Biodiversity Indicators Review - International Climate Finance Evidence Project

Harris, M.[1], Woodcock, P.[1], Wright, E.[1], Britton, A.[1], Matear, L.[1], Smith, A.[1], Vina-Herbon, C.[1], and Cheffings, C.[1]

May 2021

For further information on JNCC's report series please contact:
Joint Nature Conservation Committee
Monkstone House
City Road
Peterborough PE1 1JY
https://jncc.gov.uk/
Communications@jncc.gov.uk

This report was produced by JNCC for Defra through the International Climate Finance Evidence Project.

This report should be cited as:
Harris, M.[1], Woodcock, P.[1], Wright, E.[1], Britton, A.[1], Matear, L.[1], Smith, A.[1], Vina-Herbon, C.[1], and Cheffings, C.[1]. Biodiversity Indicators Review - International Climate Finance Evidence Project. JNCC, Peterborough.
https://hub.jncc.gov.uk/assets/376d989f-0563-4e7f-b034-c79108f63758#biodiversity-indicators-review.pdf

[1] JNCC, Monkstone House, City Road, Peterborough, PE1 1JY

Evidence Quality Assurance:

This report is compliant with JNCC's Evidence Quality Assurance Policy
https://jncc.gov.uk/about-jncc/corporate-information/evidence-quality-assurance/

Summary

The UK Government's International Climate Finance (ICF) makes use of indicators in their annual publications to set out results of any impacts from their portfolio of investments. These indicators currently cover social and climate metrics well, but do not yet incorporate information on the biodiversity impacts of projects that are funded. A major emphasis to support the triple-win is the focus on funding nature-based solutions (NbS) and ICF recognises the potential for these projects to contribute to a 'triple win.' It is therefore necessary for any biodiversity indicators adopted by ICF to be applicable to NbS projects.

Many biodiversity indicators already exist. To recommend effective indicators that could be used by ICF, it is first valuable to understand indicators that have been designed and implemented previously. This task therefore aimed to review and evaluate existing biodiversity indicators from major biodiversity frameworks, and identified in the NbS database in order to:

- Identify different types of biodiversity indicator that are already implemented, particularly considering which aspects of biodiversity are measured and which different approaches can be taken to measure biodiversity.
- Review and evaluate the design and implementation of the indicators, where information was available.
- Summarise the strengths, weaknesses, important characteristics of and likely effort required to calculate different types of indicator identified, in relation to ICF.
- Provide a short summary of the most relevant specific biodiversity indicators identified within the review for their potential relevance to ICF investment in NbS projects.

Contents

1 Introduction

Biodiversity, or biological diversity, describes the variation in life and living organisms. The Convention on Biological Diversity (CBD) defines it as "*the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems*." (CBD, 2006). Each of these aspects of biodiversity are interlinked. For example, low ecosystem diversity could lead to fewer ecological niches and therefore lower species diversity. However, the components do not necessarily correlate well and to have a holistic understanding of biodiversity each component would need to be monitored independently. For instance, an area that has a high species richness (the number of species present) would not necessarily have high genetic diversity within the population of each species present, and to understand this each aspect would need to be assessed. Biodiversity is therefore a broad and complex concept, making it difficult to capture in a single indicator.

Indicators are measurements across space and time that are used to report the state or identify change in a variable of interest, such as biodiversity. Indicators can be useful to summarise trends and communicate data to a wide range of audiences. They are not designed to incorporate all information possible about the variable of interest but are intended to be indicative of wider changes (JNCC, 2020). Indicators can be used to monitor progress against targets and to communicate results from monitoring or evaluation exercises. Good indicators should be scientifically valid (widely accepted and based on reliable data), responsive (picking up on changes within the variable of interest), easy to understand (both conceptually and in terms of presentation and interpretation), based on data (sampled/observational data or modelling) that are already available or are feasible to collect (which will allow for regular production of the indicator over time) and relevant to the needs of those who will use them (BIP, 2011).

Indicators reporting on biodiversity can make use of various types of data. Some measure the **state** of one or more aspects of biodiversity, such as counts of species richness or abundance. Others measure the extent of a known **pressure** on biodiversity, such as deforestation. Another approach is to use data on how people or organisations are **responding** to the challenge of biodiversity conservation, for example using metrics such as the amount of money spent on relevant projects. Indicators also vary in whether they rely on **direct** ecological measurements in the field, are based on **proxies**, or are **modelled** from multiple other datasets.

The UK Government's International Climate Finance (ICF) makes use of indicators in their annual publications to set out results of any impacts from their portfolio of investments. These indicators, known as Key Performance Indicators (KPIs) currently cover social and climate metrics well, but do not yet incorporate information on the biodiversity impacts of projects that are funded. Inclusion of biodiversity indicators would help ICF understand its progress towards its aspiration to provide 'a triple win for people, climate and nature'. With the conservation of biodiversity a rising priority on the international policy agenda, it could also help the UK Government to show alignment with and progress towards other international biodiversity targets and conventions, such as the Sustainable Development Goals and the Convention on Biological Diversity's planned post-2020 Framework.

A major emphasis to support the triple-win is the focus on funding nature-based solutions (NbS) and ICF recognises the potential for these projects to contribute to a 'triple win.' It is therefore necessary for any biodiversity indicators adopted by ICF to be applicable to NbS projects. While there are varying definitions of NbS, for the purposes of this report, we consider NbS to be "*actions that enlist elements of nature or natural processes to address* *particular problems faced by society*". It is easy to assume that NbS are inherently beneficial for the conservation of biodiversity, however, if poorly implemented or not specifically designed with biodiversity in mind (such as lack of consideration for native species), NbS can have negligible or even negative impacts on biodiversity. Biodiversity indicators would therefore be an important addition to NbS projects to ensure that biodiversity impacts are understood and addressed, and to facilitate reporting on the scale of benefits achieved by the ICF portfolio.

Many biodiversity indicators already exist. To recommend effective indicators that could be used by ICF, it is first valuable to understand indicators that have been designed and implemented previously. This can help ensure indicator development takes advantage of the wealth of information already available, including an understanding of the relative advantages and disadvantages of different approaches in the ICF context. This task therefore aimed to review existing biodiversity indicators in order to:

- Identify different types of biodiversity indicator that are already implemented, particularly considering which aspects of biodiversity are measured and which different approaches can be taken to measure biodiversity.
- Review and evaluate the design and implementation of the indicators, where information was available.
- Summarise the strengths, weaknesses, important characteristics of and likely effort required to calculate different types of indicator identified, in relation to ICF.
- Provide a short summary of the most relevant specific biodiversity indicators identified within the review for their potential relevance to ICF investment in NbS projects.

Note that the review is intended to be sufficiently comprehensive to address the above questions but was time-limited and is not exhaustive. The purpose of the review was to inform the choices of indicators in the ICF context, but critically there is a need to define the requirement for potential KPIs, carefully considering factors such as what is the scale of change that is anticipated at a strategic level and assessing the feasibility of detecting this sort of change. The review considers all ecosystems but there is a terrestrial bias in the indicators reviewed, so for clarity additional sections considers marine indicators, and the terrestrial subset of urban indicators in isolation. A summary of the review is presented in Concluding remarks and summary.

2 Methods

2.1 Review Scope and Search Methods

Two previous tasks within this project reviewed biodiversity indicator frameworks and a list of NbS case studies. Frameworks identified as significant and case studies identified as containing a biodiversity indicator were used to create a list of indicators for review in the current task. Each of these indicators was assessed for whether it could be applicable at the project scale, whether it could be aggregated to portfolio level and whether it defined a method that was prescriptive enough that guidance could be developed for an NbS project to be able to make use of it in a clear and repeatable manner. Where all of these criteria were met, information was extracted. The priority was on understanding the spectrum of indicators currently available, so information was not extracted from indicators that duplicated, were very similar to or were simply a disaggregation of an indicator already reviewed, or that measured something that would not affect biodiversity directly (e.g. people's perceptions of biodiversity). The review covered indicators used for terrestrial (including freshwater and urban environments) and/or marine ecosystems.

2.2 Data extraction

For indicators that passed the criteria for data extraction, the following data were recorded where information was available:

i. The framework the indicator formed a part of
ii. Whether the indicator was designed for terrestrial, urban, marine or all environments
iii. Whether the indicator was based around species, ecosystems, ecosystem services or another theme
iv. Whether the indicator was measuring a pressure, a state or a response
v. Links to key resources
vi. The scale at which the indicator is implemented in its current framework (e.g. project scale, national scale), and its applicability, or potential applicability, for reporting or aggregation at the ICF portfolio scale
vii. Data availability and any data collection that would be required if implemented as an indicator for an ICF project
viii. How applicable the indicator would be to NbS projects
ix. The taxonomic breadth that the indicator covers
x. Whether it is generic or aimed at (a) specific ecosystem type(s)
xi. The baseline the indicator is compared against
xii. The temporal scale of the indicator (e.g. gains since the project started, annual progress, difference to baseline)
xiii. The frequency of reporting of the indicator
xiv. Practices for ensuring data quality
xv. Whether the indicator takes net effects into consideration (i.e. including both positive and negative impacts)
xvi. What the indicator is measuring
xvii. Output units
xviii. A subjective assessment of the level of expertise that would be required to assess against the indicator
xix. Any risks or problems with the indicator that can be identified
xx. A summary of the indicator

Note that in practice several of the above data extraction categories involved some subjective judgement. A spreadsheet of extracted information is available upon request.

