ARPHA Preprints

PENSOFT

Project Report

Author-formatted document posted on 22/03/2023

Published in a RIO article collection by decision of the collection editors.

DOI: https://doi.org/10.3897/arphapreprints.e103765

D3.3 Identification of current monitoring workflows and bottlenecks

Alejandra Morán-Ordóñez, Pedro Beja, Sara Fraixedas, Sergi Herrando, Jessi Junker, W. Daniel Kissling, Maria Lumbierres, Anne Lyche Solheim, Gabriel Miret, Jannicke Moe, Francisco Moreira, Henrique Pereira, Joana Santana, Dani Villero, Lluís Brotons

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

EUROPABON

D3.3 Identification of current monitoring workflows and bottlenecks

31/01/2023

Lead beneficiary: The Centre for Ecological Research and Forestry Applications (CREAF)

This project receives funding from the European Commission's Horizon 2020 research and innovation programme, under Grant Agreement n.101003553

Prepared under contract from the European Commission

Grant agreement No. 101003553
EU Horizon 2020 Coordination and Support Action

Project acronym:	**EuropaBON**
Project full title:	**EUROPA BIODIVERSITY OBSERVATION NETWORK: INTEGRATING DATA STREAMS TO SUPPORT POLICY**
Start of the project:	01.12.2020
Duration:	36 months
Project coordinator:	Prof. Henrique Pereira Martin-Luther Universitaet Halle-Wittenberg (MLU) www.europabon.org
Type:	Coordination and Support Action
Call:	The Sc5-33-2020 Call: "Monitoring ecosystems through innovation and technology"

Project ref. no.	**101003553**
Project title	**EUROPA BIODIVERSITY OBSERVATION NETWORK: INTEGRATING DATA STREAMS TO SUPPORT POLICY**

Deliverable title	**D3.3 Identification of current monitoring workflows and bottlenecks**
Deliverable number	**D3.3**
Contractual date of delivery	**31.01.2023**
Actual date of delivery	**31.01.2023**
Type of deliverable	**Report**
Dissemination level	**Public**
Work package number	**WP3**
Institution leading work package	**Centre for Ecological Research and Forestry Applications (CREAF)**
Task number	**T3.3**
Institution leading task	**The Centre for Ecological Research and Forestry Applications (CREAF)**
Author(s)	**Morán-Ordóñez, A., Beja, P., Fraixedas, S., Herrando, S., Junker, J., Kissling, W.D., Lumbierres, M., Lyche Solheim, A., Miret, G., Moe, J., Moreira, F., Pereira, H.M., Santana, J., Villero, D., Brotons, L.**
EC project officer	**Laura Palomo-Rios**
Deliverable description	This deliverable identifies the main bottlenecks in current biodiversity data flows to the generation of EBVs at the European level. Results are detailed by EBV (36 EBVs from those identified at the EuropaBON EBV list – Deliverable 4.1) and summarized across EBV classes. The deliverable introduces a framework for analyses of bottlenecks on biodiversity monitoring workflows based on 16 different criteria related to data collection and sampling, modelling, interoperability and IT infrastructure and data integration.
Keywords	**Biodiversity Monitoring, Data workflows, Bottlenecks**

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EUROPABON

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

Contents

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EUROPABON

Executive Summary

The EuropaBON project seeks to design a European Biodiversity Observation Network to monitor the status and trends of European biodiversity and ecosystems in the coming years. To accomplish this, the project has brought together biodiversity monitoring experts and other relevant stakeholders from various sectors (policy, NGO, academia, business, citizen science) from all over Europe. These have agreed on a list of 70 Essential Biodiversity Variables (Deliverable 4.1) that should enable tracking the progress of biodiversity-oriented policy instruments and provide information on changes in biodiversity at different levels.

In this deliverable, we analyse the limitations (bottlenecks) in the current European biodiversity data flows and their ability to achieve the generation of those EBVs. We define a bottleneck as a factor that hinders or constrains the quantification/generation of a given EBV at the desired temporal frequency and spatial resolution/extent. We propose a framework to analyse those bottlenecks based on 16 different criteria related to a diversity of aspects related to data collection and sampling, modelling, interoperability and IT infrastructure and data integration. We assess whether each EBV on the current EuropaBON EBV list (Deliverable 4.1) could be generated from already existing biodiversity monitoring data flows coordinated at the European level. These existing flows were previously mapped and collected in the EuropeaBON biodiversity monitoring database (EuropaBON task 3.3); however, we also consulted other relevant sources of information such as the recent MarBioME assessment on *the current state of Biodiversity Monitoring in the European Union and adjacent marine waters* (Jessop et al. 2022) or the *Biodiversa + Literature survey on protocols and methods used across regions and countries* (in preparation at the same time of this EuropaBON deliverable; Silva del Pozo and Body, 2022). We compared each of the data flows in these monitoring programmes with the 16 defined bottlenecks criteria and assessed whether the criteria were fully or partially met, or not met at all. Therefore, for each combination of EBV and monitoring programme, we were able to determine whether for a given criterion there were no bottlenecks to report or whether these were partial or total, respectively.

We were able to identify bottlenecks in data flows for 36 different EBVs, mostly belonging to the EBV classes “species populations” (61% of EBVs) and “community composition” (25%) across the freshwater, marine and terrestrial realm. These data flows correspond to those of 18 different European coordinated monitoring programmes.

The **main bottlenecks** to the generation of EBVs from current biodiversity monitoring data flows - across all EBV classes and realms - **relate to data integration.** The **lack of long-term secured funding** was identified as a partial or full bottleneck for the generation of EBVs in 74% of the monitoring programs evaluated. Secured funding is not just needed to allow the maintenance of a given monitoring programme in the long term, but also to increase its geographic coverage (e.g., the total number of sampling sites) and its capacity building and training, to support volunteers or to hire specialists (e.g. modelling technicians, taxonomists, IT professionals) or to create and maintain IT infrastructure to promote the automatization of data flows. In fact, **automatization of data streams is partially achieved** in many of the monitoring programmes evaluated (54%) thanks to the design and implementation of webportals, Apps other integration and harmonization tools, but absent in 29% of them. **Models to generate spatially-explicit EBVs from monitoring data are not routinely incorporated in the data flows** and processing of most the monitoring programs

This project receives funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003553.

assessed (88%). When the models needed to generate the target EBV have been tested or are routinely used within the monitoring programme, these are mostly fit in the statistical tool R that, although it is open-software, it uses programming language and requires advanced technical knowledge and in many cases, the code used for modelling is not widely open. This is an important barrier to the generation of EBVs at the European scale considering that most (70%) integration programmes in the EuropaBON biodiversity monitoring database (Deliverable 3.1) are coordinated by non-research institutions and in these institutions, volunteers typically do a lot of the data integration work and most of them do not have a scientific or technical training. The data needed to generate the EBVs was partially available in 50 % of the monitoring programs evaluated (and not available at all in 19% of the monitoring programs evaluated).

Overall, only 52% of the biodiversity metrics and indicators currently generated by existing European monitoring programmes partially matched those described in the EBVs definitions: they differed in their specifications in terms of taxonomic resolution (i.e. not all taxa targeted in a given EBV are covered by a given monitoring program) and spatio-temporal resolution (i.e. products already in place have generally a broader spatial resolution than described in the EBV definition or generated at lower temporal frequency); in **38 % of the monitoring programs assessed, the products generated do not match the EBV definitions at all**.

We identified that **multiple bottlenecks** in monitoring data hinder specially the production of EBVs **for some taxonomic groups** such as for example, zooplanckton in freshwater systems, lichens, terrestrial artropods or crops pests for which we could not find any European-wide monitoring programme that could provide harmonized data to generate the corresponding EBVs at the desired spatio-temporal scale. Moreover, we found that monitoring data in European marine waters is very fragmented and mostly integrated (if at all) at the regional level (e.g., the NorthEast Atlantic region - OSPAR convention, the Baltic Sea - HELCOM convention) with biodiversity from Southern and Eastern European waters being scarcely represented in databases. This fragmentation hampers the capacity to generate marine EBV metrics across all European waters due to the lack of data harmonization among different monitoring schemes and regions.

This Deliverable lacks data representation for the genetic, ecosystem function and ecosystem structure EBV classes, for which we could not retrieve enough information about gaps and bottlenecks in data flows (when none at all). This is either because the data at the European level does not exist (case of genetic composition EBVs) or because the data needed to generate these EBVs is produced by remote sensing programs which data flows were hard to evaluate with the proposed bottlenecks framework, better suited to assess data flows of standard biodiversity monitoring programs.

Overall, our results suggest that improving current biodiversity monitoring flows to generate European level EBVs will require overcoming a plethora of interaction constraints that come in bundles. In this context, the hereby proposed framework can be used to understand the factors currently hindering the generation of EBVs at the desired the spatial, temporal and taxonomic resolution as well as those that have already being tackled by long-term monitoring networks to overcome bottlenecks in their data flows and integration over time (e.g. better standardization of monitoring protocols, definition of metadata standards or by developing webportal tools and apps to facilitate data harmonization and integration tasks). Despite this framework has been applied to European level monitoring networks, it could perfectly be transferred to other scales, for example

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

to assess bottlenecks in biodiversity data flows at National level or within specific contexts (e.g. integration of biodiversity monitoring data across protected areas). In fact, our framework should be viewed as a tool to facilitate monitoring workflow assessments across different taxonomic and spatial scale assessments, and to allow the coherent reporting on monitoring and EBVs generation across different contexts. Within the EuropaBON project, the results from this task represent the basis to identify actions needed to overcome current bottlenecks in biodiversity monitoring data flows to improve the current European biodiversity monitoring system so that it will be more representative temporally, spatially and taxonomically (EuropaBON task 4.3 co-design), it will maximize benefits (task 4.4 cost-efficiency) and become better integrated into wider biodiversity policy.

The team

Alejandra Morán Ordóñez and Lluís Brotons from CREAF led the analyses, with support of other members of the CREAF team (Dani Villero, Sergi Herrando, Sara Fraixedas, Gabri Miret). The leading team worked closely with: Joana Santana (ICETA), Pedro Beja (ICETA), Francisco Moreira (ICETA) on the assessment of gaps (partially incorporated into this report); Maria Lumbierres (UvA) and Daniel Kissling (UvA) on the classification of bottlenecks criteria into broad categories that link to the EuropaBON EBV co-design workflows task (EuropaBON task 4.3); Henrique Pereira, Jessi Junker (MLU; coordinating team) on the overall framework introduced in this deliverable and Anne Lyche Solheim and Jannicke Moe (NIVA) on the evaluation of freshwater EBVs.

The leading team also received valuable data and input on the overall bottlenecks analysis framework and the EBV factsheets from Adrià López-Baucells (Natural Science Museum of Granollers), Ana Cristina Cardoso (JRC), Anne Lyche Solheim (NIVA , Borja Jiménez Alfaro (University of Oviedo), Chiara Magliozzi (JRC), Diego Pavón Jordán (NINA), Eugenio Gervasini (JRC), Gabriel Gargallo (Institut Català d'Ornitologia - ICO), Joana Soares (TERINOV), Judy Shamoun-Baranes (University of Amsterdam), Nicolas Segebarth (European Commission), Laurence Carvalho (NIVA), Sophie Mentzel (NIVA), Simon Potts (UReading), Tom Breeze (UReading), Tom Langendoen (Wetlands International), Vujadin Kovacevic (European Commission).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

1. Rationale

The EuropaBON project seeks to design a European Biodiversity Observation Network to track the status and trends of European biodiversity and ecosystems in the coming years. The project is deeplyrooted in the concept of Essential Biodiversity Variables (EBVs) (Pereira et al. 2013; GeoBON 2021): a theoretical framework to harmonize and standardize biodiversity data from diverse sources into a minimum set of key variables able to inform about biodiversity change at different levels (six major classes: genetic composition, species traits, species populations, community composition, ecosystem functioning and ecosystem structure); these key variables facilitate data integration into a spatially and temporally-explicit way and represent an intermediate abstraction level between primary observations (both from field and remote sensing data) and indicators (Geijzendorffer et al., 2016; Kissling et al. 2018). EuropaBON, along with a large community of stakeholders from various sectors (policy, NGO, academia, business, citizen science) across Europe, has worked on defining and specifying a list of priority EBVs to be measured across the continent (Deliverable 4.1; Moersberger et al. 2022) that could potentially allow tracking the progress of biodiversity-oriented policy instruments (e.g., the European Biodiversity Strategy for 2030, the EU Strategy on Green Infrastructure, the New Restoration Law, the Habitats and Birds Directives, the Water Framework Directive). EuropaBON will co-design a cost-efficient framework for the future monitoring of these EBVs across Europe, including aspects related to data collection, -processing and -management, relevant workflows, and a coherent framework for the analysis and reporting of data and biodiversity trends in key ecosystem types.

This co-design process will build on the vast, standardized, transnational monitoring programmes and biodiversity data flows already existing in Europe. These have been already captured and summarized in a web-based database specifically designed for this purpose within the framework of the EuropaBON project https://monitoring.europabon.org/ (Deliverable 3.1). The database focused on monitoring efforts that are effectively coordinated at a supranational level and the European scale, and it retrieved information for each of these monitoring networks (also called “integration initiatives”) at three different levels (Figure 1): 1) data collection: what is the biodiversity/remote sensing data collected in sampling schemes in each integration initiative? Who coordinates data collection efforts and how?; 2) data integration: what institutions/projects/platforms integrate and process biodiversity data to generate EBVs, Ecosystem Services Variables (ESSVs) or any other indicators? These are called “integration nodes”, which, if they are multiple within a single integration initiative, are structured hierarchically; and 3) data flows and final products: how does information flow through different institutions and programs? (e.g., frequency of flow, the spatial resolution of data, product/indicators generated).

The main objective of this deliverable is to provide a common framework to identify and characterize major limitations (bottlenecks) in our ability to derive functional EBVs at the European scale from available biodiversity monitoring data. Past efforts aiming at describing the main factors hampering monitoring development have mostly focussed on specific constraints in isolation (e.g., taxonomic, geographic and data accessibility gaps; Wetzel et al. 2018 - EUMON database), rarely acknowledging the multifactorial nature of successful biodiversity monitoring workflows (Silva del Pozo and Body 2022). Therefore, there is a need for an integrative perspective

This project receives funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023.* DOI: https://doi.org/10.3897/arphapreprints.e103765

that adequately assesses the variety of factors at play in biodiversity monitoring and evaluates it across a wide range of monitoring flows that may be very different in nature (Kühl et al. 2020).

This deliverable builds on already existing integration initiatives to identify the main bottlenecks to generating EBVs under current monitoring and integration efforts at the European level. In the EuropaBON context, we consider a bottleneck as any aspect or factor that hinders or constrains the quantification/generation of a given EBV at the desired temporal frequency and spatial resolution/extent. These bottlenecks are assessed across different levels of data flows within the already existing European monitoring initiatives, from data collection to data integration processing and storage to data sharing.

