.R-project.org/package=xgboost

41. Baggström A, Goodsell R, Van Dijk L, Iwaszkiewicz-Eggebrecht E, Miraldo A, Tack AJM, et al. The utility of combining deep learning with metabarcoding to model biodiversity dynamics at a national scale. Ecol Inform. 2025 Dec;90:103318.

42. Carvalho F, Brown KA, Gordon AD, Yesuf GU, Raherilalao MJ, Raselimanana AP, et al. Methods for prioritizing protected areas using individual and aggregate rankings. Environ Conserv. 2020 Jun;47(2):113–22.

43. Du Z, Li L, Liang J, Kang B, Meng W, Li H. National scale biodiversity conservation priorities based on integrated multiple vulnerability features in China. Ecol Indic. 2024 Mar;160:111914.

44. McGeoch MA, Genovesi P, Bellingham PJ, Costello MJ, McGrannachan C, Sheppard A. Prioritizing species, pathways, and sites to achieve conservation targets for biological invasion. Biol Invasions. 2016 Feb 1;18(2):299–314.

45. Wilson KA, Carwardine J, Possingham HP. Setting Conservation Priorities. Ann N Y Acad Sci. 2009 Apr;1162(1):237–64.

46. Guidelines for species conservation planning : version 1.0 [Internet]. IUCN; 2017 [cited 2025 Dec 4]. Available from: https://portals.iucn.org/library/node/47142

47. GOV.UK [Internet]. [cited 2025 Dec 4]. Annex 3: Scoring Higher Tier Applications. Available from: https://www.gov.uk/government/publications/countryside-stewardship-higher-tier-manual-for-agreements-starting-on-1-january-2022/annex-3-scoring-higher-tier-applications

48. GOV.UK [Internet]. [cited 2025 Dec 4]. Priority open habitats and woodland creation: A field guide. Available from: https://www.gov.uk/government/publications/priority-open-habitats-and-woodland-creation-a-field-guide

49. Example TNFD reporting – TNFD [Internet]. 2025 [cited 2025 Oct 6]. Available from: https://tnfd.global/knowledge-hub/example-tnfd-reporting/

50. Ferrier S, Guisan A. Spatial modelling of biodiversity at the community level. J Appl Ecol. 2006 Jun;43(3):393–404.

51. Benito BM, Cayuela L, Albuquerque FS. The impact of modelling choices in the predictive performance of richness maps derived from species‐distribution models: guidelines to build better diversity models. O'Hara RB, editor. Methods Ecol Evol. 2013 Apr;4(4):327–35.

52. Businesses sharing biodiversity data via GBIF [Internet]. [cited 2025 Nov 24]. Available from: https://www.gbif.org/composition/1XtRfS0nTKs8HtRd18Q7ai/businesses-sharing-biodiversity-data-via-gbif

53. Chudy R, Barton DN. Biodiversity risk screening tools in finance fail to meet the need for local project risk screening. J For Bus Res. 2025 Sep 17;4(2):47–64.

54. PCAF. The Global GHG Accounting and Reporting Standard for the Financial Industry. 2022. Report No.: Second Edition.

55. Schipper AM, Hilbers JP, Meijer JR, Antão LH, Benítez-López A, de Jonge MMJ, et al. Projecting terrestrial biodiversity intactness with GLOBIO 4. Glob Change Biol. 2020;26(2):760–71.

56. Biodiversa + [Internet]. [cited 2025 Nov 24]. A business guide to sharing biodiversity data. Available from: https://www.biodiversa.eu/2025/10/27/a-business-guide-to-sharing-biodiversity-data/

Supplementary Materials

Part 1: Summary of corporate report analysis

Scope: We only considered datasets and tools used for nature-related risk and impact assessments, excluding those related to climate assessments. Furthermore, we only included datasets and tools that were used systematically within company-wide or portfolio-wide risk assessments, excluding those used in individual projects or presented as part of risk or impact management processes. We also excluded risk assessment frameworks and any mentions of bespoke tools or frameworks developed by commercial consultancies.