2.3 Synthesising review

Indicators from the review were grouped according to (i) whether these measured biodiversity in terms of state, pressure, or response, and (ii) whether the indicator was a direct, proxy, or modelled measure of biodiversity. The synthesis assesses the strengths and weaknesses of these different groups of indicators and their relevance to ICF investments. The second part of the synthesis examined points that will be important for biodiversity KPIs to address and how these are tackled by existing indicators, as well as the main distinctions between headline portfolio indicator and site-specific project indicators and considerations for applying indicators consistently across different types of ecosystem. Specific indicators mentioned within this report are summarised for further information in Annex A.

3 Synthesis

We considered 460 indicators, covering 25 frameworks and including large international policy frameworks, as well as smaller more project-specific indicators from frameworks and from the NbS case study database. From these, 154 were sufficiently relevant, not duplicated and with sufficient levels of information available for data extraction. Indicators covered all ecosystems, although there was a bias towards terrestrial habitats.

3.1 What do indicators measure?

Most indicators reviewed were either measuring the state of biodiversity, known pressures on biodiversity, or people's responses that aimed to improve biodiversity.

3.1.1 State indicators

State indicators measure and describe how the components of biodiversity (ecosystems, species, genes) are changing. This is an important way of monitoring biodiversity impacts because one of the main ultimate ambitions of conservation policy and funding is to improve the state of ecosystems, species, and genetic diversity. As a result, state indicators are widely used in NbS projects and in biodiversity frameworks. However, measuring changes in the state of biodiversity does not necessarily inform on *why* it is changing. This may present difficulties for NbS projects in attributing any changes in biodiversity directly to the projects (i.e. output indicators rather than outcome indicators); it is possible that biodiversity could change for reasons unrelated to the NbS taking place. In addition, some aspects of biodiversity change slowly, and so state indicators measured during a short-term project (e.g. 3-5 years) may not fully capture the expected longer-term impact of the project on biodiversity. Nearly all of the indicators reviewed focused on the state of biodiversity either at the species or the habitat level, with very few documenting ways of understanding state from a genetic perspective. An illustrative list of a subset of state indicators is presented in Table 1.

Species-level indicators can use a wide range of taxa and or can focus on a few individual species. The former often involve taking counts of species or taxa richness, or calculating well-known diversity metrics such as the Shannon's diversity index[1] or the Simpson's diversity index[2]. These indicators potentially provide very relevant information on an important aspect of biodiversity, but can also involve substantial resources for data collection if monitoring is required at the project level. Some of the species-level indicators of state proposed globally attempt to tackle this problem using modelling and/or globally available datasets (e.g. Living Planet Index, Local Biodiversity Intactness Index) – although typically designed for national reporting, these might still produce applicable metrics at the project level.

Species-level indicators that focus on a select few individual species (including flora and fauna) are sometimes used where these species are viewed as proxies for other aspects of state (i.e. assuming their abundance would reflect either habitat quality and/or populations of other species within the ecosystem), although specific rare or charismatic species are also sometimes monitored because they are of direct interest (e.g. if conserving these species is

---

[1] Example of a project that made use of the Shannon diversity index: http://nrcsolutions.org/green-city-clean-waters-program-philadelphia-pennsylvania/

[2] Example of a project that made use of the Simpson's diversity index: https://panorama.solutions/en/solution/resilient-rural-livelihoods-through-eco-restoration-and-sustainable-natural-resources

a particular goal of the project). This approach of focusing on a smaller number of taxa can help ensure feasibility if monitoring is carried out by individual projects directly in the field (particularly as the species used are often relatively easy to recognise and count) and so is popular across the NbS case studies that included biodiversity monitoring. For example, indicators to assess the impact of a Green Climate Fund project in Bhutan used tigers and snow leopards as headline species, arguing that their status at the top of the food chain means their populations must, to a certain extent, reflect health of the ecosystem as a whole. The main disadvantages of this approach are the representativeness of the indicator for other species and the potential difficulty of aggregating information on different indicator taxa across projects and countries. One example intermediate between this use of a small number of selected taxa and the more comprehensive species-level measures is the wildlife picture index, which is included in international policy frameworks and uses camera trap data to assess 278 species of tropical mammals and birds.

Habitat-level indicators often measure the extent (or changes in extent) of various habitat types. For many habitat types, data on extent can be obtained through satellite imagery and global datasets, requiring relatively low levels of expertise or survey effort from the project team to measure (although they could be supplemented with more detailed data from the project team where appropriate and available). This makes them a particularly attractive choice for monitoring biodiversity impacts.

Other indicators focused on habitat condition, degradation or connectivity. Condition and degradation typically describe how close a habitat is to a defined 'preferred' state, and so can be very informative for understanding how well a project or programme contributes to maintaining or restoring the overall functioning and structure of an ecosystem. However, these measures are often noted as particularly subjective and difficult to measure consistently[3]. Condition indicators are also more likely to be habitat specific (e.g. coral reef condition), posing an additional challenge for aggregating across projects. One of the main alternatives to measuring condition directly is to consider activities that that are expected to improve condition. For example, habitat restoration is expected to improve condition and so the area of habitat restored could be used to understand the impacts of some types of project, as well as linking with broader CBD targets on habitat restoration.

While indicators were found for a wide variety of habitat types, forest was by far the most common habitat-based indicator found in the frameworks and case studies reviewed, suggesting a potential need for greater knowledge and emphasis on other habitat types. Given the importance in protecting the diversity of ecosystems themselves it is critical to have representation of different ecosystems within biodiversity indicators, and in particular rare ecosystems may be a unique consideration where the ecosystem itself is threatened with disappearance (Keith et al. 2015). Use of an indicator focusing on only one particular habitat could lead to perverse incentives, whereby the habitat type selected as the indicator is protected at the expense of other, potentially rarer or more imperilled, habitat types.

A limitation of habitat-based metrics is that using extent gives no information about the condition or connectivity of the habitat, and vice versa. For instance, if focusing only on forest extent as an unqualified indicator, it could be possible to obtain a good indicator score by replanting an area with a monoculture plantation of non-native species that displace or do not support native species well. Similarly, an NbS attempting to improve management practices within an area of currently degraded habitat could have substantial effects on biodiversity, but these would not be apparent in an extent-based indicator. Focusing instead only on condition would miss any reduction in extent of the habitat so long as the remaining

[3] Example of an indicator of habitat degradation that acknowledges the context specific nature and complexity of measuring this: https://unstats.un.org/sdgs/metadata/files/Metadata-15-03-01.pdf

area was in good condition. Any habitat-based metric therefore needs to be well-qualified and should not be interpreted in isolation.

Very few indicators are in use for monitoring state at the level of genetic diversity, and those that are in place focus on domesticated crops and animals (Hoban et al., 2020). In part, this probably reflects a combination of challenges that are conceptual (defining suitable genetic diversity indicators) and practical (the technology required for monitoring). However, progress is being made in both areas and the recent review by Hoban et al. (2020) proposes three potential genetic diversity indicators that could be suitable for the CBD. As such, although immediate incorporation of genetic diversity indicators into ICF evaluation is probably not feasible, it will be important to consider how future methodological and conceptual developments in this area can be captured.

Table 1 Examples of state indicators. 'Where used' is indicative of major frameworks: CBD - Aichi Biodiversity Target Indicators; SDG - Sustainable Development Goals; EKLIPSE, N4C – Nature 4 Cities; LSA – LandScale Assessment Framework; RSC – Regional Seas Conventions; Other (specified).