By providing a common framework to assess and map bottlenecks in European biodiversity monitoring, we will be able to assess which monitoring efforts can deliver information to develop robust European-level EBVs at the desired spatial and temporal resolution units (as defined in the EuropaBON EBV list - Deliverable 4.1). Our approach also allows us to guide future improvements of both current monitoring networks and integration flows. This will be possible because our comprehensive framework will serve as a foundation for experts in various fields, EBV classes and taxonomic groups to improve data flows in biodiversity monitoring. In this deliverable, we provide an application of the framework to existing biodiversity monitoring workflows reported in the EuropaBON database, but we intend to apply it to the different showcase workflows for selected EBVs in the various EU policy directives evaluated in WP5. Finally, our approach will facilitate the identification of bottleneck commonalities across EBV groups that will be integrated in the future co-design efforts developed in WP4.

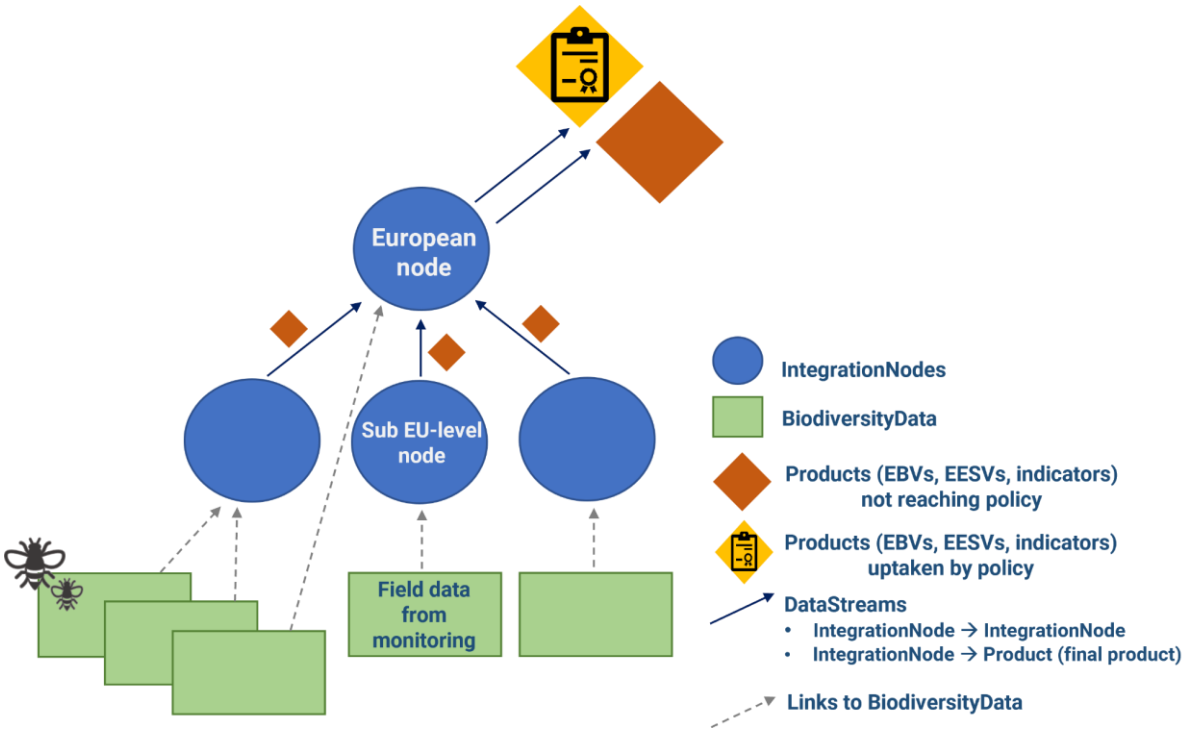


Figure 1. Schematic diagram of the elements and data flows (also call data streams) within a given biodiversity monitoring integration initiative as considered in the EuropaBON biodiversity monitoring database.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

2. Methodology: a framework for bottleneck analyses of biodiversity data flows

Our framework centres around identifying and describing types of bottlenecks that impede the development of the EuropaBON project's Essential Biodiversity Variables (EBVs) from monitoring data at the European scale. It uses this information to assess the potential for generating EBVs from existing monitoring data within the EuropaBON project.

For each EBV in the EBV list (deliverable 4.1), we searched in the EuropaBON biodiversity monitoring database (deliverable 3.1) for European integration initiatives that could potentially provide data to generate specific EBVs. For example, for the EBV "species distributions of freshwater fishes", we searched the database for monitoring initiatives collecting fish data. In addition to the information contained in the database, other available sources of information were consulted; for example, in the case of marine EBV, the assessment of t*he current state of Biodiversity Monitoring in the European Union and adjacent marine waters* and its corresponding database was used as a reference (Jessop et al. 2022).

Then, we evaluated the EBV as described in the EBV list - in terms of taxonomic coverage, spatial and temporal extent (EBV attributes; Kissling et al. 2018) -, against the already existing biodiversity data flows coordinated at the European level and evaluated 16 different criteria reflecting potential bottlenecks grouped in three categories selected to allow future integration in the co-design process to be developed in WP 4 (task 4.3). These categories were: 1) data collection and sampling, 2) modelling, interoperability, IT infrastructure and 3) data integration (Figure 2; Table 1):

1. *Bottlenecks related to data collection.* These refer to aspects/issues linked to the collection of biodiversity data itself (e.g., related to fieldwork, remote sensing approaches). These bottlenecks indicate whether the monitoring network (integration initiative) depends on elements that facilitate data collection (e.g., volunteer training) and whether the data gathered fulfil quality criteria allowing its integration (Figure 1). The criteria evaluated in this bottleneck category were:

1.1 **Standard monitoring (StMn)**: systematic monitoring programmes rely on standardized sampling protocols to collect data that facilitate data harmonization and integration (in contrast to the unstructured data that generally populates data repositories). Moreover, most systematic monitoring programmes are carried out over broad geographic areas and over long-time spans allowing to estimate trends of change and eventually understand the drivers behind those changes (extensive schemes, *sensu* Proença et al. 2017); data from long-term monitoring programmes can be especially valuable to fit robust models and EBV predictions over space and time when compared with other biodiversity data (e.g., opportunistic records) (Honrado et al. 2016; see the "Models" criteria below); therefore, the lack of standardized monitoring programs is understood as a bottleneck to data integration and therefore, EBV estimation;

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

1.2. **Capacity building (CBui)**: this assesses whether a given biodiversity monitoring network has demonstrated the ability to strengthen, create, adapt and maintain biodiversity monitoring capacity over time (e.g., by training volunteers, adding new sites to the monitoring programme, etc.); the lack of demonstrated capacity building and training by the monitoring network under evaluation was considered a bottleneck;

1.3. **Data type (DaTy)**: this evaluates the type of data collected by the monitoring programme (e.g., presence-only, presence-absence, abundance estimates, genetic diversity) and whether it matches the EBV needs (for example, community abundance EBVs can´t be estimated from presence-only data); if the data collected by a given monitoring programme does not serve to generate the EBV, this is considered a bottleneck; for example, if the monitoring program only collects species occurrence data (records of presence/absence in a given place, it would not serve to generate EBVs on species abundances, since the latter needs count data;

1.4. **Taxonomic completeness (TxC)**: EBVs are defined for specific biological entities or groups/clusters of biological entities (i.e., a list of species or ecosystem types; for example, the EBV “Distributions of marine turtle species nesting grounds” focuses on the turtle species listed in the Habitats Directive); this bottleneck criterion assesses whether all taxa considered in the EBV definition get covered by the monitoring programme or if the taxonomic coverage is incomplete;

1.5. **Geographic completeness (GeC)**: we assess the European coverage of the biodiversity monitoring programme (number of countries) and consider this to be complete/full when all EU countries + UK collect data for that specific program; if some areas of Europe are not covered by the monitoring programme, we consider this to be a partial bottleneck;

1.6. **Timely Update (TiUp)**: because EBVs are meant to track the progress of different European environmental policies, this criterion assesses whether the sampling frequency of data collection matches (or is superior to) the temporal resolution of the EBV, designed in many cases to match policy reporting obligations.

Note here that the last three criteria (TxC, GeC and TiUp) refer to three main attributes of EBVs (space, time and taxonomy; Kissling et al. 2018) and can also be used to assess gaps in biodiversity data availability; as such, these three criteria have been evaluated in further detail in EuropaBON task 3.2 “Report on gaps and important new areas for monitoring in Europe and guidelines/ recommendations for filling these gaps” but are included here again for their integration into this more general framework of the bottlenecks analysis.

This project receives funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003553.

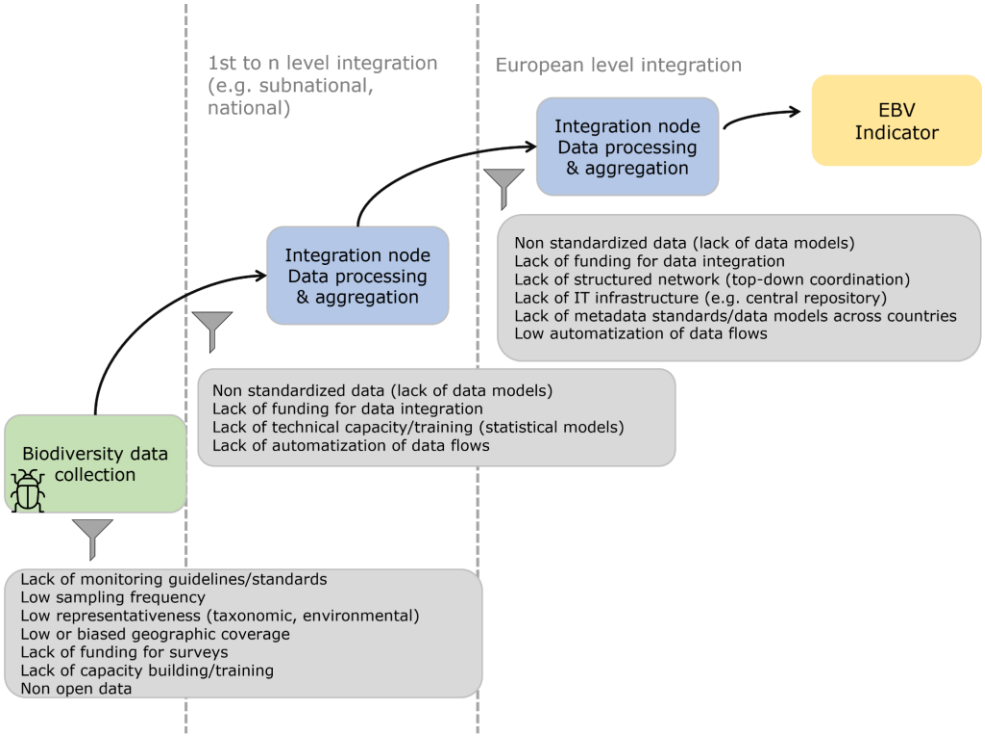


Figure 2. Potential bottlenecks to the generation of EBVs at the European level. Bottlenecks can occur from data collection to data processing (multiple integration levels) and storage to data sharing.

2. Bottlenecks related to modelling, interoperability and IT infrastructure:

2.1 **Models (Mod)**: the EBVs, as defined in EuropaBON, are spatially-explicit (i.e., they are defined over a continuous space - Europe in this case - with a given spatial resolution). However, biodiversity data are collected in situ (locally) using a variety of sampling units (point samples, linear transects, grid cells, etc.). Statistical models are therefore required to extrapolate biodiversity metrics from sampling units across a continuous space, as well as to upscale and downscale biodiversity data resolution. For example, the use of community-level species distribution models has allowed researchers to produce habitat suitability maps of the "potential area of occupancy" of habitats of Conservation Community Interest by the European Habitat Directive 92/43/ECC. These have a spatial resolution of 1-km and span the whole of Europe (Schaminée et al. 2014, 2016a, 2016 b; Jiménez-Alfaro et al. 2018). They are based on a combination of a) data on vegetation records derived from systematic monitoring programs (the European Vegetation Archive; Chytrý et al. 2016) and b) environmental predictors available at the continental scale (e.g. climatic and land use predictors). In this report, we evaluate whether a given integration initiative already uses statistical models to extrapolate biodiversity metrics and estimate EBVs or other indicators across Europe and whether these models are suited for the generation of the target EBV. If an integration initiative does not use models to develop spatially and/ or temporally seamless datasets, this hampers the generation of EBVs at the European scale and is thus regarded as a bottleneck. Note here that the use of models is not needed in all cases to generate EBVs: methods such as spatial aggregation or interpolation (e.g., kriging) might in some cases suffice to generate EBVs at the desired spatio-temporal scale;

2.2 **Metadata Standards** (**MdSt**): the use of metadata standards or data models for data integration favours the traceability of data flows and their replication. It also facilitates the

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

automatization of the data integration process by setting clear definitions and reference terms that are comparable across biodiversity data sets (e.g., concerning spatial- and temporal coverage and -resolution, measurement units, sampling strategy, and taxonomy). The lack of metadata standards is therefore understood in this framework as a bottleneck to the generation of EBVs at the continental scale;

2.3 **IT infrastructure (CRep):** the EBVs generated from biodiversity monitoring data should ideally be available in the long-term - e.g. to allow subsequent time series analyses - and therefore there is a need for a permanent or semi-permanent storage infrastructure to be in place (Kissling et al. 2018). The lack of a central repository is therefore understood in this framework as a bottleneck to data flows that generate EBV products.