Global datasets and tools used for risk/impact assessment	State of biodiversity	Corporates	Financial institutions	Corporates	Financial institutions	Total
ENCORE	No	67%	93%	38	25	63
WRI Aqueduct	No	40%	11%	23	3	26
IBAT	Yes	25%	22%	14	6	20
WWF Biodiversity Risk Filter	Yes	25%	15%	14	4	18
SBTN Materiality Screening Tool	No	16%	15%	9	4	13
IUCN Red List	Yes	19%	0%	11	0	11
WWF Water Risk Filter	No	16%	4%	9	1	10
Global Forest Watch	Yes	12%	4%	7	1	8
World Database on Protected Areas	No	12%	0%	7	0	7
World Database on Key Biodiversity Areas	Yes	9%	4%	5	1	6
SBTN High Impact Commodity List	No	7%	4%	4	1	5
Biodiversity Intactness Index	Yes	5%	7%	3	2	5
Global Biodiversity Score	No	4%	4%	2	1	3
Ecosystem Services Review	No	5%	0%	3	0	3
GLOBIO	No	0%	11%	0	3	3
ForestIQ	No	0%	11%	0	3	3
RESOLVE Ecoregions and Biomes	Yes	4%	0%	2	0	2
Ecologically or Biologically Significant Marine Areas	Yes	4%	0%	2	0	2
Forest500	No	0%	7%	0	2	2
ZSL SPOTT	No	0%	7%	0	2	2
Other global tools/datasets	Yes/No	56%	56%	32	15	47

Table S1: Relative and absolute counts of individual corporate reports that mention specific global datasets and tools used to inform company or portfolio wide risk or impact assessments. Split between corporates (total report count 57) and financial institutions (total report count 27).

Both the Integrated Biodiversity Assessment Tool (IBAT) and the WWF Biodiversity Risk Filter incorporate and combine several state of biodiversity datasets. IBAT includes the IUCN Red List, World Database on Protected Areas, World Database of Key Biodiversity Areas, Species Threat Abatement and Restoration metric, and Rarity-Weighted Species Richness. WWF Biodiversity Risk Filter includes the Biodiversity Intactness Index, World Database of Protected Areas, World Database of Key Biodiversity, and more. For a full list see: https://panda.maps.arcgis.com/sharing/rest/content/items/60743c9fff314ba1b4bd2564401764e2/data

Internal, local or national datasets used for risk/impact assessment	Corporates	Financial institutions	Corporates	Financial institutions	Total
Restoration, conservation, and habitat management data layers (state/national/regional)	23%	19%	13	5	18
Community, supplier and internal surveys	19%	7%	11	2	13
Impact assessments and environmental audits	19%	0%	11	0	11
(Operational) environmental pressure monitoring	14%	0%	8	0	8
Ecological monitoring	12%	0%	7	0	7
Water resource monitoring	11%	0%	6	0	6
Remote sensing/monitoring	5%	4%	3	1	4
No internal, local or national datasets mentioned	39%	74%	22	20	42

Table S2: Relative and absolute counts of individual corporate reports that mention specific categories of internal, local or national datasets and tools used to inform company or portfolio wide risk or impact assessments. Split between corporates (total report count 57) and financial institutions (total report count 27).

Restoration, conservation and habitat management data layers (state/national/regional): Third party data layers from Data from restoration projects and biodiversity management plans

Community, supplier and internal surveys: Data collected via internal surveys, supplier surveys, community engagement or other participatory processes.

Impact assessments and environmental audits: Data collected as part of formal environmental and social impact assessments, biodiversity impact assessments, environmental audits at company or supplier sites etc. conducted before or during projects, construction, etc.

(Operational) environmental pressure monitoring: Data from ongoing monitoring of pressures (e.g. emissions, waste, pollution, land disturbance, etc.) from operational sites.

Ecological monitoring: Data from ongoing monitoring of biodiversity, habitats or ecological conditions around operational sites.

Water resource monitoring: Data on water withdrawal, quality, etc. collected at or near operational sites.

*Remote sensing/monitorin*g: Data from remote sensing platforms (e.g. satellites, aircraft) to monitor environmental changes on the ground.

No internal, local or national datasets mentioned: No mentions for a systematic assessment of nature-related risks or impacts across the company or portfolio(s). Any reports mentioning the collection of internal, in-situ or local data collection for an individual site or project or for risk management and operational monitoring purposes are also covered in this category.