Indicator	Where used	Main strength (+) /weakness (-)
Native/indigenous species richness	EKLIPSE	+ Can be a good indication of species diversity in an area -- Would require expertise and field measurements -- Does not take into account population sizes of each species, only presence -- Difficult to aggregate to portfolio level -- Risk of sampling bias
Shannon and Simpson's diversity indices - No. of species present, their relative abundances & their evenness	N4C	+ Can be good indications of species diversity in or between areas -- Would require expertise and field measurements -- Difficult to aggregate to portfolio level -- Risk of sampling bias
Abundance of indicator species	LSA; common as project level indicator	+ Would require less resource and expertise than sampling all species present -- Would only act as a proxy for biodiversity -- Difficult to aggregate to portfolio level -- Risk of sampling bias
Extent (area) of natural habitat (e.g. forest, coral, etc)	CBD; LSA	+ Easily measured -- Difficult to define -- Does not take habitat quality into account
Proportion of land degraded over total land area	CBD; SDG	+ Does take habitat quality into account -- Difficult to define -- Does not capture differing degrees of degradation

3.1.2 Pressure indicators

Pressure indicators measure and describe how human impacts on biodiversity are changing. Policy ambitions often focus on reducing pressures as a means to facilitate biodiversity recovery (i.e. improvements in the state of biodiversity). As such, monitoring pressures can serve two purposes: (i) to assess how effective actions have been in reducing pressures, and (ii) as a proxy to understand the state of biodiversity. Examples of pressure-based biodiversity indicators include water quantity and quality, nitrogen pollution, the presence or introduction of invasive species, illegal killing, take and trade in wild species, and human appropriation of primary productivity.

Indicators that track progress in reducing pressures are usually pressure-specific. This can pose a problem where a diverse range of projects are being considered, because individual projects are likely to focus on different pressures. For example, an indicator of water quality would not be relevant to a project focusing on reducing illegal wildlife trade. Meaningful aggregation across the many different types of pressures that could be affected by NbS might therefore be challenging.

Despite this challenge, pressure indicators can have important advantages for monitoring NbS projects. Firstly, changes in pressures are sometimes easier to monitor and/or respond more rapidly than changes in state (e.g. a 3-year project might be able to document a reduction in illegal trade of wildlife but perhaps not how this has an impact on populations). Secondly, some NbS projects may have more indirect effects on biodiversity that would be otherwise hard to track. For example, the use of drought-resistant seed varieties or crop wild relatives probably would not have substantial on-site biodiversity benefits but may reduce pressures less directly (e.g. lower water abstraction for irrigation, reduced need for herbicides or pesticides). In this way, pressure indicators could be the best (or only) approach for capturing the biodiversity impacts of some NbS. Ultimately, an important factor in the use of pressure indicators would be to provide the context to communicate whether the pressure indicators are (a) important to measure and track in their own right, or (b) important as a means of providing meaningful information about the state of biodiversity.

An illustrative list of a subset of pressure indicators are presented in Table 2.

Table 2 Examples of pressure indicators. Where used is indicative of major frameworks: CBD - Aichi Biodiversity Target Indicators; SDG - Sustainable Development Goals; EKLIPSE, N4C – Nature 4 Cities; LSA – LandScale Assessment Framework; Other (specified).

Indicator	Where used	Main strength (+) /weakness (-)
IUCN STAR - Potential reduction in species extinction risk based on changing threats	LSA	+ Takes into account multiple pressures + Data available globally, with optional field data to improve specific application to a project -- Global scale data is available at 5 km^2 resolution, which may be too coarse for application to NbS projects
Ecotoxicology factor - Effects of pollutants on populations within an area of interest	N4C	+ Measures a factor with a clear impact on biodiversity -- Would require expertise and field data collection -- Only takes into account one pressure on biodiversity
Number of invasive alien species (or variations including no. of introduction events / no. of eradication events)	CBD; EKLIPSE; N4C	+ Measures a factor with a clear impact on biodiversity -- Would require expertise and field data collection -- Only takes into account one pressure on biodiversity
Number of instances of illegal activities (burning, illegal killing etc.)	Other – individual project case studies	+ Measures a factor with a clear impact on biodiversity + Records may already be collected -- Only illegal activities that are caught would be recorded -- Only takes into account a small subset of pressures on biodiversity

3.1.3 Response indicators

Response based indicators are those which measure factors relating to actions people have taken to improve biodiversity, for example developing a conservation plan or designating a protected area. Several such indicators, especially those that came from NbS case study projects, focused on the number of plans, policies, or institutional structures into which biodiversity had been integrated as part of the project. Some also measured the amount of money spent on biodiversity.

From an evaluation perspective, response indicators should require little time or expertise to assess, as the indicators simply rely on the existence of the response, not an assessment of how well it was carried out (i.e. it is likely to be an activity indicator rather than an outcome indicator). These kinds of indicator have a less direct link with biodiversity than the other indicators reviewed, as they do not require the response to be successful; only to have taken place. For example, if a project has developed a management plan to prevent the local extinction of a particular species, but this was poorly implemented, the indicator would show a success even if the aim was not achieved. Responses can lead to changes in biodiversity if these plans, policies or protected areas are well implemented, but the fact that they exist does not prove this.

Many other indicators, largely from international frameworks, also had a strong focus on understanding the extent of protected areas or designated key biodiversity areas. However, protected area-based metrics are likely to be of little relevance to NbS projects. An increase in protected area would likely result from legislative solutions rather than nature-based solutions. Where NbS may have significant benefits is in reducing pressures in or around protected areas, but their success in improving condition would not be captured in extent metrics.

Some response indicators used scores calculated based on compliance with criteria. In these cases, meeting the criteria is often a binary case of yes or no, so this does not give a quantifiable metric, and does not necessarily give an indication of the scale of impacts (a project including the addition of a 'green roof' may contribute the same to the indicator as a project restoring dozens of hectares of priority habitat). Often such indicators require a certain number of the criteria to be met for a good final score, which could lead to comparability issues across different projects. Where such indicators are used, great consideration must be given to which criteria are applied, to ensure the indicator is meaningful.

An illustrative list of a subset of response indicators are presented in Table 3.

Table 3 Examples of response indicators. Where used is indicative of major frameworks: CBD - Aichi Biodiversity Target Indicators; SDG - Sustainable Development Goals; EKLIPSE, N4C – Nature 4 Cities; LSA – LandScale Assessment Framework; RSC – Regional Seas Conventions; Other (specified).

Indicator	Where used	Main strength (+) /weakness (-)
Rio Marker on Biodiversity - ODA spend on biodiversity	Other - Cadre logique de la Cadre d'Intervention Transversal Biodiversité 2013-2016	+ Clear and easy to understand + Internationally agreed metric -- No information is given on how well the spend is actually improving biodiversity
No. of plans, policies, or institutional structures into which biodiversity has been integrated as part of the project	CBD; RSC	+ Data collection would be straightforward -- No information is given on how well the plans and policies are actually improving biodiversity
Number of seedlings planted	Other – individual project case studies	+ Clear and easy to understand + Data would probably be easily available -- No information is captured on how many seedlings survive to form a new habitat -- No information is given on which species were used (e.g. they could be planting invasive species)
Extent of protected areas	CBD; RSC; LSA	+ Widely used and globally accepted method -- No information is captured on how well managed these protected areas are -- Designating protected areas is not likely to be relevant to NbS

3.1.4 Other indicators

Some indicators measured factors such as people's perceptions of and attitudes towards biodiversity, amount of data available on biodiversity and the skills and capacity to monitor biodiversity. These are important in enabling biodiversity conservation. For example, it is difficult to conserve biodiversity if public attitudes are indifferent or if there is a lack of capacity to understand how and where it is changing. As a result, indicators on public perceptions, capacity etc. are included within several biodiversity frameworks. Although these types of indicators were largely excluded from the review because they do not describe a project or programme impact on biodiversity *per se*, they may be worthwhile to consider as important project-level goals and for alignment with international goals and targets.

3.2 What different approaches to producing indicators are there?

Most indicators reviewed were either measuring aspects of biodiversity directly or measuring factors that could be used as a proxy for biodiversity impacts. Others modelled biodiversity by making use of several different types of input data.

3.2.1 Direct indicators

Some of the indicators reviewed had directly measured the relevant aspects of species or ecosystems (e.g. through field data collection or remote sensing). If carried out correctly, this approach should give reliable information on the state of biodiversity. However, direct assessments rely on suitable data collection and pose challenges for monitoring NbS projects in practice.

Directly measuring biodiversity at the species level will generally involve field surveys. Whilst these can give a good picture of biodiversity and are carried out by some projects as part of evaluation, substantial resources and expertise may be required – particularly if the project area is large or there is a need for comprehensive information (e.g. species richness, Shannon's Diversity Index). This may present problems for their practical implementation within ICF projects. These kinds of indicators also risk sampling bias as it is not possible to count all species or all individuals within a species in any given project area or sample site.

In contrast, directly measuring ecosystems could use Earth Observation (EO) data (including freely available global satellite coverage where the area of interest is large enough for the pixel resolution to be meaningful). Indicators of habitat extent that use this approach might require no (or limited) site visits and so be feasible with limited resources. Direct measures of habitat condition are more challenging but potentially still possible to generate from EO data.

3.2.2 Proxy indicators

Proxy indicators make use of a factor that is likely to be linked to biodiversity to infer a correlation – if the proxy changes, it is assumed that biodiversity changes. Proxies are usually used when direct biodiversity measurements are difficult or resource-intensive to obtain and so can represent a pragmatic approach. However, all proxies are also imperfect to some degree. As such, if proxies are used it is important to properly understand their relationship with the underlying aspect of biodiversity to be represented. Proxies are probably most commonly considered at the species-level, because of the challenges of directly measuring this aspect of biodiversity. The three main proxies for understanding species-level biodiversity are habitat-level metrics, pressure metrics, and selected species.