3. Bottlenecks related to the data integration process:

3.1 **European Integration node (EuInt)**: because the focus of this analysis is the integration of EBVs at the European scale, the lack of an integration node for the target taxa at the European level is seen as a bottleneck for the generation of EBVs. By integration node, we refer here to an institution/ structure responsible for a given monitoring flow and its associated product/ EBV;

3.2 **EBV match (EBVm):** this criterion provides an overall assessment of the match between the EBV definition (in terms of taxonomic focus, temporal and spatial resolution) and the integration initiative under evaluation. If the integration initiative already generates EBV products at the desired taxonomic and spatio-temporal scale, then the EBV match is 100%;

3.3 **Software (Sofw)**: this criterion evaluates whether a given integration initiative uses a software solution for modelling EBVs that is easy to access and operate by non-scientists (note here this criterion only applies if biodiversity models are routinely in place within the framework of the monitoring program - see **Mod** criterion above). This is especially important considering that most (70%) integration initiatives in the EuropaBON biodiversity monitoring database (Deliverable 3.1) are coordinated by non-research institutions and in these institutions, volunteers typically do a lot of the integration work (e.g., data aggregation at the national level of the Pan- European Common Bird Monitoring Scheme - species abundances and trends - are estimated by national coordinators). The use of complex modelling approaches that require specific technical or statistical skills is therefore understood as a bottleneck to EBV generation in this framework;

3.4 **Open and reproducible code (OpC)**: this criterion also relates to the use of statistical models to extrapolate EBVs across space (Mod). Currently, there is growing interest in using free-software solutions for statistical computing (e.g., R, Phyton) that allow users more flexibility in the types of input data and formats, as well as in the types of models that can accommodate these. Users must know the programming code to implement these software solutions. However, it is more and more becoming standard practice to make code available in public repositories, thereby facilitating the use of these types of software and the reproducibility of the analyses (e.g., GitHub or Gitlab). For our framework, we considered integration initiatives that use complex modelling approaches without making the programming code publicly available, a bottleneck in data flows to EBV generation;

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

3.5 **Automated data streams (Auto)**: the automatization of data flows (from biodiversity collection to data integration at different levels) is key to speeding up EBV generation and minimising errors linked to data handling and processing. The fast development of new Apps is already assisting this automatization in biodiversity monitoring. For example, the European Bird Portal reports daily observation of bird species across Europe (aggregated distribution maps at 30 x 30 km resolution) thanks to the automatic integration of species lists that get reported via Apps to data portals such as the national nodes of ornitho, liveatlas.nl or the eBird. The lack of automatization is therefore interpreted here as a bottleneck in data flows to EBV generation;

3.6. **Funding (Fnd)**: financial security is pivotal to promoting the generation and maintenance of EBV products in the long-term. Funding supports coordination and integration tasks, maintains IT infrastructure, and allows the development of tools such as Apps, etc. The lack of financial security in the long term of a given integration initiative is understood as a bottleneck to EBV generation;

3.7 **Open Data (OpDat)**: data availability is an important requirement for EBV generation. In many cases, raw data and biodiversity data sets might not be 'findable, accessible, interoperable and reusable' (FAIR principles; Wilkinson et al., 2016), and this restricts the data flows to generate EBV products. This criterion assesses whether raw data is available or not and under which conditions.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Table 1. Bottlenecks evaluated for each EBV. The description reflects the questions we investigated in relation to already existing biodiversity data flows at the European level.

Acronym & bottleneck		Description
Data collection and sampling		
StMon	Standard Monitoring	Does the monitoring network integrate data from standardized monitoring schemes?
Cbui	Capacity Building	Does the monitoring network have capacity building and provide training?
DaTy	Data type	Does the data type collected by the monitoring network fit the EBV requirements? (e.g., presence, presence- absence, abundance data)
TxC	Taxonomic completeness	Do data at European level exist for all taxa/habitats/ecosystems targeted in the EBV?
GeC	Geographic Coverage	Does the monitoring network collect data across all EU27 + UK?
TiUp	Timely updated	Does the timing of the data workflows within a given monitoring network match the EBV desired temporal resolution?
Models, Interoperability & IT infrastructure		
Mod	Statistical models	Are statistical models routinedly used for data integration and processing?
MtSd	Metadata standards	Are there any metadata forms available for a standardized recording and integration of data and metadata (e.g., sampling process, locality info)?
CRep	Central Repository	Is there a central repository for data storage and processing that follows standardized data management protocols?
Data Integration		
EUInt	European integration	Is the monitoring network coordinated at the European level?
EBVm	EBV match	Do the products already generated/integrated by the monitoring network match the EBV specifications?
Fnd	Funding	Does the European-level node have funding secured for keeping this integration initiative running in the mid- long- term?
Auto	Automated data streams	Is there software/app already in place that facilitates data flows at various levels over the data workflows (e.g., from data collection to data integration)?
SoFw	User-friendly software	Is the software used for modelling user friendly? (i.e., a software solution that even non-technical people find accessible and easy to operate - easy to update, visually appealing interface, intuitive operation...)
OpC	Code	Is the software code for data integration and modelling shared? This relates to modelling/data integration: is the code used for modelling open to everyone (reproducible code, version control, R packages...)?
OpDat	Open data	Is the raw data collected by the monitoring network open/accessible?

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

We assessed whether each of the 16 criteria were met, partially met, or not met at all by the current available data flows of European level monitoring programmes/initiatives. The latter two indicate the existence of bottlenecks in the dataflows that might hamper the generation of the EBV at the European level. This was done by summarizing the retrieved information on these bottlenecks for each monitoring programme into EBV fact sheets in a narrative fashion: for each EBV, we generated a fact sheet where we described in detail what the monitoring programme does, whether it fulfils the criteria under consideration, what the current bottlenecks in data flows are and - when information was available - why the latter exist. This information is then interpreted to assign a score for each bottleneck criteria by answering the questions in Table 1. If the criteria were met, we assumed the lack of bottlenecks. For some criteria this was a straightforward interpretation; for example, the GeC criteria was only completely fulfilled when all European countries were covered by a given monitoring programme. If one or more European countries were not covered by a given monitoring programme, this was scored as “partial bottleneck”. The absence of a monitoring program for a particular taxon across all countries will be considered a full bottleneck for the GeC criterion. For other criteria that are harder to quantify with numbers (e.g., capacity building), the existence of bottlenecks (or the lack of) was interpreted in relation to the information that could be gathered from the monitoring programme and from expert consultation (this is detailed in each of the EBV fact sheets). For some of these more difficult criteria (i.e., funding), other EuropaBON tasks include some more in-depth discussion derived from additional data collected via formularies and personal interviews with the main coordinators of these European monitoring programmes (task 3.4 - cost-effectiveness of current monitoring programs).

To ease interpretation of EBV bottlenecks, we visually represented the 16 bottlenecks criteria as petals in a flower plot (Figure 3). When a criterion is met or partially met by a given monitoring initiative, the petal is fully coloured or half-coloured, respectively. A petal showing grey indicates the presence of a full bottleneck. This kind of visualisation could be adapted in the future to a more quantitative representation in case that bottlenecks can be measured or divided in a larger number of categories.

Finally, we summarized the retrieved data on bottlenecks across the six broad EBV classes (genetic composition, species traits, species populations, community composition, ecosystem functioning and ecosystem structure) and across realms (terrestrial, freshwater and marine) to identify common challenges in EBV generation.

This project receives funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003553.

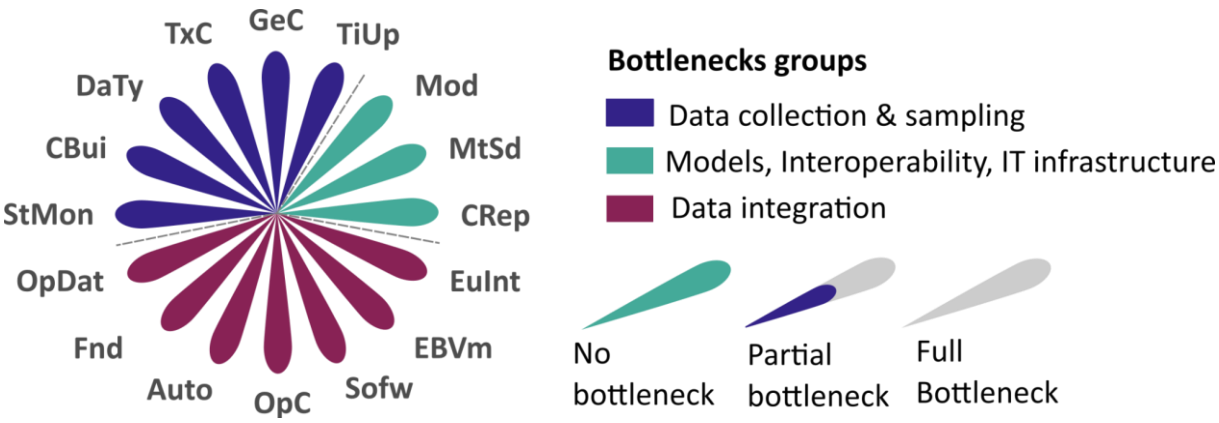


Figure 3. Visual representation of the bottlenecks criteria evaluated for each EBV. Each criterion is represented as a petal in the flower. A fully coloured petal indicates a criterion is fulfilled by an integration initiative for a particular EBV (i.e., no bottlenecks were identified for that criteria). Petals coloured in grey indicate the presence of bottlenecks. Acronyms of the 16 criteria are detailed in Table 1.

3. Framework of bottleneck analyses in practice

While the main goal of this deliverable is to assess current bottlenecks for European level biodiversity data flows, we will first illustrate the process of applying the proposed bottleneck framework, and demonstrate how it could be used to understand how already existing European monitoring programmes (integration initiatives) have overcome bottlenecks in their data flows over time. For this, we have selected two key and contrasting examples of monitoring initiatives coordinated at the European level which are also relevant for two of the showcases in WP 5. These two monitoring initiatives have undergone significant improvements towards a European EBV in recent years and therefore provide an opportunity to test the capability of our framework to visualise these using a common perspective. For the rest of the initiatives in this deliverable, only one flower figure will be derived from the currently available monitoring flows. The two initiatives used as illustration cases were: 1) the Atlas of European Breeding Birds, coordinated by an association of naturalists (the European Bird Census Council – EBCC), and 2) the proposal of the EU pollinator monitoring scheme (EuPoMS), promoted by the European Commission. These two integrative initiatives could potentially contribute data to the generation of two EBVs, “Species distributions of terrestrial birds” and “Species distributions of terrestrial priority invertebrates and key pollinators”, respectively.

This project receives funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023.* DOI: https://doi.org/10.3897/arphapreprints.e103765

3.1 EBV "Species distributions of terrestrial birds" and the European Atlases of Breeding Birds

The specifications of this EBV are the following (definition of the EBV in EuropaBON deliverable 4.1):

EBV: Species Distributions Terrestrial Birds	
Realm	Terrestrial
EBV class	Species Populations
EBV name	Species distributions of terrestrial birds
Definition	The presence/absence or probability of occurrence of each European terrestrial bird species within contiguous spatial units (grid cells) across the EU over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	1 × 1 km – 10 × 10 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	All terrestrial birds of the EU (taxonomy based on the HBW and BirdLife Taxonomic Checklist, with focus on those bird species that are officially recognized in the List of birds of the European Union).

The Second European Breeding Bird Atlas (EBBA2) (Keller et al. 2020) is the most recent and comprehensive European-wide integration initiative mapping the species distributions of all European terrestrial birds and as such, the product generated by this initiative matches quite well the EBV under consideration (distribution maps for terrestrial birds at 10 x 10 km resolution - although only for 222 breeding species). Data presented in this atlas corresponded to a sampling period between 2013 – 2017. The atlas was coordinated by the EBCC. This same association published in 1997 the first Atlas of European Breeding Birds (EBBA1; data was collected mostly in the 1980s (Hagemeijer and Blair 1997). Having reference and comparable data of species distributions in these two time periods has allowed the EBCC to estimate spatially-explicit trends of change in European birds' distributions over the last 30 years (Keller et al. 2020).

When considering the EBV details in terms of desirable spatial (1 x 1 km – 10 x 10 km) and temporal resolution (3 – 6 years), we can use the analytic framework proposed in this deliverable to identify the bottlenecks in data workflows to the generation of this EBV in each of the two EBCC atlases and assess whether and how these have also changed over time.

1) Data collection and sampling. The EBBA2 has mapped the distribution and abundance of 596 breeding birds (both native and non-native) across Europe at the 50 x 50 km grid resolution. For this purpose, an extensive collection of data from multiple sources was conducted at national level and then integrated at European level. Abundance and breeding likelihood were also assessed at 50 x 50 km for all species. Moreover, species distribution maps at 10 x 10 km resolution were developed for 222 species using statistical modelling. EBBA2 targeted surveys (breeding period; 10 km2 squares) followed a standardized protocol (time surveys 60 – 120 min, carried out between 2013 – 2017) which derived from different ongoing monitoring programs depending on each country (data from national atlases, national breeding surveys, etc.).
The EBBA1 data mapped the distribution of breeding bird species at 50 x 50 km being the data reported at insufficient resolution to map or model species distribution at finer scales: in fact, not all

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

50 x 50 km squares were sampled following standardized monitoring protocols but general guidelines and, in some cases, already existing data (e.g., National atlases) were the only data source available for a specific grid cell (these qualify as partial bottlenecks in **DaTy** and **StMon**: Figure 4a). The geographic coverage of EBBA1 was more limited than that of EBBA2 (data gaps in Eastern European Countries) but in both cases, all current EU countries were well covered. Thanks to external funding (by the MAVA foundation and others, as there was no EU funding in support of this initiative), the EBBA2 project contributed to capacity building in different countries; however, this capacity building was more limited during the first Atlas which counted on a smaller budget (**CBui** partial bottleneck in EBBA1; Figure 4a). In any case, it will be difficult to replicate the sampling intensity and effort dedicated to EBBA2 with a frequency lower than 15-20 years for all species (**TyUp** bottleneck); however, and because of the joint work between EuropaBON and EBCC, species distributions for a set of 50 European bird species (in particular farmland species) will be generated to test the capacities and limitations to update maps and analyse change with a frequency of 5 years (EBBA Live Farmland initiative; this initiative will be showcased in EuropaBON Deliverable 5.1).

2) Models, Interoperability and IT infrastructure. The EBBA2 used statistical models to generate 10 x 10 km species distribution maps in EBBA2 (probability of occurrence); on the contrary, maps in EBBA1 reflect a simple aggregation of species presence/abundance estimates at the 50 x 50 km grid cell (**Mod** bottleneck in EBBA1; Figure 4a). Both EBBA1 and EBBA2 designed metadata standards for data reporting and aggregation, but the EBBA2 improved the capacity of the EBBA1 project of centralizing and harmonizing data at the European level thanks to the automatization of dataflows (see next section on data integration) (**CRep** partial bottleneck in EBBA1).

3) Data integration. The 10 x 10 km grid species distributions models in EBBA2 were fitted using R (a priori, a non-user-friendly software, and the code used to fit the models is not openly available; **OpC and Sofw** bottlenecks; Figure 4b). These models are complex (especially the site occupancy models) and require of specific data from the observation process not always available because of the type of surveys carried out; there is therefore, and additional bottleneck related to the need of advanced technical programming skills and model knowledge to replicate the species distributions built on EBBA2. Both the EBBA1 and EBBA2 were supported by external funding; in the EBBA2 the funding was directed to training, capacity building and bird sampling in countries not previously covered by any survey protocol, but it also supported the overall coordination of the EBBA2 project. In the EBBA1, the funding was especially directed to handling and data processing. Both in EBBA1 and EBBA2, the funding support ended with the publication of the atlas: this hinders the continuity of sampling programs in some countries/regions as well as of data integration tasks (**Fnd** bottleneck: Figure 4a, b). While EBBA2 has made a huge effort to automate data streams compared to EBBA1 (e.g., data aggregation at both National and European level has been automated using MapViewers), the automatization of data streams from sampling plots to European coordinators was lacking (**Auto** partial bottleneck in EBBA2 and full bottleneck in EBBA1). All the data generated in EBBA2 is available either open access (50 x 50 observed occurrence data) or upon request (10 x 10 km modelled probability of occurrence data, 50 x 50 km breeding evidence data, 50 x 50 km abundance data and 50 x 50 km EBBA1 - EBBA2 change data). In the latter case, requests should be approved by an EBCC committee and data handling fees can be charged EBBA2 50 x 50 occurrence data are available under license CC BY 4.0, the rest of EBBA2 maps are © EBCC. Following EBCC data policy, raw 10 x 10 km

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023.* DOI: https://doi.org/10.3897/arphapreprints.e103765

sampling data is not owned by the EBCC but by its national partners, which could be asked for these data (**OpDat** partial bottleneck).