Use case	Corporates	Financial institutions	Corporates	Financial institutions	Total
Screening for potential impacts and dependencies	75%	85%	43	23	66
Identifying priority locations (impacts)	65%	33%	37	9	46
Identifying priority locations (dependencies)	51%	30%	29	8	37
Identifying priority engagement targets/topics	0%	41%	0	11	11
Supply chain due diligence and engagement	5%	0%	3	0	3
Quantifying supply chain risk/impact	4%	4%	2	1	3
Scenario analysis	2%	7%	1	2	3
Monitoring/tracking performance or progress (against targets)	2%	4%	1	1	2
Benchmarking impacts and dependencies	0%	7%	0	2	2
Develop commercial strategy and customer support options	0%	7%	0	2	2
Investment integration	0%	4%	0	1	1

Table S3: Relative and absolute counts of individual corporate reports that mention specific use case categories for the global, national, local or internal datasets and tools used to inform company or portfolio wide risk or impact assessments. Split between corporates (total report count 57) and financial institutions (total report count 27).

Screening for potential impacts and dependencies: High level mapping of impacts and dependencies across all activities within the company, supply chain or portfolio to identify which activities are relatively more/less dependent on nature or contributing to pressures.

Identifying priority locations (impacts): Singling out individual sites, activities or companies based on their potential impacts, for further in depth analysis.

Identifying priority locations (dependencies): Singling out individual sites, activities or companies based on their potential dependencies, for further in depth analysis.

Identifying priority engagement targets/topics: Singling out individual companies or activities based on their potential impacts, for further in depth analysis.

Supply chain due diligence and engagement: Using tools and datasets to screen/audit suppliers or identify priority suppliers for engagement.

Quantifying supply chain risk/impact: Using tools or datasets to quantify dependencies, pressures, risks or impacts within the own operations or supply chains typically for the purpose of reporting. These can be both negative or positive e.g. from restoration efforts.

Scenario analysis: Using tools or datasets to project future risks based on different scenarios or pathways.

Monitoring/tracking performance or progress (against targets)

Benchmarking impacts and dependencies: Comparing impact and dependencies within the own portfolio against a benchmark portfolio.

Developing commercial strategy and customer support options

Investment integration: Integrating nature-related considerations into investment decision-making, including divestments

Detailed corporate report analysis

The full list of reports reviewed for this analysis is available on GitHub (*https://github.com/ronquistlab/finbio-ebv*). It contains information about the reporting entity, type of report, year published, sustainability accounting standards board (SASB) sector, organisation type and headquarter country. For each report it maps the global datasets and tools used for risk/impact assessment; in-situ, local or national datasets used for risk/impact assessment; and the use cases for the datasets.

Part 2: Model fitting

Tables of hyperparameter values used for during the optimisation step (**S4 & S5)**, and the final hyperparameters used for each model **(S6**).

Table S4. Stage 1 Hyperparameters test values

Hyperparameter	Meaning	Values Tested
Learning rate (Eta)	Controls how fast the model learns. Smaller values make learning slower but more accurate.	Twenty values between 0.01:0.300
Number of boosting rounds (Nrounds)	Total number of trees added to the model. More rounds increase modelling power but risk overfitting.	100, 300, 500, 1000,1200

Table S5. Stage 2 Hyperparameters test values

Hyperparameter	Meaning	Values Tested
Max tree depth (Max Depth)	Maximum depth of each tree. Higher depth captures more interactions but increases overfitting risk.	3:12
Data subsample rate (SubSample)	Fraction of total data sampled for each tree	0.8, 1.0
Column subsample rate (ColSample)	Fraction of predictors sampled for each tree. Lower = more regularisation, reduces correlation between trees.	0.25, 0.35, 0.50, 0.75
L2 regularisation (Lambda)	Ridge penalty on leaf weights. Helps stabilise large weights and reduces overfitting.	0.1, 0.5, 1.5, 3.5
L1 regularisation (Alpha)	Lasso penalty on leaf weights. Encourages sparse trees and feature selection.	0.1, 0.5, 1.5, 3.5

Table S6. Final hyperparameters used for model fits

Country	EBV	Max Depth	Eta	Nrounds	ColSample	SubSample	Lambda	Alpha
Madagascar	FD	7	0.01	2000	0.35	0.8	0	0
	FE	6	0.056	500	0.35	0.8	3.5	0.5
	GSH	11	0.056	2000	0.35	0.8	3.5	1.5
	LCBD	11	0.1	1500	0.5	1	3.5	0
	SR	6	0.071	1000	0.5	0.8	0	0.5
Sweden	FD	10	0.025	1000	0.35	1	3.5	0
	FE	6	0.071	1500	0.35	0.8	3.5	0
	GSH	11	0.041	1000	0.5	1	1.5	1.5
	LCBD	12	0.12	1500	0.75	1	0.5	0
	SR	4	0.086	2000	0.5	1	1.5	0.5

Part 3. Assessing model fit

To assess the fit of individual XGboost models we plot observed vs predicted values for each model fit. Figure S1 displays these values for each EBV and country combination. Overall each model displays good levels of fit (i.e. high concordance between observed and predicted values. Some models do appear to be overfit as concordance between these values are so high. However as we tuned hyperparameters for out-of-sample prediction, the predictive performance of these models is still acceptable.