Habitat metrics (extent, condition, etc) are potentially helpful proxy indicators for species (in addition to being important aspects of biodiversity in their own right), because of the relationship between species and habitat extent and quality. For example, many forest species cannot survive outside of forested environments, so if forest extent or quality is significantly reduced, species populations and diversity will also reduce. However, such metrics by no means give the full picture because other factors can influence species diversity without affecting habitat extent and quality (e.g. hunting).

Pressure metrics can be effective proxies for species where there is a relationship between the magnitude of the pressure and biodiversity (i.e. as pressures increase, biodiversity will decrease). For example, a reduction in hunting pressure might be used as a proxy if directly monitoring changes in species populations is difficult. Again however, relationships are assumed and there will generally be some uncertainty over how great the impact of a particular pressure on biodiversity is, or how significant each pressure (as measured separately) is relative to other pressures on biodiversity within a given area.

Another example of a proxy indicator is the use of selected species (e.g. top predators or habitat specialists in ecosystem). In some cases, these species are assumed to be representative of a wider range of taxa that would be difficult to measure directly. Selected species can also be used to infer habitat quality – e.g. habitat specialists might only be found in good quality habitat. While using selected species as proxies potentially requires lower levels of expertise and associated cost than more comprehensive direct measures, it may also present comparability issues when aggregating results to the portfolio level, as each project would need to select their own species based on the habitat and circumstances of their project area. As with other proxies, there needs to be a strong correlation between the selected species and the broader aspect of biodiversity they are intended to indicate.

3.2.3 Modelled indicators

Some of the indicators that were reviewed were based on modelling aspects of biodiversity by combining multiple data types. For example, some indicators put pressure-based metrics together with global scale data on factors such as species' global ranges and extinction risks. Examples of this include the IUCN STAR metric, the GEOBON Global Biodiversity Change indicators and Defra's Biodiversity 2.0 metric (see Annex A).

This approach has the advantage of making use of existing global scale data (in many cases on a 1 km grid), allowing for greater ease of calculation. At the same time, while still not a direct measurement of biodiversity, it goes much further than simply assuming that a pressure can act as a proxy, by linking pressures to their likely effects on biodiversity. In some cases, multiple pressures are combined in the models based on their relative significance, giving a more holistic picture than a single pressure metric alone.

Such metrics appear to be relatively new compared to metrics described in the previous two sections. As such, they are less well developed and tested, and indeed some explored are not yet operational, though notably they are already widely used in marine frameworks, for example, metrics used under the Extent of physical damage to marine habitats indicator which uses modelled pressure and habitat data combined with survey data or the Live coral cover, which uses remote sensing to calculate the percentage of hard corals of the surface area of the coral community.

By using a model that incorporates the effects of multiple pressure types, an indicator would be applicable to a wider range of NbS than a single pressure indicator would be. However, with many such options making use of global scale data, small scale NbS may not be picked up. For example, while large-scale reforestation or wetland construction could be picked up at a scale of 1 km^2, NbS such as verge planting, rain gardens and green manuring could not. Modelled indices can also be less straightforward to interpret and communicate than some other approaches (e.g. a % change in an index that combines multiple properties vs. a measure of how much habitat has been added) and depend on the strength of the underlying data and understanding of the system.

Table 4 Overview of different indicator types

Type of indicator	Examples	Advantages	Disadvantages
Direct	♦ Native species richness ♦ Shannon and Simpson's diversity indices ♦ Abundance of indicator species ♦ Extent of natural habitat	+ If measuring habitat extents, could rely on EO data and therefore not need expensive field data collection + If carried out correctly gives most accurate representation of biodiversity	- If measuring species, would require expensive field data collection - Risk sampling bias
Proxy	♦ Abundance of indicator species ♦ Number of instances of illegal activities (burning, hunting etc.) ♦ Rio Marker (spend on biodiversity)	+ In some cases, measuring proxies will be easier and require less expertise than carrying out direct measurements	- A correlation between the proxy factor and biodiversity is assumed rather than measured
Modelled	♦ Biodiversity Habitat Index ♦ Global Ecosystem Restoration Index ♦ Local Biodiversity Intactness Index ♦ Species Habitat Indices	+ Most are available globally on a 1 km grid so would not require local data collection + May take into account multiple factors so can give a more holistic picture than single metrics	- May be difficult to interpret results - Novel - 1 km is too coarse a scale for some NbS projects - Ultimately only as good as underlying field data and ecological understanding

3.3 Important points for consideration in indicator selection and development

Across all types of indicator (whether state, pressure, response, direct, proxy or modelled), a number of factors identified within the review are particularly important for an ICF KPI to capture or address. These include net effects, displacement, a defined baseline to allow for comparative assessment, a timescale of relevance to the project being carried out, prescriptive methods and appropriate quality assurance processes. It will also be important for an indicator to acknowledge its own limitations, such as the inability to distinguish between impacts of the project and chance events such as drought.

3.3.1 Net effects

Evaluating the net effects of an NbS project means taking into account all positive and negative impacts of an action, rather than only considering the positive impacts. For example, measuring the area of trees planted is not an example of an indicator addressing the net effects of the project; it would be possible that within a project area the same extent of forest has also been felled, leading to no overall change in the factor that the intervention was aiming to address. To get a more holistic perspective, an indicator that takes into account total tree cover change within the project area would be required to capture the project's net effects. Although net effects are not always easy to capture, almost twice as many indicators within the review were found to account for net effects compared to those that did not do so, highlighting the understanding of the importance of considering net effects. This will be important to consider in an ICF KPI in order to ensure that results are meaningful and not misleading.

3.3.2 Displacement

Displacement (or leakage) refers to cases where stopping an impact in one place simply means it starts or increases in intensity somewhere else. Indirect land use change is a significant displacement effect, whereby land lost to production in one area leads to additional land conversion in another area. This is of particular relevance to NbS, many of which focus on restoration of natural habitats. For example, if tree planting or wetland reconstruction takes place in an area that was previously used for the production of commodities, the action of reconstructing these habitats does not reduce the overall demand for these commodities. Therefore, because of market pressures, a similar amount of commodity will likely end up being grown elsewhere. This may take place in previously undisturbed habitat leading to land conversion and no real difference caused by the NbS at a global scale. As an indirect effect, this will be difficult to measure.

Within the review, many indicators did not take into account potential displacement effects. For instance, one indicator was based on the proportion of total agricultural land under organic management processes. While very likely to improve biodiversity within the area in question itself due to factors such as the reduced use of fertilisers and pesticides, organic farmland generally has lower productivity than conventional farming. Therefore, to meet demand, either other areas will need to intensify their production, or additional land outside of the area of interest will need to be converted to agricultural land, both of which would negatively affect biodiversity and may cancel out the positive effects within the organic area itself (Tuomisto et al. 2012).

On a related note, off-site effects of any intervention, which are often overlooked when assessing impacts, should also be considered. For example, in an NbS project may successfully make use of biocontrol organisms to remove target species within the project area and have no apparent effects on native wildlife. However, if not monitored, an unintended effect may be that the biocontrol agent could disperse outside of the project area and have unknown consequences on native wildlife elsewhere. In many cases, the spatial boundaries of the project will not be the same as the spatial boundaries of the project's impacts.

Ideally, displacement and leakage will be important to capture in order to avoid potential perverse incentives or unintended consequences of indicator use, and to again make certain that the indicator is meaningful when considered in a holistic context rather than in isolation.

3.3.3 Baselines and additionality

A baseline defines the point of comparison against which progress can be measured. This may be an initial assessment at the start of the project, or a historical reference point.

Additionality (or attribution) describes comparisons between project impacts and what would have happened without the project. This may take the form of a business-as-usual projection, or a control site elsewhere, which can again be used to distinguish any impacts as a result of project's actions from any changes that would have happened anyway in absence of the project. From a biodiversity perspective, determining additionality will require a good understanding of the ecosystem and wider socio-political context, because it involves not only predicting a business-as-usual scenario for the project area but also predicting how that scenario would then affect different aspects of biodiversity.

Having a defined baseline and/or a defined way to take into account additionality will be necessary for any ICF KPI to determine change caused by the funding. However, within the review, many indicators did not define a baseline or temporal scale. Those that did most commonly used the project start as the baseline and reported annually throughout the project timeframe. However, other baselines were also used, including the pre-industrial era value, the value at a defined year within the last century, or the predicted values at each point in time if the project had not taken place. In most cases the most appropriate point of comparison will depend on what is being measured. As is already used by a number of ICF KPIs, the project start would make the most sense to use to show change on the scale of interest for reporting the impact of ICF investment. Predicted values if the project had not taken place could also be useful if it is possible to calculate accurately and relatively simply, as is recommended in the developmental ICF KPI 8 Hectares Indicator, which reports on likely avoided loss and degradation of forests. Note that the indicators used by international policy frameworks reflect the combined effect of all policy decisions (e.g. conservation, economic etc.) and so as normally presented, trends in these indicators are not designed to be attributable to specific policies or projects – though this does not necessarily preclude adapting such indicators to understand project level impacts.