So, while EBBA1 and EBBA2 have already generated continuous predictions of the distributions of all birds across Europe, EBBA products only partially match the definition of this EBV (with divergences in taxonomic completeness (only 222 species were modelled at 10 km resolution), temporal and spatial resolution demanded) (**EBVm** partial bottleneck). However, the extra funding, the experience gained in coordination tasks and the technical advances in 30 years have allowed to overcome some bottlenecks related to modelling, standardization, automatization of data streams, among others in EBBA2 compared to EBBA1.

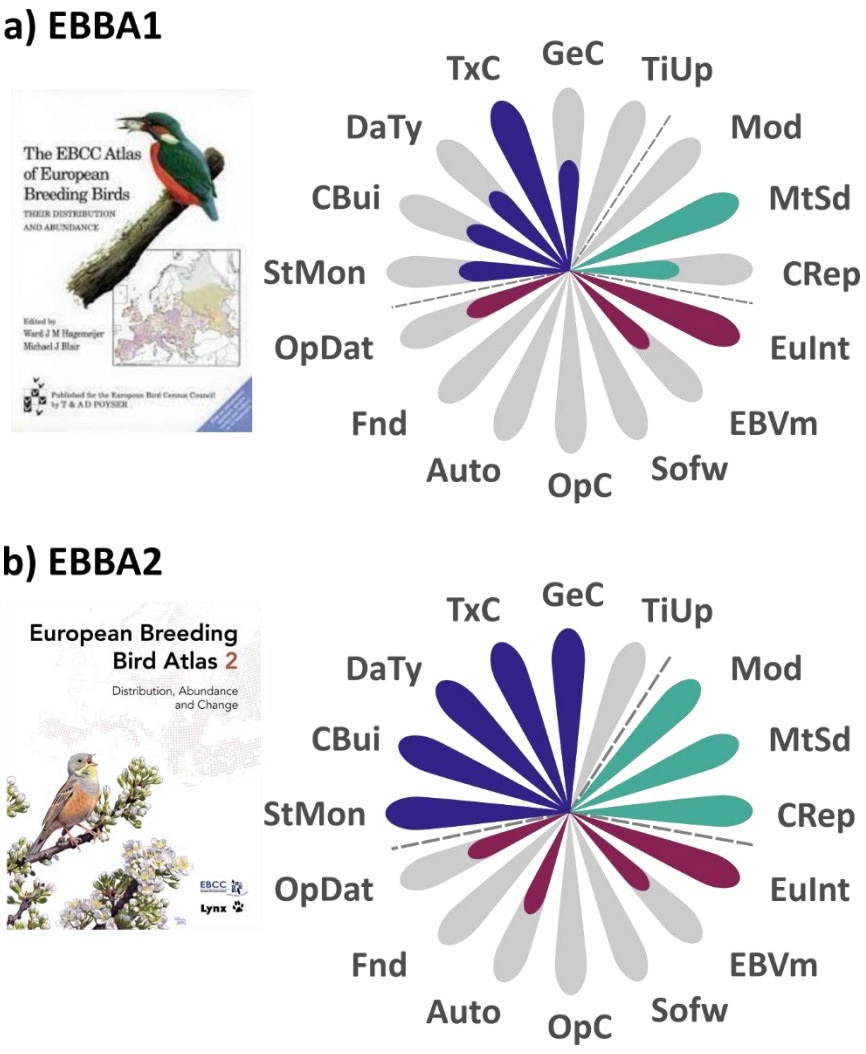


Figure 4. Bottlenecks to the generation of the EBV "Species distributions of all terrestrial birds" by the first and second atlases of European Breeding birds (EBBA 1 and EBBA2, respectively).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

3.2 EBV "Species distributions of key pollinators" and the EU PoMS

The specifications of this EBV are the following (definition of the EBV in EuropaBON deliverable 4.1):

EBV: Species distribution of terrestrial priority invertebrates and key pollinators	
ID	50
Realm	Terrestrial
EBV class	Species populations
EBV name	Species distribution of terrestrial priority invertebrates and key pollinators
Definition	The presence/absence or probability of occurrence of priority invertebrates and key pollinator species within contiguous spatial units (grid cells) across the EU over time.
Metric	Binary presence/absence Probability of occurrence
Spatial resolution unit	10 × 10 km – 50 × 50 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	- Priority invertebrates as listed in the Annex II and Annex IV of the Habitats Directive - Key pollinator species as specified by the EU Pollinator Monitoring Scheme (EUPoMS)

While the description of the EBV also includes priority invertebrates from those listed in the Habitats Directive, in this example we focus on the identification of bottlenecks to the generation of distributions of key pollinators (bees, hoverflies, and moths).

The 2021 report on the design of an European Pollinator Monitoring scheme (EU PoMS) (Potts et al. 2021), when implemented in practice, will collect the data to generate this EBV since the abundance estimates, along with the location/spatial coordinates of systematic surveys will be recorded and made available for this purpose. The EU PoMS design has already identified current bottlenecks that could hinder the estimation of pollinator abundances across Europe which also apply to species distribution models. These relate to:

1) Data collection & sampling. The EU PoMs proposal has identified more than 76 pollinator monitoring schemes already collecting pollinators' data across Europe; however, the variety of sampling methods used makes it difficult to combine the already existing data in these programs to generate pollinator indicators or to estimate abundance trends at the European level. Moreover, for some important pollinator taxonomic groups, the data is limited (**TxC** partial bottleneck); this relates to the limited capacity building of some countries where there is lack of taxonomic resources and experts (**Cbui** and **GeC** partial bottlenecks). The EU PoMs proposal describes in detail how systematic surveys will be carried out for each taxonomic group (surveys are initially planned on an annual basis and assessments every three years). There are ongoing Preparatory Actions already working on overcoming these bottlenecks: the SPRING project seeks to strengthen taxonomic and citizen science capacity with regard to pollinating insects, the ORBIT and Taxo-FLY , seek to create a more centralized taxonomic EU facility for the identification of wild bees and to develop resources for European hoverfly inventory and taxonomy, respectively and an expert working group is developing a rare species module for integration with the main monitoring scheme.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

2) Models, interoperability, and IT infrastructure. Some tests have already been run on using models to develop species distributions from data collected across this monitoring network, including integrated distribution models accounting for occupancy and process-based models for predicting abundance and visitation (**Mod** partial bottleneck). The EU PoMs pilot proposal envisions data to be centralized in a repository at the EEA, European Commission (DG ENV), JRC or Eurostat. Data will be submitted or shared in a standardized form (following metadata standards) via an online platform to the European coordination facility, where pan-European analyses will be made. Currently, none of the latter two elements exist (**MtSd** and **CRep** bottlenecks), though they are specifically being co-developed between an expert working group, DG ENV, EEA and Member States with delivery due in late 2023.

3) Data integration. The EU PoMs proposal presents options to set up a Pan-European pollinator monitoring network, which does not exist so far (**EuInt** bottleneck). The proposal cites the European Butterfly Monitoring (eBMS) as the closest initiative to what the EU PoMs wants to set up at the European level. Initially, EU PoMs expects to be integrating data at the national level to estimate trends in pollinators abundances and taxonomic diversity and therefore there is not a perfect match between the integration product generated by EU PoMS and this EBV (**EBVm** partial bottleneck): however, the abundance estimates collected along with the location/spatial coordinates of systematic surveys could be made available for making spatially explicit predictions of pollinators distributions. There are not yet clear guidelines about whether the code used for data integration and potentially modelling will be openly shared (although tests have been run with R, which is not a user-friendly software; **OpC** and **Sofw** bottlenecks). The EU PoMS will put in place metadata standards to facilitate data integration and there is a plan to make the data openly available upon request, following the eBMS model (**OpData** partial bottleneck). One of the actions proposed to automate data flows is the development of a pan-European internet identification platform for pollinators, which is constantly maintained and updated. The monitoring of pollinator populations by EU PoMS will prove key to tracking the goals set in the proposed Nature Restoration Law of halting the decline of pollinator populations by 2030 and achieve thereafter an increasing trend of pollinator populations. This regulation (Nature Restoration law) will impose an obligation to monitor pollinator communities and therefore member states will have to ensure that sufficient funding is allocated for this purpose (therefore the change in **Fnd** bottleneck between Figures 5a and 5b).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023.* DOI: https://doi.org/10.3897/arphapreprints.e103765

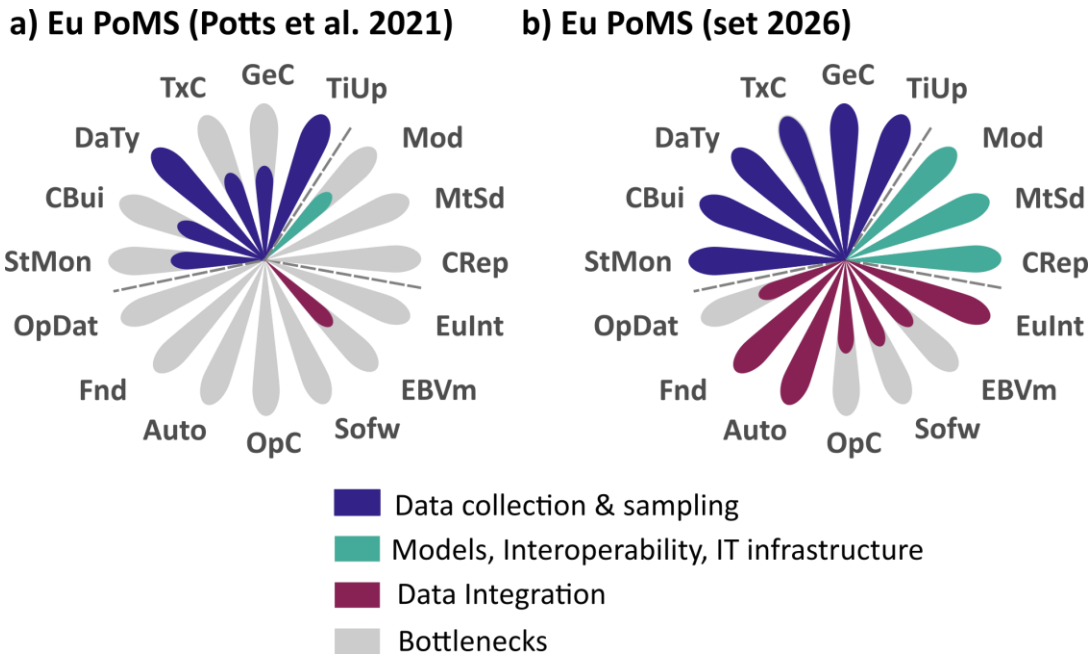


Figure 5. Bottlenecks to the generation of the EBV "Species distributions of key pollinators" comparing **a)** the information that is currently collected by European monitoring networks (as evaluated in the 2021 proposal for an EU pollinator monitoring scheme; Potts et al. 2021) and **b)** the EU-wide monitoring program envisioned by EU PoMS (Potts et al. 2021).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

4. Current bottlenecks to the generation of European-wide EBVs

We were able to reliably retrieve detailed information about bottlenecks in data flows for a total of 36 EBVs among those defined in the EuropaBON EBV list (deliverable 4.1): 14 Freshwater EBVs (Annex I), 4 Marine EBVs (Annex II) and 18 Terrestrial EBVs (Annex III). The European coordinated monitoring programmes evaluated here (18 in total) collect data that could mainly contribute to the generation of EBVs within the class "species populations" (61%) and "community composition" (25%), being the other EBV classes scarcely represented. For example, we could not find any monitoring initiative coordinated at the European scale that collects data to generate EBVs related to genetic composition in any of the three realms and only 5 monitoring programs collecting data that could contribute to the generation of the EBV classes species traits, ecosystem's structure or function.

Table 2. Number of EBVs from EuropaBON EBV list (deliverable 4.1) for which detailed information about bottlenecks in biodiversity monitoring data flows was retrieved (i.e., EBVs for which all criteria in Table 1 were evaluated). For some EBVs, not all bottlenecks' criteria could be assessed but still it was possible to collect some information about monitoring programmes that could potentially inform the EBV. These are indicated in this table in grey font.

	Realm		
Ebv class	**Freshwater (AnnexI)**	**Marine (Annex II)**	**Terrestial (Annex III)**
Genetic composition			
Species populations	7	4 (2)	11
Species traits	1		2
Community composition	5		4
Ecosystem structure		3	1
Ecosystem function	1		
Total	14	4	18

Main bottlenecks to EBV generation at the European scale

The **main bottlenecks to the generation of EBVs** from current biodiversity monitoring data flows - across all EBV classes and realms - **relate to data integration aspects.** There are already many biodiversity monitoring networks in place collecting and integrating data across taxa and countries to generate different products at the European level. Some examples include, the multi-species trends indicators generated by the European Butterfly Monitoring Scheme or by the Pan-European Common Bird Monitoring Scheme , the maps of species distributions for EUNIS habitats developed with data from the European Vegetation Archive or the Ecological Quality Ratios generated for different freshwater taxonomic groups within the data flows of the Water Framework Directive (WISE-2 Biology data). However, the metrics and indicators generated in 52% of the monitoring programs assessed showed a partial mismatch with those described in the EBVs definitions: their specifications differ in terms of taxonomic resolution (i.e. not all taxa targeted in a given EBV are covered by a given monitoring program) and spatial resolution (i.e. products already in place have generally a broader spatial resolution than described in the EBV definition) or temporal resolution (**EBVm** criterion, partial bottlenecks; Figure 6); 38 % of the monitoring programs assessed, the products generated do not match the EBV definitions at all (EBVm criterion, full bottlenecks). For example, the eBMS

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

generates multi-species trends indicators at the national and European scale, and while the project collects raw data that could potentially be used to generate spatially-explicit species distributions of key pollinators (count data at the transect level), the eBMS does not generate such product (EBV ID 50; FIgure 6).