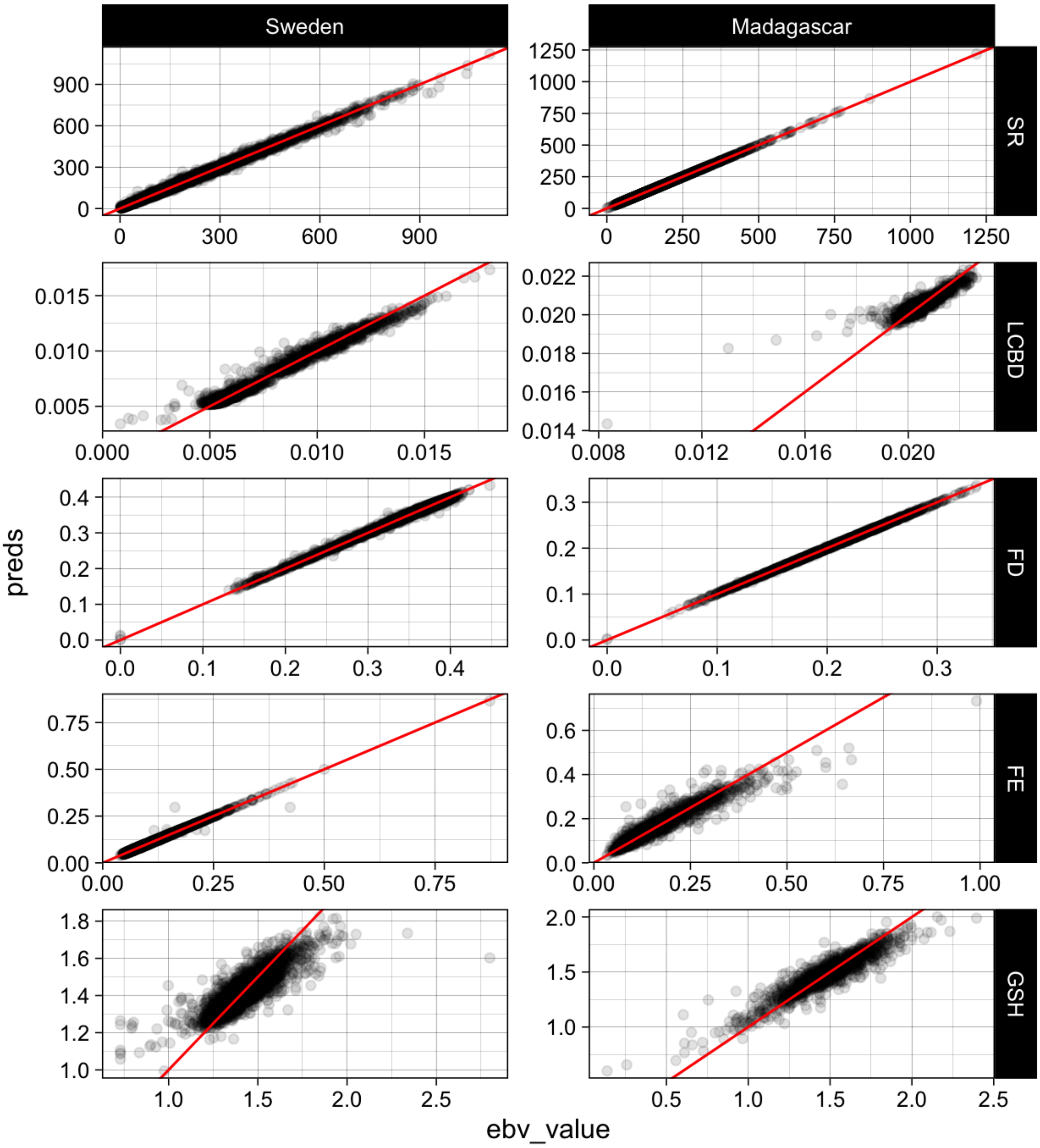

Figure S1. Observed vs predicted values of EBVs for each country for within sample prediction. The red lines represent a 1:1 relationship.