3.3.4 Timescales

Ideally, indicators would be sensitive to change on a timescale that matches the project. However, biodiversity changes are typically slow. For short-term projects it is likely that the greatest biodiversity effects would take place after the projects end (for example once the trees had grown to maturity and there had been time for dispersal events to take place from other habitat patches, perhaps decades later). Any ICF KPI would need to consider this problem carefully and perhaps look into extended monitoring post-project if wishing to fully capture relevant effects. Many of the project level indicators for funding generally measured factors that do change more quickly (such as habitat creation) as proxies for likely later species impacts. While this is a pragmatic way to measure project impact it does not consider biodiversity holistically, and would for instance mask an inappropriate or poorly implemented habitat restoration project as a success, as there would be no monitoring of whether the expected species community developed within it. By contrast, indicators used in international policy frameworks are sometimes designed to capture longer-term trends (e.g. progress towards 10-20 year goals) and so are potentially more suitable for describing the lifetime impacts of a project but less sensitive to effects that occur while a project is being actively carried out and monitored.

3.3.5 Spatial scales

Some indicators will apply more effectively at a local or project scale than others. Many indicators that are based on habitat or land cover would be unable to measure the benefits of NbS that happen within that same land cover. Similarly, many of those that rely on models or on global datasets will be based on data at too coarse a resolution for project level implementation. This is especially the case for some NbS such as verge planting and rain gardens, which take place at a fairly small scale. Spatial scale is a particularly important consideration if using indicators from prominent international policy frameworks, which are often designed primarily for national (or global) level reporting.

3.3.6 Prescriptive methods

Prescriptive methods ensure that processes are repeatable and comparable when carried out across different projects. This is likely to be important for an ICF KPI, as in order to aggregate indicators across the portfolio they will need to be applied across projects in a consistent manner. Many of the indicators reviewed were not presented alongside an associated prescriptive methodology, leading to their exclusion during data extraction. While this approach allows more flexibility for projects to produce indicators in a manner that makes most sense within the context of their project, it is not an approach that will work for a context where comparison of the indicator across projects is key. A few indicators gave information about training and workshops designed for users to attend prior to applying indicators, which not only builds capacity but helps to ensure the consistent application of processes and definitions.

Prescriptive methods also need to include precise definitions to ensure terms are applied consistently across project. For example, 'natural habitat' can be a difficult term to interpret if using an indicator such as the area of natural habitat created. Deciding which areas of land (which may be on a whole spectrum of semi-natural habitat types) should count towards this indicator would likely lead to inconsistencies between projects unless the methods made it explicitly clear exactly what 'natural habitat' meant. A solution may be that users are referred to a specific global dataset with a list of appropriate land classes to choose from. This would need to be countered against the value in being project-context specific.

3.3.7 Data quality and quality assurance

Processes will need to be put in place to ensure the quality of any data provided as an ICF biodiversity KPI. Many of the indicators reviewed had such processes in place, including peer review, standardised logic checks, comparisons with other data sources, review of documentation by a third-party authority, adherence to international standards and quality assurance protocols (e.g. FAO's Statistics Quality Assurance Framework), spot checks for selected projects and stakeholder review. Being able to quantify uncertainty will also be important to better understand the confidence you should have when interpreting results. Data quality and assurance processes add costs to the monitoring system, so there may be a trade-off to ensure that some take place, but that they are not excessive and causing unnecessary expense.

3.3.8 Distinguishing project impacts from chance events (also see Baselines and Additionality)

Distinguishing the effects of chance events such as drought or disease outbreaks from the direct impacts that a project has on an area is complex. For example, a three-year Darwin project included aims to improve the status of an Endangered marshland bird. However, although conservation actions were implemented, the project coincided with an extreme drought. As such, it was impossible to use population trends for the species within the timeline of the project to determine if the conservation actions were unsuccessful, or if the lack of increase in breeding success was due to the chance weather event. This influences conclusions on whether the actions would have legacy beyond the project duration or not. Such stochastic events may also lead to incomparability between projects, or even between years within the same project. This issue is not easily addressed by quantitative indicators. In some cases, it may be possible to model the effects of some common factors such as water availability, and in some cases it may be possible to understand inconsistencies in results through the use of expert opinion advising that results in a certain situation should be discarded as incomparable. However, the limitations that this inconsistency brings across all indicator contexts should be recognised and clearly communicated as a limitation of using indicators under any circumstance.

3.4 Notable differences for marine and urban specific considerations

At both the framework and indicator level there was a mix of those indicators that can be applied or try to capture elements of all domains, and those which are specific to certain ecosystems. Even across ecosystems, many of the considerations of suitable biodiversity indicators such as those discussed above still apply, but there is a terrestrial bias in indicators. In the ICF context, there has been interest in considering marine and the terrestrial subset of urban NbS case studies specifically, alongside general (and restricted) terrestrial examples. This section highlights particular considerations in marine and urban environments.

3.4.1 Marine

The review of marine indicators highlighted that those relating to NbS were predominantly relevant to coastal regions, focussing on activities such as coral, seagrass and saltmarsh restoration. Only a small proportion of the reviewed indicators were found to relate to wider marine habitats and species, sometimes with less relevance to NbS, or with limitations restricting their use and application. Although terrestrial and marine indicators can share similar foci (assessment of habitats, species, or pressures etc.), the review highlighted specific differences between terrestrial and marine, which may have attributed to the low representativity of indicators to assess NbS and ICF investment in non-coastal marine areas.

For example, state indicators (also in some cases impact indicators where changes of condition cannot be measured directly) are currently commonly used in marine environments, to assess spatiotemporal changes in components of biodiversity, including habitat condition, or ecological functions, such as changes in trophic chains. However, due to a paucity of marine data, and a limited understanding of state-pressure-impact interactions within ongoing prevailing environmental conditions, state indicators may require the combination of a variety of quantitative, or semi-quantitative methods to enable the successful evaluation of impacts from pressures on sensitive biodiversity receptors (Table 5). Such knowledge gaps can limit the applicability of indicators to assess NbS in coastal and other marine areas. Although, the use of pressure and/or impact methods can help to measure if a feature is improving or failing, particularly in areas where pathways between pressure sources and receptors are not fully understood and/or data is poor and therefore condition cannot be measured directly (e.g. links between reducing, and/or ceasing pollution discharges into the sea and habitat sensitivity).

In addition, the availability of data required to inform certain marine indicator assessments was found to vary dependent on reporting requirements unique to specific geographic locations. For example, limited access to fishing data at appropriate spatial resolution hindered the ability to assess the impacts of fishing pressure on coastal biogenic reefs. Limitations relevant to marine indicators, such as data availability, and the technical capacity of projects to collect, process and measure indicators on the ground were commonly identified with terrestrial environments. However, as a result of complexities associated with marine assessments and monitoring, some limiting factors were considered unique to marine environments. For example, species recovery might be less successful in marine areas if there are issues associated with larvae dispersal or be made more challenging due to marine-specific constraints, such as effects from fluctuations of environmental parameters or human pressures outside of NbS project areas. It should also be noted that marine environments are impacted by both land-based activities (e.g. pollutant discharges) and activities occurring within marine waters (e.g. fishing).

The review highlighted that limitations which may reduce indicator applicability to assessing NbS or ICF investment, such as a lack of baseline or refence condition, could be mitigated by designing marine indicators to measure improvements as a trend from the inception of the project or investment. Additionally, the use of a mixture of datasets could also help to address limitations such as the inaccessibility of marine environments and associated survey costs, which make undertaking detailed assessments of environmental improvements challenging. For example, the use of remote sensing, including aerial lidar, drone photogrammetry and satellites can enable data acquisition in inaccessible areas, such as coastal habitats that are too shallow to access by boat. Remote sensing is also currently used in combination with in-situ sampling methods as a cost-effective approach for data acquisition, enabling the measurement of a wide range of environmental variables, such as turbidity, temperature and the presence and footprint of human activities, which can impact the condition of marine ecosystems. However, its usage and application may be dictated and limited by environmental characteristics such as water depth.

Drawing from the aforementioned marine-specific constraints and potential mitigation measures, the review identified a series of key methodological elements which could enhance the applicability of marine indicators to assessing NbS and ICF investment. Notable marine examples which could be further explored for ICF evaluations are simple habitat extent-based indicators, such as Live Coral Cover (Table 5), which can be implemented at the project-level and aggregated to wider portfolio-scale resolutions. Another example is the proportion of fish stocks within biologically sustainable levels (state indicator for SDG target 14.4), which highlighted that the use of a well-established definition of ‘sustainable’ in the context of maximum sustainable yield made methods easier to implement and data easier to measure.