Automatization of data streams is partially achieved in many of the monitoring programmes evaluated (54%) (**Auto** partial bottlenecks; Figure 6): these programmes already count on web platforms or APPs where national coordinators or data contributors can upload the data, and where data then gets integrated after some automatic and manual quality check and harmonization procedures. For example, the second European Breeding Bird Atlas developed Map Viewing tools that automatically generated maps from the count data reported by national coordinators and that will flag inconsistencies in reported abundances along border areas between neighbour countries. However, within this same integration initiative there was little automatization in terms of data collection or in data flows at the national level. A similar example is the reporting tool of the WISE-2 Biology data (ReportNet) where Member States can upload the data flowing to the reporting of the Water Framework Directive. A few monitoring initiatives are already in the process of automatizing the flows of data from data collection to data integration. Some examples are the EuroBirdPortal that automatically integrates and harmonizes data collected in citizen science portals (many of them already operated via smartphone applications) (EBVs 45b and 58a in Annex III) or the "IAS Europe'' smartphone App developed by EASIN to promote the reporting of sightings of Invasive Alien Species of Union Concern by citizens and their integration into the EASIN GeoDatabase (see for example EBV ID 7 in Annex I) (only 16% of the monitoring programs evaluated could count on fully automated data flows). On the other hand, **around 29 % of the EBVs could be generated with data from monitoring programs that have not automatized their data flows** yet (e.g. the EURING programme collects the data needed to generate the *Phenology of migration of terrestrial birds EBV* - date of arrival and departure for different migratory species across Europe - EBV 58a Annex III; however ringing centres across Europe send their ringing observations (data exports) once a year to the EURING central repository manually).

The lack of long-term secured funding was identified as a partial or full bottleneck for the generation of EBVs in 74% of the monitoring programs evaluated. Many of these monitoring networks are coordinated and run by NGOs and research institutions or a mix of both (83% of European integration nodes in the EuropaBON monitoring database); many of the developments made by these institutions (e.g., the EBBA2 or EMMA2 atlases, the EBP map viewer or the interactive visualization tools developed by GLOBAM to follow bird migration detected by weather radars) have been funded via competitive calls (e.g., LIFE and BiodiVERSA projects) or with sporadic and short-term contributions from private foundations (2 -3 years). Funding budgetary constraints have been already identified by stakeholders consulted by EuropaBON as one of the main challenges to carrying out long-term monitoring programmes (Moersberger et al., 2022). Secured funding is not just needed to allow the maintenance of a given monitoring initiative in the long term, but also to increase its geographic coverage (e.g. number of sampling sites) and its capacity building and training, to support volunteers or to hire specialists (e.g. modelling technicians, taxonomists, IT professionals) or to create and maintain IT infrastructure to promote the automatization of data flows. Therefore, **to secure funding is key to minimize other bottlenecks in data flows** across all the three categories considered in this assessment (data collection and sampling, models,

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

interoperability and IT infrastructure and data integration; see also task 3.4 of EuropaBON). In our assessment, only those EBVs that could be generated from data flows linked to the mandatory reporting of quality status (EQR) by Member States to the Water Framework Directive could be considered to have funding secured in the long - term.

The use of **models to generate spatially-explicit EBVs from monitoring data is not routinely incorporated in the data flows and processing of most the monitoring programs assessed (88%; Figure 6**). In some of these programmes statistical models are in place to generate indicators but these models are not suited to generate the EBV under consideration, and therefore this was considered a full bottleneck. On such approaches, the TRIM model (TRends and Indices for Monitoring data) is used to estimate species trends in the PEBCMS, the IWC or the eBMS; however, this model is not suited to generate predictions of species distributions. A few monitoring programmes have already made some modelling tests to generate the EBV under consideration; for example, habitat suitability models have been already fit for the freshwater invasive species *Elodea nuttalli* from EASIN data across EU (Steen et al. 2019); however, this is not a product regularly generated by this initiative, but a case study. Examples like the latter (tests, cases studies) were considered partial bottlenecks in modelling capacity (**Mod**) and were found for 24% of the monitoring programs. The lack of modelling capacity does not only hamper the generation of spatially-explicit EBVs (predictions over continuous grid cells in Europe), but also the capacity to predict EBV values to unsampled areas or of making inference about how different taxa/indicators might respond to changes in environmental conditions (e.g. to predict how a given EBV value will change after a restoration intervention).

In the few examples in which models have been tested or are routinely used within the framework of a monitoring program, these are mostly fit in the statistical tool *R* that, although it is open-software, it uses programming language and requires advanced technical knowledge; moreover, and in many cases, the code used for modelling is not widely open (**OpC** bottleneck). This is an important barrier to the generation of EBVs (**Sotfw** bottleneck in 95% of the EBVs evaluated; Figure 6) at the European level considering that most (70%) integration initiatives in the EuropaBON biodiversity monitoring database (Deliverable 3.1) are coordinated by non-research institutions and in these institutions, volunteers typically do a lot of the data integration work and most of them do not have a scientific or technical training.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

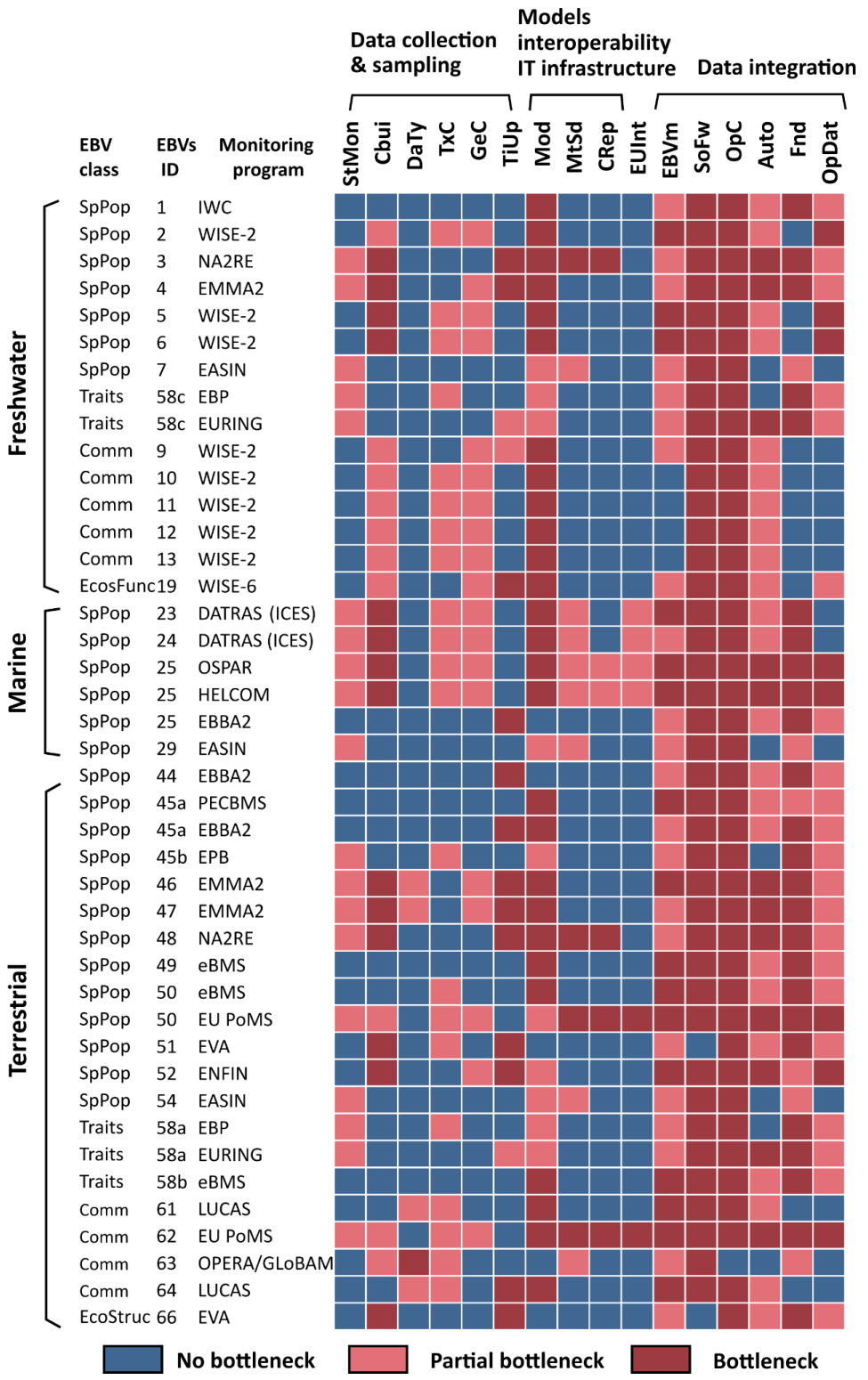


Figure 6. Summary of bottlenecks found across the 36 EBVs evaluated, identified in the table through the EBV ID (Deliverable 4.1) and grouped by EBV class (SpPop- Species populations, Traits - Species traits, Comm - Community composition, EcosFunc - Ecosystem function and EcoStruc - Ecosystems structure). The bottlenecks refer to the data flows of the main biodiversity monitoring programmes which data could potentially serve to generate each EBV (identified by its acronym "monitoring program" column). See Table 1 for a full description of the bottlenecks criteria evaluated (here acronyms shown in columns and classified into three main categories: 1) data collection and sampling; 2) models, interoperability and IT infrastructure and 3) data integration).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

The data needed to generate the EBVs was partially available in 50 % of the monitoring programs evaluated (e.g., available upon demand and only after the agreement of national coordinators and/or after the payment of some data handling fee). This partial availability is strongly linked to two factors: 1) the data ownership: many of these monitoring programs rely on the in unpaid and selfless work of many national coordinators and volunteers that ultimately hold the governance of the data and the decision about whether they agree in sharing their data - and the effort put into collecting and integrating it - to a specific purpose (e.g., scientific research); 2) the sensitivity of part of the data: many of the databases generated in these monitoring programmes contain data on the location of sensitive and threatened species. In 19% of the programmes evaluated data is not available at all: for example, raw monitoring data collected at the waterbody level and used to report water quality status or EQR values within the framework of the Water Framework Directive and the WISE-2 biology data flows are not available at the European level. This data might be available for a few of the national agencies responsible for the data reporting and ultimately may be available for EBV production; however, the **fragmentation in data availability** (i.e. the fact that individual requests have to be made to each of the Member States' environmental agencies to access data) is certainly an impediment to the generation of European-wide EBVs because it may enhance some specific bottlenecks such as the development of modelling applications (see for example Shamoun-Baranes, et al. 2022).

The bottlenecks in data collection and sampling differed on a case-by-case basis, but many of the criteria considered (Geographic coverage - **GeC**, Taxonomic Coverage - **TxC**, whether the data is adequate to generate the EBV or not -**DaTy** or the Temporal resolution- **TiUp**) have been evaluated with more detailed in the EuropaBON Deliverable 3.2 “Report on gaps and important new areas for monitoring Europe” and in the Jessop et al. (2022) “Assessment on t*he current state of Biodiversity Monitoring in the European Union and adjacent marine waters “* for the Marine realm. We highlight here the **bottlenecks in monitoring data for some taxonomic groups,** such as, zooplankton in freshwater systems, lichens, terrestrial arthropods, or crops pests for which we could not find any European-wide monitoring initiative that could provide data to generate the corresponding EBVs at the desired spatio-temporal scale (see Annex IV for EBVs not evaluated in this report). This does not mean there is no data at all for these taxonomic groups but that it is probably only available at the local, subnational, or national level (and as such, better captured in previous efforts to collect data on monitoring capacity at all scales in Europe, e.g., EUMON Wezzel et al. or Moussy et al. 2022). These results only reflect that currently there are no European initiatives making the efforts to retrieve and harmonized all available data for those taxa at the European scale. In fact, there are already good examples of national monitoring network that are collecting data to generate some of these EBVS; for example, the Rothamsted Insect Survey (RIS) in UK has been running two trap networks since 1964 and provide information on aphids, larger moths and many other migrating insects (i.e. collects data suited to generate the EBV “species abundances of terrestrial crop pests”). This initiative generates predictions for the timing of the start of the aphid migration and aphid abundance in spring and early summer. These forecasts are targeted at brassicas, potatoes, and cereal crops and therefore, developed to support farmers and decision-makers.

This project receives funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003553.

Bottlenecks in the generation of marine EBVs

We found partial or total bottlenecks to the generation of marine EBVS for most of the criteria evaluated (Figure 6 - see marine EBVs). One of the main bottlenecks across all marine EBVs assessed is the lack of European integration and data harmonization (**EuInt**). **Monitoring data in European marine waters is very fragmented** and mostly integrated (if at all) at the regional level (e.g. the NorthEast Atlantic region - OSPAR convention, the Baltic Sea - HELCOM convention). Moreover, Southern and Eastern European waters are less well covered by biodiversity monitoring programmes as shown in the *assessment report of the current state of Marine Biodiversity Monitoring in the European Union and adjacent marine waters* (Jessop et al. 2012). This fragmentation hampers the capacity to generate marine EBV metrics across all European waters due to the **lack of data harmonization among different monitoring schemes and regions** (different sampling protocols, metadata standards, different indicators). We could only retrieve information for species populations EBV classes (4 EBVs in total) and lacked the capacity to describe bottlenecks in data flows for some taxonomic groups like hard coral, seagrass, macroalgae or turtles. This is probably explained because the number of monitoring programmes for these taxa is very limited or because of their restricted distribution (i.e. monitoring only concentrates in the small areas where these taxa occur and therefore, the integration of these data only applies at the national level). All in all, our evaluation only corroborates the findings in gaps and data integration challenges detailed in the assessment that Jessop et al. (2022) have recently made on EBVs and gaps in marine biodiversity monitoring and we refer the readers to that report for further details. We envisage as a future priority the application of our framework to the future planning of more comprehensive marine monitoring efforts.

5. Limitations, opportunities and future directions

Scarce representation of EBV classes

One of the limitations of this deliverable is the lack of data representation for the genetic, ecosystem function and ecosystem structure EBV classes, for which we could not retrieve enough information about gaps and bottlenecks in data flows (when none at all).

In some cases, this is explained by the generalized lack of data and data integration at the European level to generate the EBVs under evaluation. This is for example, the case of genetic composition EBVs (three EBVs: genetic diversity of freshwater, marine and terrestrial species). A recent global review on the feasibility of generating genetic EBVs (Hoban et al. 2022), has emphasized the lack of systematic global data collection and monitoring on genetic biodiversity (even for the most studied species taxa such as mammals) as one of the main bottlenecks to operationalize the generation of these EBVs at large scales. In this review, Hoban et al. 2022 mention other hampering aspects to operationalization of genetic EBVs such as the lack of georeferenced sequences in data that is already available and archived (key to generate spatially-explicit predictions of genetic composition EBVs), the lack of harmonization of genetic measurements, the lack of standardized minimum sampling size and spatiotemporal sampling guidelines or the high costs of gathering and processing genetic data.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

EBVs on ecosystems composition and structure were scarcely represented in our assessment as they require data for their generation that is generally not retrieved through standard biodiversity monitoring protocols but rather by remote sensing monitoring programmes. For example, the EBV "*Fire disturbance per habitat type*" (EBV class ecosystem function) could be generated from data collected by the European Forest Fire Information System (EFFIS) that supports the services in charge of the protection of forests against fires in the EU and neighbouring countries. Developed under the umbrella of Copernicus Land Monitoring Service (CLMS) , the EFFIS provides updated and reliable information on wildland fires in Europe. For example, the EFFIS has developed an interactive current situation viewer that detects fires in real-time (updated 6 times daily). The products available in this viewer are burnt area and active fires (from MODIS and Sentinel satellite imagery) but there is not a product that directly informs of the land cover affected by those fires - targeted in this EBV; however, this could potentially be inferred by crossing EFFIS products with other Copernicus land cover products. Similarly, the CLMS produces and disseminates High-Resolution Vegetation Phenology and Productivity product suite (HR-VPP) at high spatial resolution and temporal resolution that matches well the EBV *Terrestrial Ecosystem Phenology* requirements description (EBV class ecosystem function). However, data flows from remote sensing products could not be evaluated in detail with the bottleneck's framework proposed in this deliverable and some of the criteria considered like the use of Standard Monitoring protocols - StMon - or taxonomic completeness - TxC were difficult to interpret in the context of these products.