The types of indicators relating to NbS occurring in coastal environments, including coral, seagrass and saltmarsh restoration could be measured using existing methodologies. Moreover, it was outlined that to maximise indicator applicability, indicators which are already operational in similar contexts to those used in existing ICF investments could also be adapted for future ICF purposes. Marine indicators found to relate to ICF and NbS were not extensive, and the applicability of such methods to countries on the ODA list needs further review.

Table 5 Examples of marine state, impact, response, and pressure indicators. Where used is indicative of major frameworks: CBD - Aichi Biodiversity Target Indicators; SDG - Sustainable Development Goals; RSC – Regional Seas Conventions; Other (specified).

Indicator	Type	Where used	Main strength (+) /weakness (-)
Live coral cover **-** % cover of hard corals of the surface area of a coral community or assemblage	State	CBD; SDG	+ Standardised method + Primarily based on local data collection + BIP indicator - Data not available for all regions - Requires field data collection
Protected area coverage	Response	CBD; SDG; RSC	+ Can be used to monitor additional MPA designations + BIP indicator - Does not indicate if the protected areas are positively impacting biodiversity - No information is captured on how well managed these protected areas are - Designating protected areas is not likely to be relevant to NbS
Marine Trophic Index	Pressure	CBD; RSC	+ Indicates whether fish stocks are being sustainable managed + Time series data is available + BIP indicator - Primarily a global and national indicator - Declines in trophic level can be masked by development of offshore fisheries
Extent of Physical damage to marine habitats	State/Impact	RSC	+ Can be used with simple datasets on the extent of pressure and the coverage of vulnerable habitats -- Testing in ODA locations will be required to fully assess the feasibility of implementation
Catch certified by the Marine Stewardship Council	Impact	CBD	+ Easily implemented at local scale +Includes elements on the biology of targets species alongside impacts on ecosystems and habitats + Availability of defined methodologies - Variable data availability - Not representative of landings within EEZ

In summary, the review outlined that due to complexities of marine environments, the suitability of some indicators as a practical mechanism for reporting on ICF portfolio of investment may be limited, not least because of the issues of being applicable to a variety of projects in different marine ecosystems globally. Additionally, the majority of NbS identified relevant to marine was very limited, only covering a small proportion of marine habitats and species in coastal regions, therefore, overall marine biodiversity considerations may be poorly represented in the ICF portfolio. If assessing performance of marine NbS is a priority, the highlighted challenges should be considered, and the definition of suitable marine indicators should be informed by the types of planned investments and the priority geographical and environmental areas for application.

3.4.2 Urban

NbS taking place in urban environments most commonly focus on a different scale and type of biodiversity compared to more general terrestrial NbS. Examples of urban NbS include urban street tree installation, verge planting and creation of rain gardens. They are therefore often on a much smaller scale than interventions such as habitat restoration, creating microhabitats rather than macrohabitats. This has implications on how they can be measured. For example, any EO based metric or metric based on global scale data will likely not be appropriate to apply in this context, placing a greater need on field measurements to be made by the project team.

The biodiversity effects of such interventions are also likely to be significantly different. For example, while a project planting several hectares of trees could potentially attract any wildlife that relies on forest ecosystems (if well designed to ensure connectivity with other habitat patches to allow for migration), planting street trees does not create a forest environment (including, for example, deadwood, litter, forest soils, shrub and understory species), but rather a set of individual trees. Therefore, it could only attract organisms relying on the tree rather than the holistic forest ecosystem. While street trees may support small organisms that remain in the same tree throughout their life cycle and highly dispersive organisms such as some bird species that are able to travel between individual trees, it will not support the full range of species that a forest could. In this context, the benefits from NbS are heavily swayed to people, and measurable benefits to biodiversity may be extremely limited. If assessing performance of urban NbS is a priority, it may be more suitable to use separate indicators and existing frameworks designed for such projects (e.g. EKLIPSE), rather than attempt to develop indicators that can meaningfully capture impacts from urban NbS projects and larger-scale NbS (e.g. habitat restoration) within the same metric.

3.5 Portfolio level indicators vs project specific indicators

Other ICF KPIs are designed to ensure they can be aggregated across projects for reporting at the portfolio level. They also aim to be relatively simple to communicate and are largely based on a single metric. However, as explored above, biodiversity is a very diverse concept with many different possible metrics to use to measure it. Therefore, any single measure will not capture everything of relevance. In addition, the fact that the majority of biodiversity indicators are context or location specific presents challenges with ensuring that any portfolio level indicator could capture evidence across all project sites. Potential NbS projects capture a wide and varied range of interventions. It will be important to consider whether it is realistic to try and design something that captures, for example, the biodiversity benefits of urban trees *and* the biodiversity benefits of reef restoration within the same indicator.

One solution for capturing this variety would be to select a subset of generic indicators for aggregated use at the ICF portfolio level, but also require projects to design their own more relevant and specific indicators based on their individual project aims. This is an approach taken by many other biodiversity indicator frameworks for funding evaluation, as discussed in more detail in the accompanying framework review. Projects often have very specific aims, and indicators that are able to best meet these aims are likely more effective in those individual cases than any generic indicator could be. Even where indicators are developed at the project level, many funding bodies also require reporting on which Aichi targets and which SDG targets the project is contributing towards, which helps put the project impacts within the context of global objectives.

Another mechanism, which is used by some of the current ICF KPIs, would be to broadly group indicators together so that a range of more specific biodiversity indicators could be applied at a project level, but at the portfolio level an aggregation such as ‘number of projects improving biodiversity' or similar could be possible, with any project fulfilling one of the defined indicators able to count towards the total.

4 Concluding remarks and summary

Biodiversity is a complex concept and projects seeking to contribute to and measure biodiversity are wide-ranging. This review has shown the different ways that measuring it has been tackled by the extensive suite of biodiversity indicators already in use, and highlights the implausibility of being able to report against all aspects of biodiversity, particularly within a discrete set of KPIs. Indeed, in the CBD post-2020 documentation there are 57 indicators identified for the components of proposed Goal A alone. The most appropriate indicators will depend on reporting requirements, in this case to monitor biodiversity impacts of the ICF portfolio of investment across a broad suite of NbS, and will be influenced by the practicality of implementation for both project level reporting and aggregation to report against the entire ICF portfolio of investments. Equally however, practicality for monitoring and aggregation must be balanced against the need to provide meaningful information on biodiversity: neither is helpful without the other, so some compromises are likely for both. Identifying the relative advantages and disadvantages of the different approaches that are currently used to measure biodiversity has provided many lessons that will be useful to consider during the indicator selection and development process, including context specificity, data availability, expertise required for calculation and applicability to NbS. EO is likely to be a significant source of data (in terrestrial and coastal systems) as it gives a general low cost, high value response. In broader marine there will be more reliance on measuring change in pressures. Understanding whether this is a reasonable proxy for species changes will be key. It is likely that indicator development will involve selection of a subset of indicators for use at the portfolio level, but the most meaningful understanding of biodiversity benefits would also require project-level indicators which would be more relevant and context specific.

4.1 Summary of indicators review

4.1.1 Types of biodiversity indicator:

Biodiversity indicators can be grouped by what they measure:

Species state indicators capture biodiversity most holistically but would be expensive to measure and are less easily linked to NbS interventions. Habitat state indicators would be relatively easy to measure, but differ in relevance across contexts, with only indirect links to some aspects of biodiversity.

Pressure indicators may be easy to measure, but their relevance also differs greatly between contexts and links to biodiversity are indirect.

Response indicators prove intention but not effect and have varying degrees of relevance to NbS.

Biodiversity indicators can also be grouped by how they are measured:

Direct indicators may require expertise and expensive in-field data collection but would be simpler if based on habitat rather than species.

Proxy indicators give a likely indication of biodiversity effects and are intended to be more practical than alternatives, but do not provide a direct measurement.

Modelled indicators often make use of global data (reducing the need for expensive data collection) and may take into account multiple aspects of biodiversity to give a more holistic viewpoint but are generally more novel than other indicators discussed, are more opaque and can be harder to communicate.

4.1.2 Important points for an indicator to capture or address:

- Capturing net effects (all positive and negative impacts of an action) increases the likelihood that an indicator is meaningful and not misleading.
- Displacement (cases where stopping and impact in one place simply means it starts somewhere else) and leakage (effects that take place outside of an area of interest) should be addressed to avoid potential perverse incentives or unintended consequences of indicator use, but this can be prohibitively difficult to achieve.
- Baselines should be well defined to allow for comparative analysis and account for additionality (differences between project impacts and what would have happened without the project).
- Timescales of measurement should be appropriate for the expected timescales associated with the NbS's effect on biodiversity. There is often a mismatch between the duration of a project and measurable change in biodiversity, which is likely to impact the appropriateness of an activity indicator (generally easier to measure) vs an outcome indicator (representing actual biodiversity responses).
- Methods should be prescriptive enough to allow for comparable repetition across projects, but will almost certainly require some flexibility to represent the wide range of interventions.
- Methods should have appropriate quality assurance procedures.
- Indicators (particularly for short term projects) may risk being unable to entirely distinguish between project impacts and chance events.