Relevant European biodiversity data flows not considered

While the WISE-2 Biology database and its data flows to the reporting of the Water Framework Directive (WFD) have been considered as the main integrative initiative collecting data to generate freshwater species populations and community composition EBVs, the data flowing to the reporting of the Habitats (92/43/EEC) and Birds (2009/147/EC) Directives were not considered in our evaluation of bottlenecks. This is because the latter data does not originate from standardized monitoring surveys in all cases (e.g., expert opinion is frequently used to assess the conservation status of specific species/habitats) and it integrates a miscellanea of monitoring programs and species data that do not follow the same standard harmonization procedures in most cases (i.e. it cannot be applied to our framework). In these cases, it can be difficult to understand the flows of data and the criteria and methods used for data collection and integration at local and subnational levels in the EuropaBON biodiversity monitoring data set (as previously noted in Deliverable 3.1). Additionally, it is uncertain if these methods will be suitable for generating the desired terrestrial EBVs at the required spatial and temporal resolution and with the appropriate taxonomic focus. This is in contrast to the reporting of data streams for freshwater systems (such as WISE-2 Biology and WFD), where significant efforts have been made to standardize and harmonize monitoring and data aggregation methods (EQR values).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Future directions and opportunities

Overall, our results suggest that improving current biodiversity monitoring flows to generate European level EBVs will require overcoming a plethora of interaction constraints that come in bundles. In this EuropaBON task, we have developed a framework for the analyses of bottlenecks to the generation of EBVs at the European scale on the basis of the current data flows of European coordinated monitoring networks. **The framework can be used to understand the main factors currently hindering the generation of EBVs** at the desired the spatial, temporal and taxonomic resolution (i.e. the EBV specifications that have been agreed over the course of the EuropaBON project in liaison with stakeholders - Deliverable 4.1). This analysis is shown in detail for 36 different EBVs in Annex I - III and complements efforts conducted by the Biodiversa + platform aiming at similar objectives (Silva del Pozo and body, 2022). Moreover, and as shown in section 3 of this report ("Framework of bottlenecks in practice"), the framework can serve to understand what factors have resulted key for different monitoring networks to overcome bottlenecks in data flows and integration over time (e.g., better standardization of monitoring protocols, definition of metadata standards or by developing webportal tools and apps to facilitate data harmonization and integration tasks). Despite **this framework** has been applied to European level monitoring networks, it **could perfectly be transferred to other scales**, for example to assess bottlenecks in biodiversity data flows at National level or within specific contexts (e.g., integration of biodiversity monitoring data across protected areas).

Within the EuropeaBON project, the results from this task represent the basis to identify actions needed to overcome current bottlenecks in biodiversity monitoring data flows to improve the current European biodiversity monitoring system so that it will be more representative temporally, spatially and taxonomically (EuropaBON task 4.3 co-design), it will maximize benefits (task 4.4 cost-efficiency) and become better integrated into wider biodiversity policy. This framework will be also used in the policy showcases (WP5): for example, the temporal evolution of EBBA Atlases (section 3.1 in this report) will represent as a starting point for the development o the EBBA live atlas that seeks to improve the spatio-temporal resolution of current distribution maps developed by the EBCC and contribute to the reporting of the Birds Directive with standardized data and methods across the whole of Europe (task 5.1, Birds Directive showcase).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

6. References

Chytrý, M., Hennekens, S.M., Jiménez-Alfaro, B., Knollová, I., Dengler, J., Jansen, F., Landucci, F., Schaminée, J.H., Aćić, S., Agrillo, E. and Ambarlı, D. (2016) European Vegetation Archive (EVA): an integrated database of European vegetation plots. *Applied vegetation science*, 19(1), pp.173-180.

Geijzendorffer, I. R., Regan, E. C., Pereira, H. M., Brotons, L., Brummitt, N., Gavish, Y., ... and Walters, M. (2016) Bridging the gap between biodiversity data and policy reporting needs: An Essential Biodiversity Variables perspective. *Journal of Applied Ecology*, 53(5), 1341-1350.

GEO BON (2021). What are EBVs? [Online]. URL.

Hagemeijer, E.J.M., and Blair, M.J. (editors) (1997) The EBCC Atlas of European Breeding Birds: their distribution and abundance. T & A.D. Poyser, London.

Hoban, S., Archer, F. I., Bertola, L. D., Bragg, J. G., Breed, M. F., Bruford, M. W., ... and Hunter, M. E. (2022) Global genetic diversity status and trends: towards a suite of Essential Biodiversity Variables (EBVs) for genetic composition. *Biological Reviews*, 97(4), 1511-1538.

Honrado, J. P., Pereira, H. M., and Guisan, A. (2016). Fostering integration between biodiversity monitoring and modelling. *Journal of Applied Ecology*, *53*(5), 1299-1304.

Jiménez-Alfaro, B., Suárez-Seoane, S., Chytrý, M., Hennekens, S. M., Willner, W., Hájek, M., ... and Tsiripidis, I. (2018). Modelling the distribution and compositional variation of plant communities at the continental scale. *Diversity and Distribution*s, 24(7), 978-990.

Jessop, A., Chow, C., Dornelas, M., Pereira, P., Sousa-Pinto, I., Hernández Chan, S., Junker, J., Soares, J., Ratnarajah, L., Fernández, M. and Mendo, T. (2022) MarBioME. Overview and assessment of the current state of Biodiversity Monitoring in the European Union and adjacent marine waters. European Commission. Directorate General for Research and Innovation.

Keller, V., Herrando, S., Voříšek, P., Franch, M., Kipson, M., Milanesi, P., ... and Foppen, R. P. B. (2020). European breeding bird atlas 2: Distribution, abundance and change. European Bird Census Council and Lynx Editions, 967 pp.

Kissling, W. D., Ahumada, J. A., Bowser, A., Fernandez, M., Fernández, N., García, E. A., ... and Hardisty, A. R. (2018) Building essential biodiversity variables (EBV s) of species distribution and abundance at a global scale. *Biological Reviews*, 93(1), 600-625.

Kühl, H. S., Bowler, D. E., Bösch, L., Bruelheide, H., Dauber, J., Eichenberg, D., ... and Bonn, A. (2020) Effective biodiversity monitoring needs a culture of integration. One Earth, 3(4), 462-474.

Moersberger, H., Martin, J. G., Junker, J., Georgieva, I., Bauer, S., Beja, P., ... and Bonn, A. (2022). Europa Biodiversity Observation Network: User and Policy Needs Assessment. URL.

Moussy, C., Burfield, I.J., Stephenson, P.J., Newton, A.F., Butchart, S.H., Sutherland, W.J., Gregory, R.D., McRae, L., Bubb, P., Roesler, I. and Ursino, C. (2022). A quantitative global review of species population monitoring. *Conservation Biology*, 36(1), p.e13721.

Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H., Scholes, R. J., ... and Wegmann, M. (2013. Essential biodiversity variables. *Science*, 339(6117), 277-278.

Proença, V., Martin, L. J., Pereira, H. M., Fernandez, M., McRae, L., Belnap, J., ... and van Swaay, C. A. (2017). Global biodiversity monitoring: from data sources to essential biodiversity variables. *Biological Conservation*, 213, 256-263.

Shamoun-Baranes, J., Bauer, S., Chapman, J. W., Desmet, P., Dokter, A. M., Farnsworth, A., ...and Leijnse, H. (2022). Meteorological data policies needed to support biodiversity monitoring with weather radar. *Bulletin of the American Meteorological Society*, 103(4), E1234-E1242.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Silva del Pozo, M. and Body, G. (2022) Literature survey on biodiversity monitoring protocols used across countries. MS55 European Biodiversity Partnership (Biodiversa+) Report. URL.

Schaminée J.H.J., et al. (2014) Vegetation analysis and distribution maps for EUNIS habitats. Report EEA/NSV/14/006. EEA, Copenhagen. URL.

Schaminée J.H.J., et al. (2016 a) Review of grassland habitats and development of distribution maps of heathland, scrub and tundra habitats of EUNIS habitats classification. Report EEA/NSV/15/005. European Environment Agency, Copenhagen. URL.

Schaminée J.H.J, et al. (2016 b) Development of distribution maps of grassland habitats of EUNIS habitat classification. Report EEA/NSV/16/005. EEA, Copenhagen.URL.

Wetzel, F.T., Bingham, H.C., Groom, Q., Haase, P., Kõljalg, U., Kuhlmann, M., Martin, C.S., Penev, L., Robertson, T., Saarenmaa, H. and Schmeller, D.S. (2018). Unlocking biodiversity data: Prioritization and filling the gaps in biodiversity observation data in Europe. *Biological conservation*, 221, pp.78-85.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1), 1-9.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ANNEX I: Freshwater EBV fact sheets

Freshwater species populations

EBV: Species abundances of wetland birds	
ID	1
Realm	Freshwater
EBV class	Species populations
EBV name	Species abundances of wetland birds
Step in identification process	Expert workshop
Definition	The estimated count of individuals / modeled relative abundance of European Wetland bird species within contiguous spatial units (grid cells) over time.
Metric	- Estimated count of individuals in winter - Modeled relative abundance in winter
Spatial resolution unit	Wetlands as defined by The Critical Site Network Tool CSN which is an online resource for the conservation of 312 species of waterbirds and the important sites upon which they depend in Africa and Western Eurasia.
Temporal resolution unit	1 year
Taxonomic/ ecosystem focus group	Wetland birds (taxonomy based on the HBW and BirdLife Taxonomic Checklist, with focus on those bird species that are officially recognized in the List of birds of the European Union, and wetland affiliation defined as the linkages of species and habitat types to MAES [wetland] ecosystems)

Main European initiatives and description of current bottlenecks

The International Waterbird Census (IWC) collects abundance data for 554 migratory waterbird populations across the African-Eurasian census on an annual basis. It explicitly estimates population trends of waterbirds within the European Birds Directive. In addition, the design of the IWC ensures that all wetland types are represented in the census, as well as wetlands with different conservation status (national, international, unprotected).

Possible bottlenecks in the generation of this EBV from existing data collected by the IWC network are:

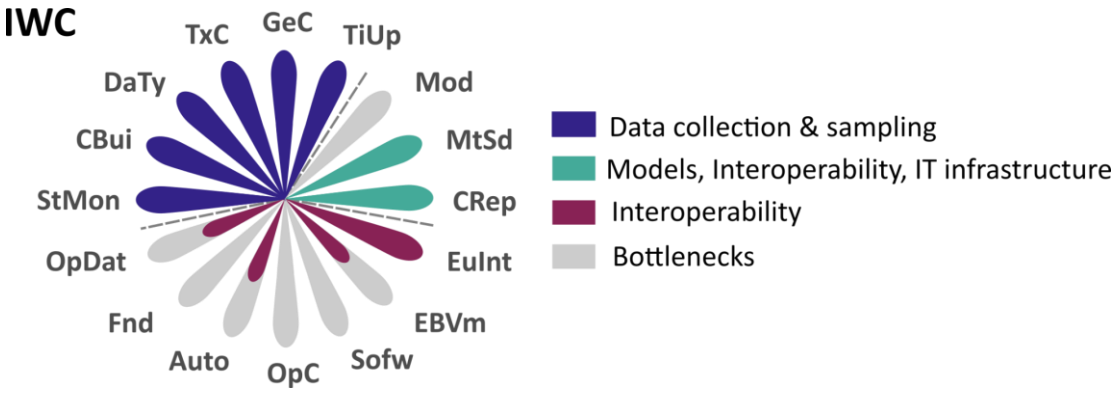
1) Data collection. Waterbird winter censuses are based on long-term counts conducted by citizen scientists on the main European water bodies and are organized at the national and subnational level. In this regard, the type of data collected by this monitoring program matches well the data needed to generate this EBV. Surveys are conducted every year in winter (most of them in January) following standard protocols. Data gets reported at the site level so it could potentially be aggregated or modelled to generate this EBV at the desired spatial resolution. The initiative offers regular training to volunteers and national coordinators contributing to capacity building of the monitoring program.

2) Models, data interoperability and IT infrastructure. The IWC uses models (TRIM) to estimate species trends at the national level following the same methods used by the Pan-European Common Bird Monitoring Scheme (PECBMS). However, no spatially explicit abundance models are generated within the framework of this

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

program (**Mod** bottleneck). Metadata standards are followed to allow harmonization and integration of data from different countries who get stored in a single central repository (global IWC database) managed by the Wetlands International.

3) Data integration. Species counts at the site level get aggregated at the national level and species trends are also estimated at the national scale. Therefore, there is a mismatch between the EBV description, and the products generated by this integration initiative (species trends at site and national level; **EBVm** partial bottleneck). Since this integration initiative does not explicitly model species abundances over space, there is no open code or user-friendly software to report in this regard (**Sofw** and **OpC** bottlenecks**)**, although as with the PECBMS program, it is to be expected that if this initiative were to derive spatially explicit abundance estimates, it would do so by relying on open software and code, as it currently does for estimating species population trends (using the *rtrim* package and TRIM model: TRends and Indices for Monitoring data). Funding is also a limitation: the international coordination, data management and systems are supported by Wetlands International members and the EU LIFE NGO grant; however, in its 2020 annual report, the IWC states "Unfortunately no sustainable funding has been secured for this important work, which depends on the 15,000 volunteers that will be collecting the data in the coming days." (**Fnd** bottleneck). The field monitoring (species counts) is mostly funded by national governments or organizations, but the Waterbird Fund supports counts when these primary funding sources fail. Data flows are partially automated: National coordinators can submit counts directly to the IWC online platform. All counts must be submitted in a standardized excel format and the system performs several automatic checks. Additional manual checks are also undertaken regularly by Wetlands International staff. The IWC has no plan to develop a system to collect data directly from counters at sites because (a) IWC depends on the network of national coordinators and their teams to perform essential data quality control and coordination for the data collected in their countries and (b) there are already many different apps available to collect waterbird count data in the field. These are widely used by counters but still the IWC relies on the coordinators to collect and review data reported on these apps to ensure they meet the standard of an IWC count (**Auto** partial bottleneck). Data is available upon request and after the payment of a data handling fee but always requires the authorization of national coordinators, who hold the governance data rights (**OpDat** partial bottleneck).