4.1.3 Notable differences between Defra's domains of interest (marine and urban)

- The review of marine indicators highlighted that those relating to NbS were predominantly relevant to coastal regions.
- Although there are scientifically and technically well-defined indicators for the marine environment, there are additional considerations which need to be explored further when considering the assessment of NbS and ICF investment.
- Due to complexities of marine environments, the suitability of some indicators as a practical mechanism for reporting on ICF portfolio of investment may be limited.
- The types of indicators relating to the NbS occurring in coastal environments could be measured using existing methodologies, and the use of operational indicators with similar applications to those used in ICF investments should be considered.
- If assessing performance of marine NbS is a priority, highlighted challenges should be considered, and indicator definition should be informed by planned investments and geographical and environmental application.
- Urban NbS in the main take place at a smaller scale than those more generally, leading to a need for data at an appropriate scale to feed into any indicators measuring biodiversity.

4.1.4 Portfolio level indicators vs project specific indicators:

A single portfolio level indicator will not capture all relevant aspects of biodiversity. It will also not necessarily be relevant to all of the kinds of NbS projects that could be carried out. As such, a subset of portfolio level indicators and a requirement for projects to also implement project specific indicators could be an effective approach to take.

References

BIP. 2011. Biodiversity Indicators Partnership – What makes a successful indicator? [online] https://www.bipindicators.net/national-indicator-development [Accessed 4 December 2020].

CBD. 2006. Convention on Biological Diversity. Convention Text Thursday // 11.2.2006 [online] https://www.cbd.int/convention/articles/?a=cbd-02 [Accessed 4 November 2020].

Hoban, S., Bruford, M., D'Urban Jackson, J., Lopes-Fernandes, M., Heuertz, M., Hohenlohe, P.A., Paz-Vinas, I., Sjögren-Gulve, P., Segelbacher, G., Vernesi, C., Aitken, S., Bertola, L.D., Bloomer, P., Breed, M., Rodríguez-Correa, H., Funk, W.C., Grueber, C.E., Hunter, M.E., Jaffe, R., Liggins, L., Mergeay, J., Moharrek, F., O'Brien, D., Ogden, R., Palma-Silva, C., Pierson, J., Ramakrishnan, U., Simo-Droissart, M., Tani, N., Waits, L. and Laikre, L. 2020. Genetic diversity targets and indicators in the CBD post-2020 Global Biodiversity Framework must be improved. Biological Conservation **248**: 108654 https://doi.org/10.1016/j.biocon.2020.108654

Keith, D.A., Rodríguez, J.P., Brooks, T.M., Burgman, M.A., Barrow, E.G., Bland, L., Comer, P.J., Franklin, J., Link, J., McCarthy, M.A., Miller, R.M., Murray, N.J., Nel, J., Nicholson, E., Oliveira-Miranda, M.A., Regan, T.J., Rodríguez-Clark, K.M., Rouget, M. and Spalding, M.D. 2015. The IUCN Red List of Ecosystems: Motivations, Challenges, and Applications. Conservation Letters, **8**: 214-226. https://doi.org/10.1111/conl.12167

JNCC. 2020. UK Biodiversity Indicators [online] https://jncc.gov.uk/our-work/uk-biodiversity-indicators/ [Accessed 4 December 2020].

Tuomisto, H.L., Hodge, I.D., Riordan, P., Macdonald, D.W. 2012. Does organic farming reduce environmental impacts? – A meta-analysis of European research. *Journal of Environmental Management*, **112**: 309-320, https://doi.org/10.1016/j.jenvman.2012.08.018.

Appendix 1 Summaries of Named Indicators

A selection of the indicators that were reviewed have been summarised here. These reflect specific examples mentioned in the report and represent some of the breadth of currently available indicators. Inclusion here is for informative purposes only, and is not an endorsement of these indicators as ICF KPIs. Indicators are listed in alphabetical order. Major aspects of elements included in the framework are depicted using the key in Box A1. Note the level of expertise is subjectively assessed as low, medium or high input indicator (based on technical expertise and/or amount of data collection required) and does not translate into literal monetary values:

	Marine		Terrestrial		Urban
	Habitat based		Species based		Pressure based
$ $$ $$$	Subjective assessment of the level of expertise required to calculate		CBD indicator		Website link

Box A1. Key design or capture elements of indicators

Abundance of indicator species

$$$

Variations of this indicator are commonly applied at the project level. This direct state indicator involves the selection of a few select species of interest that are thought to reflect likely population responses in other biodiversity within the same ecosystem (known as indicator species). The populations of these species are then monitored through field studies throughout the course of a project. The species chosen would vary per project (making it difficult to aggregate meaningfully at the portfolio level), and generally require field surveys for data collection.

Biodiversity Habitat Index

$$

This modelled state index predicts the effects of habitat loss, degradation and fragmentation on terrestrial biodiversity, by taking into account the spatial distribution, condition and rarity of ecologically similar habitat. It can optionally be translated into an estimate of the proportion of native species expected to persist over the longer term within the area of interest, which can be used as a modelled indicator. It is calculated globally on a 1 km^2 grid. Habitats are considered ecologically similar based on generalised dissimilarity modelling that takes into account both species composition, turnover and distribution records, and abiotic factors such as climate, terrain and soil type. Data on the loss, degradation and fragmentation of habitats are generated from global remotely sensed datasets. The index can be calculated for an area of any scale, using a weighted geometric mean of the scores obtained for all grid cells within the area of interest. It covers a wide taxonomic range of plants, invertebrates, amphibians, reptiles, mammals and birds. Whilst originally developed for forests, it is now applicable across all terrestrial ecosystems. The index is reported on a scale of 0-10, with 10 representing 100% of natural habitat retained and 0 representing none. The data is available on request and updated annually. The indicator was designed for measuring and reporting progress in relation to the Convention on Biological Diversity's Aichi Target 5, and can be aggregated from project to portfolio level.

Biodiversity Metric 2.0

$$

This modelled state indicator is an aggregated biodiversity metric aiming to determine the change in biodiversity following a development or land use change intervention. As most NbS constitute land change interventions in one form or another, it would be widely applicable in this context. It takes into account changes in habitat type, area, condition, connectivity and strategic significance. The current method is based on the UK Habitat Classification system, so it would need adapting to a global habitat classification system if applying it in ODA countries in the context of ICF projects. It would require some local data collection, but most of the information could be derived from land cover maps or other global sources. The indicator was designed as a Defra and Natural England tool. It could be used as a project indicator that could be aggregated to a portfolio level indicator.

Catch Certified by the Marine Stewardship Council

$$

This indicator measures the catch of fisheries certified by the Marine Stewardship Council, as a percentage of total wild catch. Catch figures are based on data from the Food and Agriculture Organization of the United Nations data portal. The indicator relates to Aichi Target 6 and SDG 14, and can be aggregated from project to portfolio level.

Ecotoxicology factor

$$$

This pressure indicator measures the effects of pollutants on populations within an area of interest. Values must not exceed the level of pollutant at which more than 50% of the population of species of interest is affected (the EC50). It also measures the time that it takes for 50% of the pollutant to disappear from the area of interest (DC 50). Pollutants are one of many pressures on biodiversity.

Extent of natural habitats

$

Many frameworks included habitat extent related metrics as direct state indicators. The habitats specified were most commonly forest, but varied widely including water related ecosystems, mountain green cover, saltmarsh extent, seagrass extent and coral reef extent. Others simply specified 'natural ecosystems' as a generic point, sometimes asking for this to be broken down by 'type' (but leaving it up to the user how to do this). Often, the extent of restored habitat was asked for instead of the total habitat area. Many extent metrics could be determined through EO based metrics and/or records that projects would keep anyway, so should be relatively easy to calculate. However, issues may arise around consistently applicable definitions of what counts as 'natural habitat,' or of each more specific habitat type recorded. Indicators measuring the extent of natural habitats make most sense at a national or global level than a project or portfolio level, as the project starting conditions would affect the outcome. However, indicators measuring the extent of restored habitat could be used at a project level and aggregated to the portfolio level.

Extent of physical damage to marine habitats

$$

This indicator calculates the extent and intensity of potential damage to a given seafloor habitat. The indicator uses and combines two types of information to produce a map of habitat disturbance: the distribution and sensitivity of habitats; and the distribution and intensity of human activities and pressures that cause physical damage. The indicator is used to assess progress against targets for the OPSAR convention and the UK Marine Strategy and can be aggregated from project to portfolio level.

Global Ecosystem Restoration Index

This modelled composite state index is based on three datasets addressing different aspects of land restoration: change in land productivity (calculated as the ratio between net primary productivity and precipitation), change in energy balance (evapotranspiration, which a functional ecosystem should optimise) and identity transitions in land cover (to act as the structural aspect of restoration and to put the other two aspects into context). It is calculated globally on a 1 km^2 grid, so could be used at a project scale and aggregated to portfolio level. Data is due to be publicly available once a web interface has been developed by the German Centre for Integrative Biodiversity Research (iDiv) computational infrastructure but is not available at the time of writing. It would be applicable to any restoration based NbS project. It may be more of an indication of ecosystem health than biodiversity, but poor ecosystem health would constitute a significant pressure on biodiversity. The indicator was developed to assess against the Convention on Biological Diversity's Aichi Target 15.

IUCN STAR

The Species Threat Abatement Restoration metric (STAR) uses data from the IUCN Red List of Threatened Species to estimate the potential reduction in species extinction risk that could be or has been achieved, for example, across a corporate footprint, within a country or on a project site. It can also be used to set local or global species extinction risk targets, and measure progress towards those targets. It is a modelled state indicator based on using pressures as inputs. It is calculated globally on a 5 km^2 grid, so could be used at a project scale (where projects are larger than this) and aggregated to portfolio level. For each pixel within this grid, data on species conservation status (IUCN Red List category), area of habitat (both current and historical, calculated using species distribution polygons, habitat associations from the IUCN Red List, and land cover maps) and the threats they face (from the IUCN Red List threat classification hierarchy) are used to estimate the relative contribution of each threat affecting species present to their extinction risk. It is designed to be aggregated, for example across a portfolio or programme. The fact that this calculates both potential and achieved effects for a site mean that it could be used both at the project selection stage and in project target setting, as well as in project monitoring and evaluation. It is currently based on mammals, amphibians and birds due to good data availability, although they are looking to expand it to wider taxonomic groups in the future. It would be widely applicable across NbS projects, responding to any project leading to changes in extent of suitable habitat, or changes in any of the Red List threats. Whilst the final version of the methods is not yet complete, there are worked examples of the current version of the method, so it would be feasible to calculate now if allowances were made for further improvements to the method in future. There may be issues around update frequency of the Red List data if applying to a project scale, although the method does expect some local data collection as well which should improve this situation.

Live coral cover

This global indicator measures live hard coral cover, the primary indicator of the health of coral reefs. It uses data collected in the field and from remote sensing to calculate the percentage of hard corals of the surface area of the coral community. The Global Coral Reef Monitoring Network collates the data and the indicator also contributes to Essential Ocean and Biodiversity Variables. The indicator is aggregated from site level data so could be disaggregated to the project level if data is available. However, projects may require data collection to set baselines and for future monitoring.

Local Biodiversity Intactness Index

This modelled state index provides an estimate of the amount of a site's original species richness that persists following land use change caused by humans, and pressures related to such changes. It complementary in approach to the Biodiversity Habitat Index (described above), with a focus on average affects at a local and site-specific scale, instead of on overall diversity of a larger region. It is calculated globally on a 1 km^2 grid, so could be used at a project scale and aggregated to portfolio level. Data from 2005 are published alongside hindcast and projected data in a time series from 1500-2095. It makes use of PREDICTS as a global database of local biodiversity surveys recording responses to land use change and other pressures, as well as high resolution global land-use data. The PREDICTS database currently includes data on 45,000 species of plant, vertebrate and invertebrate species. In the context of applying this index to ICF projects, recalculations would need to be made to incorporate changes caused by the project. It would be applicable to any NbS causing land use change, for example restoration projects.

Marine Trophic Index

$$

This index measures the mean trophic level of fisheries catches for all large marine ecosystems and can indicates whether fish stocks, especially of large bodied fish, are being overexploited and fisheries are being sustainably managed. Data is available from Sea Around Us at the University of British Columbia. The indicator relates to the Aichi Target 6, the SDGs and the Regional Sea Conventions core indicator set. This indicator can be disaggregated from the global level to the sub national if data is available.

Native species richness

$$$

A number of project specific indicators reviewed contained counts of direct state metrics of native species richness. Measuring this would require field surveys and significant local expertise to identify native species present, and so would likely be costly. There is a high probability that some species would be missed even with an experienced surveyor. Some projects specified a particular taxonomic group rather than identifying all species and some were interested in taxa richness as opposed to species richness. It would be difficult to meaningfully aggregate to a portfolio level as different ecosystems will naturally support a different species richness and therefore different projects would not be comparable.

Number of incidences of illegal activities

$$

Some project specific indicators took into account the number of illegal events recorded within the project area, as a response indicator. Examples included illegal hunting and illegal burning. This indicator would only be of relevance to projects targeting these pressures, which may be difficult to do using NbS. It could be possible to aggregate results to a portfolio level.

Number of invasive alien species

$$$

Invasive species are a significant pressure biodiversity, particularly in specific ecosystems such as islands. Counts of the number of invasive species, or similarly the number of eradication events have therefore been used in some projects and frameworks. Monitoring the number of invasive species could be considered either a state indicator or a pressure indicator, while monitoring their removal would constitute a response indicator. This would be of most relevance to projects taking place in geographically discrete areas such as islands where it is clear which species are native and which have been introduced and are causing harm. Other indicators instead measured the number of introduction or eradication events taking place. This could be used as a project level indicator. Whilst possible to aggregate as a portfolio level indicator, there may be issues with comparability.

Number of plans, policies, or institutional structures into which biodiversity has been integrated

$

Some frameworks included a generic response indicator identifying the number of plans, policies or institutional structures that have been implemented as a result of a particular project or initiative, which could be aggregated to a portfolio level. For example, this was included in the German Government's IKI (International Climate Initiative) framework for ODA spend on biodiversity and climate. The indicator does not give an indication of whether the plans adopted were successful in conserving biodiversity.

Proportion of land degraded over total land area

$$$

This is a direct state indicator from the SDGs that measures degraded land as a proportion of total land. It is a combined metric from three sub indicators, covering land cover / land cover changes, productivity and carbon stock. It recognises its own limitations in that land degradation is context-specific so it is difficult for a single indicator to capture, and that as it is based on a binary degradation value (degraded or not degraded) it is unable to capture differing degrees of degradation. Despite these limitations, a similar indicator will need to complement any extent-based indicators selected to ensure the quality of any restoration based projects as well as just the area restored. It could be measured at a project level and aggregated to a portfolio level.

Protected area coverage

$$

This indicator measures the policy response to biodiversity loss for both terrestrial and marine systems. Data is available for individual protected areas at Protected Planet, so could therefore be used at the project level to track changes in protected area designations, and the portfolio level if data is aggregated. It is unlikely to be relevant to NbS.

Rio Marker on Biodiversity

$

This is an internationally recognised response metric measuring how much ODA has been spent on biodiversity per year, based on criteria for what constitutes spend on biodiversity. It could be used at the project level and aggregated to the portfolio level.

Shannon and Simpson's diversity indices

$$$

These are two direct state indicators that are commonly used in ecology to assess the species diversity of a community of organisms. Both indices take into account the number of species present, their relative abundances and their evenness, with the evenness having more effect in the Shannon index and the abundance of dominant species having more effect in Simpson's index. Their use in ICF projects would require field measurements and ecological expertise of relevance to the project area. They could be used at the project level, but would be difficult to meaningfully aggregate to a portfolio level as different ecosystems will naturally support a different species richness and therefore different projects would not be comparable.

Species Habitat Indices

$

These modelled state indices provide data on the change in suitable habitat of single species, which can be used to infer their potential population loss and extinction risk. Indices on all species occurring within an area of interest can be aggregated to give a total biodiversity value for that area. Global habitat suitability maps are created on a 1 km² grid based on information from literature and experts about species' habitat restrictions and remotely sensed land cover products. Data are available for over 20,000 species of terrestrial plant, vertebrate and invertebrate, although at the time of writing the dashboard is currently in a test phase and improvements in the ease of navigation will be needed. This indicator could be applied to any NbS project affecting change in habitat extent (e.g. restoration projects) but would not pick up any changes in projects that improved habitat quality or reduced any other pressures on biodiversity. It could be used at the project level and aggregated to the programme level.

Wildlife Picture Index

This direct state index is calculated using presence/absence data from camera traps and sound sensors to measure changes in species variation over time at a landscape scale. It is sensitive to changes in the number of species, their relative occurrence and evenness over time. The index currently takes into account 278 species of tropical mammal and bird species, but could be adapted to include any species within a site of interest that it is possible to record with a camera trap or sound sensor (likely larger vertebrates). It would require local data collection. It could be used at the project level but would be difficult to meaningfully aggregate to a portfolio level as different ecosystems will naturally support different species and species compositions and therefore different projects would not be comparable.