Many passerine bird species with wetland affiliation as defined as the linkages of species and habitat types to MAES [wetland] ecosystems) are monitored by the Pan European

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Common Bird Monitoring Scheme (PECBMS). However, the sampling design does not specifically focus on wetlands, surveys are mostly conducted during the breeding period and it is difficult to estimate the coverage of these ecosystems by PECBMS sampling transects. A detailed description of the bottlenecks to generation of spatially-explicit abundance estimates from PECBMS data is detailed in the fact sheet *Species Abundances Terrestrial Birds: COMMON BIRDS.*

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species distributions of freshwater fishes	
ID	2
Realm	Freshwater
EBV class	Species populations
EBV name	Species distributions of freshwater fishes
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of each European freshwater fish species within lakes and rivers catchments over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	Lakes and rivers as delineated in ECRINS (European catchments and rivers network system)
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Freshwater fishes listed in the European Red List of Freshwater Fishes (currently 531 native and described European species)

Main European initiatives and description of current bottlenecks

The **WISE-2** (Water Information System for Europe - Biology data), is an ongoing European monitoring data collection scheme established by the European Environment Agency (EEA) to obtain a harmonized flow of biology monitoring data from all EEA member states reported as Ecological Quality Ratios (EQRs) from all surface water categories: rivers, lakes, transitional and coastal waters. This data flow is tightly linked to the reporting duties of Member States within the regulations of the Water Framework Directive.

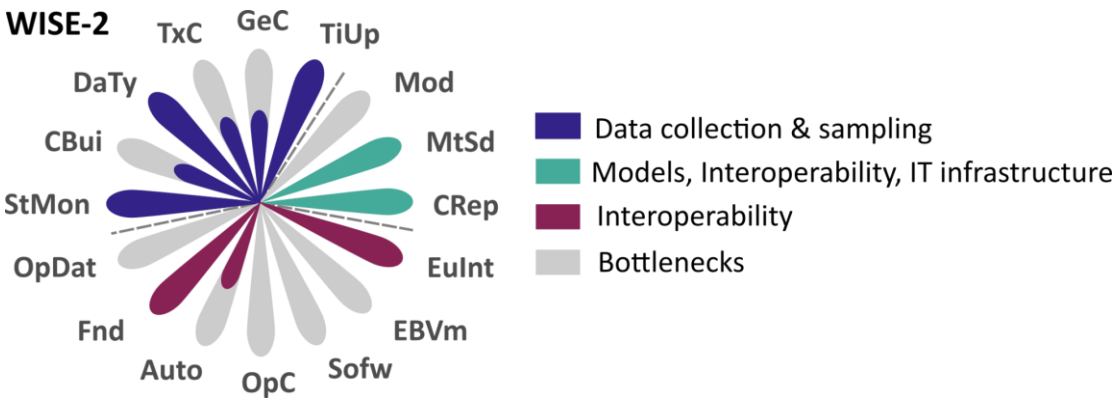
Possible bottlenecks in the generation of this EBV from existing data collected by WISE-2 data relate to:

1) Data collection and sampling. Currently, only 10 European countries report FISH_EQRs to the WISE-2 so there is a gap in the reporting of fish data to the European Commission, that probably translates into a gap in taxonomic coverage given many fish species are endemic of specific river catchments (**GeC, TxC** partial bottlenecks). However, 27 countries collect data on an annual basis on fish species for the Water Framework Directive flows, so probably the geographic and taxonomic monitoring gaps are smaller than expected by WISE-2 data flows. The WISE-2 data flow, as an integration initiative, organizes yearly webinars for all member states with information and Q&A sessions on data reporting, but does not have capacity building on monitoring. If this exists, it is the capacity building of the member states (**CBui** partial bottleneck). While the reported data are EQRs, the underlying raw data collected (composition, abundance, and age structure of fish fauna at the waterbody level) could be potentially used to generate this EBV.
2) Models, interoperability, and IT infrastructure. Currently, there are no models in place to map the distribution of individual freshwater fish species across lakes and rivers within the framework of the WISE-2 flows (**Mod** bottleneck). The EEA centralizes (CDR - Central Data Repository) the EQRs reported by member states and integrates the reporting values to obtain overall EU water quality values representing the specified BQE (biological quality element) and pressure type (e.g. eutrophication). Reporting by Member States follows metadata standards (the structure and contents of the data are described in a Data Dictionary).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023.* DOI: https://doi.org/10.3897/arphapreprints.e103765

3) Data integration. Freshwater fish data reported to the WISE-2 data flow system (EQRs) does not match the definition of this EBV (**EBVm** bottleneck). Countries report FISH_EQRs but these are based on fish records at monitoring sites. Initially, species distributions could be retrieved from raw data/monitored data at the waterbody level and use that information to infer/model species distributions at the European scale. Data flows are not fully automated: the EEA provides empty templates for download from the data dictionary as Excel file and XML Schema.There is a reporting tool (ReportNet) where to upload these data.
An automatic quality control (QC) process is implemented on the Central Data Repository, to check the structure and content of the data file(s) uploaded by each member state. This first step of QC is followed by three more QC steps that involve (2) cross-checking against previously reported data, (3) statistical tools and other algorithms for identifying suspicious values, and (4) expert-based checking and communication with data reporters. Automation of data flows between field data collection and national coordinators yet to be explored (some tests in Norway are underway) (**Auto** partial bottleneck). Since there are not currently models in place to map the distribution of freshwater species, there is no open code or user-friendly software already being used to generate this EBV from WISE-2 data (**OpC** and **SoFw** bottlenecks). Funding must be secured by member states to comply with reporting obligations. While fish EQR ratios at the waterbody level are openly available, the raw data that would serve to generate this EBV are not (species lists collected at each sampling site; **OpDat** bottleneck).



Another integration initiative at the European level retrieving information on freshwater fishes distributions is the IUCN Red List of Freshwater Species. However, most of the time IUCN maps are not linked to a monitoring program (but expert based) and represent a snapshot of the species distributions so this initiative couldn´t contribute to the quantification of this EBV at the desired spatio-temporal scale. Part of the IUCN evaluation is based on the European Handbook of Freshwater fishes covering 546 native and 33 introduced species.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species distributions of amphibians and freshwater reptiles	
ID	3
Realm	Freshwater
EBV class	Species populations
EBV name	Species distributions of amphibians and freshwater reptiles
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of each European amphibian and freshwater reptile species within contiguous spatial units (grid cells) over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	10 × 10 km – 50 × 50 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	All European amphibians and freshwater reptiles
Main European initiatives and description of current bottlenecks	

The New Atlas of Amphibians and Reptiles of Europe (NA2RE, 2014) is the only Europe-wide coordinated initiative that has attempted to map the distribution of amphibians and reptiles across the continent at 50 x 50 km grid. It generated maps for 218 taxa (73 species of amphibians and 145 of reptiles) updated as of 2014 (no information about temporal dynamics).

Possible bottlenecks in the generation of this EBV from data flows in NA2RE relate to

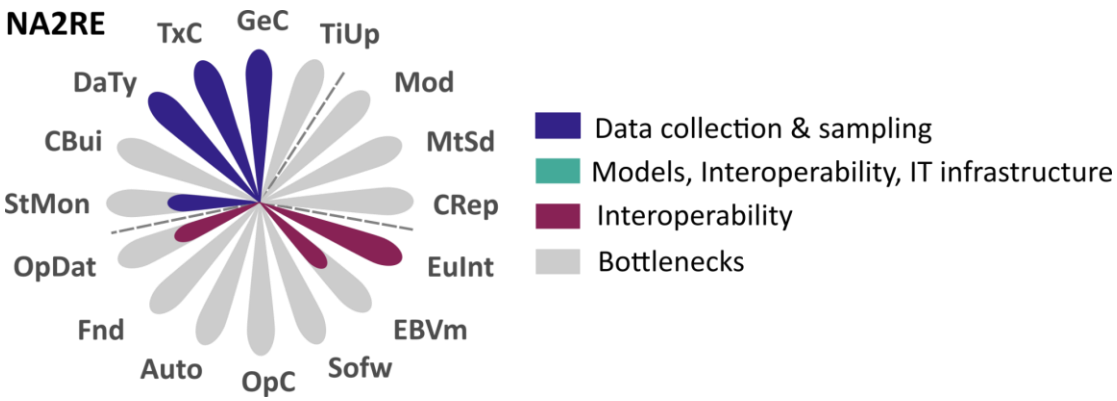
1) Data collection and sampling. Most of the data used for the Atlas compilation did not come from systematic monitoring programs (**StMon** partial bottleneck). Data sources included: (1) data published (in books or websites) or ongoing national atlases, (2) personal data kindly provided to the *Societas Europaea Herpetologica*, (3) the 1997 European Atlas, and (4) the Global Information Facility (GBIF). While data is available across most European countries (23; i.e., it does fulfil the criteria set to qualify for full Geographic Coverage in this assessment), the data coverage within each country was uneven and very scarce in many cases, with lack of funding and personnel being identified as limiting factors to set up national databases in many countries. The initiative did not have capacity building (if it existed, it was that of the national nodes) (**CBui** bottleneck). The Atlas is a snapshot of the distribution of amphibians and reptiles and therefore, does not match the temporal resolution criteria of this EBV (**TiUp** bottleneck).

2) Models, interoperability, and IT infrastructure. Maps in NA2RE reflect an aggregation of the raw data at 50 x 50 km and there is no modelling involved in predicting species distributions across areas not covered by data (**Mod**, **OpC** and **Sofw** bottlenecks). Data from different sources (collected using different standards) were centralized and harmonized for the atlas (**MtSd** bottleneck); the authors of the atlas recognize the challenge of finding better ways to gather species occurrence data across Europe and centralized it (**CRep** bottleneck) (Sillero et al. 2014).

3) Data integration. The maps generated by this integration initiative do not perfectly match the EBV definition (especially the criteria of temporal resolution; **EBVm** partial bottleneck). Data streams were not automated (**Auto** bottleneck). Funding

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

supported the publication of the Atlas, but it is not available in the long-term to generate this EBV with the temporal frequency needed (**Fnd** bottleneck). The NA2RE maps are freely available for download but the raw data are not (**OpDat** partial bottleneck).



Other integration initiatives at the European level retrieving information on amphibians and freshwater reptiles distributions are the IUCN European Red Lists of Amphibians and Reptiles. However, most of the time IUCN maps are not linked to a monitoring program (but expert based) and represent a snapshot of the species distributions, so this initiative couldn´t contribute to the quantification of this EBV at the desired spatio-temporal scale.

References:

Sillero, N., Campos, J., Bonardi, A., Corti, C., Creemers, R., Crochet, P. A., ... and Vences, M. (2014). Updated distribution and biogeography of amphibians and reptiles of Europe. *Amphibia-reptilia*, 35(1), 1-31.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species distributions of freshwater mammals	
ID	4
Realm	Freshwater
EBV class	Species populations
EBV name	Species distributions of freshwater mammals
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of each European freshwater mammal species within contiguous spatial units (grid cells) over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	10 × 10 km – 50 × 50 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Freshwater mammal species listed in the Annex II and Annex IV of the Habitats Directive (*Lutra lutra*, *Galemys pyrenaicus*, *Mustela lutreola*)

Main European initiatives and description of current bottlenecks

The Atlas European Mammals is the first coordinated initiative that has attempted to map the distribution of mammals at 50 x 50 km resolution across Europe (including the freshwater mammals focus of this EBV). The First Atlas of European Mammals was published in 1999 (Mitchell-Jones et al. 1999) but the Second Atlas of European Mammals (EMMA2) is expected to be released in 2023.
Possible bottlenecks in the generation of this EBV from existing data collected by EMMA2 data relate to:

1) Data collection and sampling. Most of the data used for the EMMA2 compilation will not come exclusively from systematic monitoring programs (**StMon** partial bottleneck). Data sources include national and regional databases, literature records and data portals. The integration initiative does not have capacity building (if it exists, it depends on national initiatives). As a result of the former points, large gaps of information are expected, particularly in some areas (**GeC** partial bottleneck; 21 countries). The spatial resolution of EMMA2 maps matches the maximum resolution allowed in the definition of this EBV; however, the temporal resolution (frequency) of atlas updates is too low (> 10 years) (**TiUp** bottleneck), therefore the products generated by this integration initiative only partially match the EBV definition (see **EBVm** partial bottleneck in *Data Integration*).

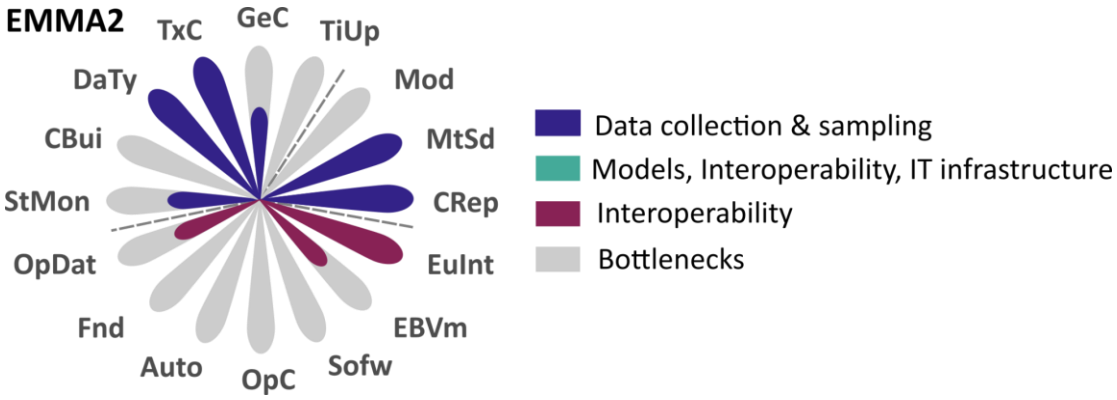
2) Models, interoperability, and IT infrastructure. EMMA2 maps will reflect the confirmed presence of the species aggregated at 50 x 50 km and there is no modelling involved in predicting species distributions across areas not covered by data (**Mod, OpC, Soft** bottlenecks). Metadata standards are defined (with the intention of following Darwin Core) and data get integrated in a central repository by the European Mammal Foundation.

3) Data integration. The products generated by EMMA2 only partially match the definition of this EBV (there is a mismatch in terms of temporal resolution but also, maps are only available at 50 x 50 km and not at 10 x 10 km; **EBVm** partial bottleneck). As there are no models in place within the framework of this initiative, there is no user-friendly software or open code to report in this regard (**Sofw** and **OpC** bottlenecks). While metadata standards are defined, data flows are not

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

automated (**Auto** bottleneck): records are to be collected into a CSV file or Access database (or similar) and submitted via email to coordinators who harmonize and integrate the data from the different sources. Funding has been secured for the publication of the Atlas but there is no funding to support the long-term continuation of this initiative (**Fnd** bottleneck). The EMMA2 original records will be held in a system that is accessible to researchers, with appropriate safeguards for sensitive records (rare species, vulnerable habitats or other reasons), though the great majority of records will be freely accessible at a higher resolution than in the Atlas, for example in national or regional atlases (**OpDat** partial bottleneck).



It should be noted here that one of the target species of this EBV (*Galemys pyrenaicus*) is endemic to the Iberian Peninsula and therefore, there is no coordinated initiative at the European level for the monitoring of the species. However, the species is the subject of several regional monitoring programs.

Also, noteworthy here is the development of new integration initiatives such as European Observatory of Wildlife that could contribute data to the generation of this EBV. The EOW (a project by ENETWILD) aims to enhance collaboration among the stakeholders that monitor, conserve, and manage wildlife in Europe (with a special focus on mammals). It seeks to develop a framework where data will be comparable, interoperable, and openly accessed at the European level, providing guidelines for monitoring mammals in Europe (e.g., protocols to estimate wildlife density), training (e.g., on new tools for data processing and analysis) and compiling and harmonize existing databases on mammals. Data gathered through this portal following standardized monitoring protocols could eventually be used to map the distribution of mammal species across Europe.

References:

Mitchell-Jones et al. (Eds) 1999 The Atlas of European Mammals. *Academic press*, p 484

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species distributions of freshwater invertebrates	
ID	5
Realm	Freshwater
EBV class	Species populations
EBV name	Species distributions of freshwater invertebrates
Step in identification process	Internal review process
Definition	The presence/absence or probability of occurrence of invertebrate species within lakes and rivers catchments over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	Lakes and rivers as delineated in ECRINS (European catchments and rivers network system)
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Freshwater invertebrate species listed in the Habitats Directive Annex II: Dragonflies: *Coenagrion hylas, C. mercuriale, Cordulegaster trinacriae, Gomphus graslinii, Leucorrhina pectoralis, Lindenia tetraphylla, Macromia splendens, Ophiogomphus cecilia, Oxygastra curtisii* Bivalves: *Margaritifera margaritifera* and *Unio crassus* Freshwater invertebrate species listed in the Habitats Directive Annex IV: Dragonflies: *Aeshna viridis, Cordulegaster trinacriae, Gomphus graslinii, Leucorrhina albifrons, L. caudalis, L. pectoralis, Lindenia tetraphylla, Macromia splendens, Ophiogomphus cecilia, Oxygastra curtisii, Stylurus flavipes* and *Sympecma braueri* Bivalves: *Lithophaga lithophaga, Pinna nobilis, Margaritifera auricularia* and *Unio crassus* Pollution-sensitive benthic invertebrates that are monitored for the Water Framework Directive: Mayflies (Ephemeroptera), Stoneflies (Plecoptera), Caddisflies (Trichoptera)
Main European initiatives and description of current bottlenecks	

Dragonflies

The only integration initiative at the European level retrieving information on dragonflies distributions are the IUCN European Red Lists of Dragonflies (2010). However, the IUCN maps shown in that report (overall species richness and richness of endemic species) are not linked to a monitoring program (but based on expert consultation and literature review) and represent a snapshot of the species distributions, so this initiative couldn´t contribute to the quantification of this EBV at the desired spatio-temporal scale.

There is no European wide integration initiative monitoring bivalves that we could report here.

Pollution-sensitive benthic invertebrates

The **WISE-2** (Water Information System for Europe - Biology data), is an ongoing European monitoring data collection scheme established by the European Environment Agency (EEA)

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

to obtain a harmonized flow of biology monitoring data from all EEA member states reported as Ecological Quality Ratios (EQRs) from all surface water categories: rivers, lakes, transitional and coastal waters. This data flow is tightly linked to the reporting duties of Member States within the regulations of the Water Framework Directive.

Possible bottlenecks in the generation of distribution models of pollution-sensitive benthic invertebrates from existing data collected by WISE-2 data relate to:

1) Data collection and sampling. Currently, only 25 European countries report benthic invertebrates in rivers to the WISE-2 data flows so there is a geographic coverage gap that probably translates into a taxonomic coverage gap (**GeC, TxC** partial bottlenecks). However, 28 countries collect data on an annual basis on freshwater invertebrates on rivers and lakes for the Water Framework Directive flows, so probably the geographic and taxonomic monitoring gaps are smaller than expected by WISE-2 data flows. The WISE-2 data flow, as an integration initiative, organizes yearly webinars for all member states with information and Q&A sessions on data reporting, but does not have capacity building on monitoring. If this exists, it is the capacity building of the member states (**CBui** bottleneck). While the reported data are EQRs, the underlying raw data collected (composition, abundance of benthic invertebrates at the waterbody level) could be potentially used to generate this EBV.

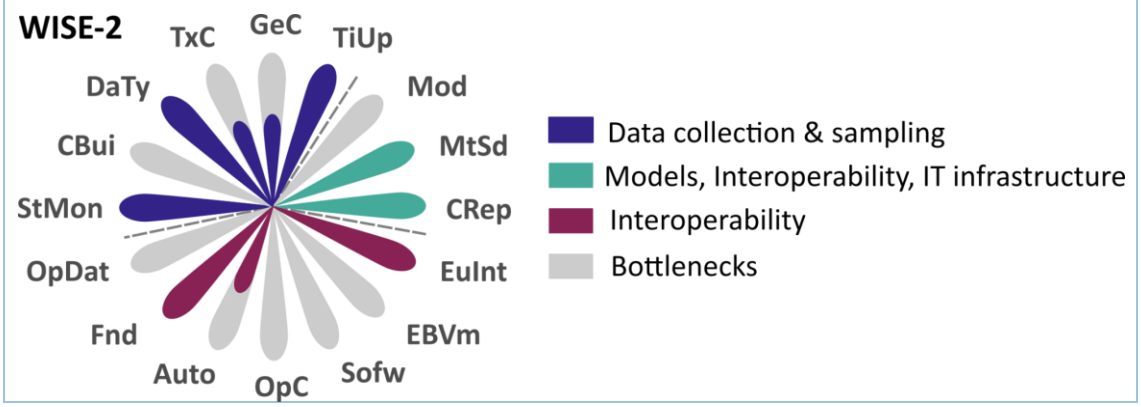
2) Models, interoperability, and IT infrastructure. Currently, there are no models in place to map the distribution of benthic invertebrates across lakes and rivers within the frameworks of the WISE-2 flows (**Mod** bottleneck). The EEA centralizes (CDR - Central Data Repository) the EQRs reported by member states and integrates the reporting values to obtain overall EU water quality values representing the specified BQE (biological quality element) and pressure type (e.g. eutrophication). Reporting by Member States follows metadata standards (the structure and contents of the data are described in a Data Dictionary).

3) Data integration. Benthic invertebrate data reported to the WISE-2 data flow system (EQRs) does not match the definition of this EBV (spatially-explicit species distributions; **EBVm** bottleneck). Countries report EQRs but these are based on benthic invertebrates' records at monitoring sites. Initially, species distributions could be retrieved from raw data/monitored data at the waterbody level and use that information to infer/model species distributions at the European scale. Since there are not currently models in place to map the distribution of freshwater species, there is no open code or user-friendly software already being used to generate this EBV from WISE-2 data (**OpC** and **SoFw** bottlenecks). Data flows are not fully automated: the EEA provides empty templates for download from the data dictionary as Excel file and XML Schema. There is a reporting tool (ReportNet) where Member States can upload this data. An automatic quality control (QC) process is implemented on the Central Data Repository, to check the structure and content of the data file(s) uploaded by each member state. This first step of QC is followed by three more QC steps that involve: (2) cross-checking against previously reported data, (3) statistical tools and other algorithms for identifying suspicious values, and (4) expert-based checking and communication with data reporters. Automation of data flows between field data collection and national coordinators yet to be explored (some tests in Norway are underway) (**Auto** partial bottleneck). Funding has to be secured by member states to comply with reporting obligations. While benthic invertebrates EQR ratios at the

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

waterbody level are openly available, raw data are not (species lists; **OpDat** bottleneck).



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species distributions of freshwater macrophytes	
ID	6
Realm	Freshwater
EBV class	Species populations
EBV name	Species distributions of freshwater macrophytes
Step in identification process	Internal review process
Definition	The presence/absence or probability of occurrence of European freshwater macrophyte species within lakes over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	Lakes as delineated in ECRINS (European catchments and rivers network system)
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	European macrophytes

Main European initiatives and description of current bottlenecks

The **WISE-2** (Water Information System for Europe - Biology data), is an ongoing European monitoring data collection scheme established by the European Environment Agency (EEA) to obtain a harmonized flow of biology monitoring data from all EEA member states reported as Ecological Quality Ratios (EQRs) from all surface water categories: rivers, lakes, transitional and coastal waters. This data flow is tightly linked to the reporting duties of Member States within the regulations of the Water Framework Directive.

Possible bottlenecks in the generation of this EBV from existing data collected by WISE-2 data relate to:

1) Data collection and sampling. Currently, only 15 European countries report macrophytes in lakes to the WISE-2 data flows so there is a geographic coverage gap that probably translates into a taxonomic coverage gap (**GeC**, **TxC** partial bottlenecks). However, 19 countries collect data on an annual basis on freshwater macrophytes in lakes for the Water Framework Directive flows, so probably the geographic and taxonomic monitoring gaps are smaller than expected by WISE-2 data flows. The WISE-2 data flow, as an integration initiative, organises yearly webinars for all member states with information and Q&A sessions on data reporting, but does not have capacity building on monitoring. If this exists, it is the capacity building of the member states (**CBui** bottleneck). While the reported data are EQRs, the underlying raw data collected (composition - presence/absence and abundance of aquatic flora at the waterbody level) could be potentially used to generate this EBV.

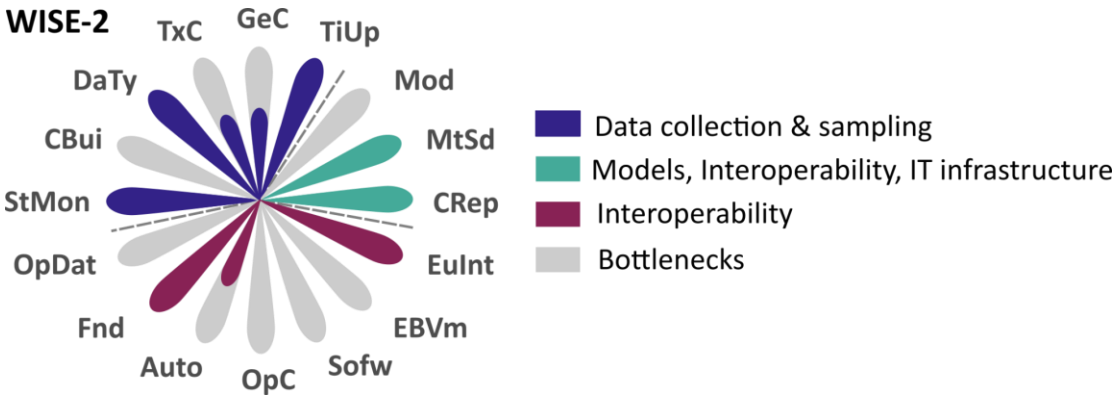
2) Models, interoperability, and IT infrastructure. Currently, there are no models in place to map the distribution of macrophytes across lakes within the frameworks of the WISE-2 flows (**Mod** bottleneck). The EEA centralizes (CDR - Central Data Repository) the EQRs reported by member states and integrates the reporting values to obtain overall EU water quality values representing the specified BQE (biological quality element) and pressure type (e.g., eutrophication). Reporting by

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

Member States follows metadata standards (the structure and contents of the data are described in a Data Dictionary).

3) Data integration. Macrophytes data reported to the WISE-2 data flow system (EQRs) does not match the definition of this EBV (spatially-explicit species distributions; **EBVm** bottleneck). Countries report Macrophytes EQRs in lakes but these are based on macrophyte records at monitoring sites (composition - presence/absence and abundance of aquatic flora). Initially, species distributions could be retrieved from raw data/monitored data at the waterbody level and use that information to infer/model species distributions at the European scale. Since there are not currently models in place to map the distribution of freshwater species, there is no open code or user-friendly software already being used to generate this EBV from WISE-2 data (**OpC** and **SoFw** bottlenecks). Data flows are not fully automated: the EEA provides empty templates for download from the data dictionary as Excel file and XML Schema. There is a reporting tool (ReportNet) where Member States can upload these data. An automatic quality control (QC) process is implemented on the Central Data Repository, to check the structure and content of the data file(s) uploaded by each member state. This first step of QC is followed by three more QC steps that involve: (2) cross-checking against previously reported data, (3) statistical tools and other algorithms for identifying suspicious values, and (4) expert-based checking and communication with data reporters. Automation of data flows between field data collection and national coordinators yet to be explored (some tests in Norway are underway) (**Auto** partial bottleneck). Funding has to be secured by member states to comply with reporting obligations. While Macrophytes EQR ratios at the waterbody level are openly available, raw data are not, hampering its access for the development of distribution models for this taxonomic group (**OpDat** bottleneck)



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species distributions of invasive alien freshwater taxa of European concern	
ID	7
Realm	Freshwater
EBV class	Species populations
EBV name	Species distributions of invasive alien freshwater taxa of European concern
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of invasive freshwater species within lakes and river catchments over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	Lakes and river catchments as delineated in ECRINS (European catchments and rivers network system)
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Freshwater species specified in the Consolidated List of Invasive Alien Species of Union Concern

Main European initiatives and description of current bottlenecks

The European Alien Species Information Network (EASIN) is an integration initiative coordinated at the European level (Joint Research Centre, European Commission) that aggregates, integrates, and harmonizes spatio-temporal data for alien species (AS) and IAS across Europe. The species catalogue and geodatabase are regularly updated following the continuous revision of species reports in the literature, data published by EASIN data partners and the official reports by Member States competent authorities. The EASIN catalogue (v9.0- 19.07.22) includes 740 freshwater invasive alien species (IAS), of which 39 are AS of Union Concern (EU Regulation 1143/2014). EASIN facilitates the exploration of data and information from exiting monitoring networks and programs available from a variety of distributed information sources by providing tools and interoperable web services, compliant with internationally recognized standards.

Possible bottlenecks in the generation of this EBV from existing data collected by EASIN data relate to:

1) Data collection and sampling. EASIN carries out the systematic collection of data on IAS in Europe, does not come exclusively from standardized monitoring programs but from a variety of sources, including literature review, occasional observations and data portals (**StyMon** partial bottleneck); EASIN has demonstrated capacity building through different activities including: i) ad hoc assessments, such as the Baseline Distribution of IAS of Union concern; ii) support to the establishment of surveillance systems in compliance with EU AS policy; iii) citizen science programs, (e.g. the development of the smartphone app 'IAS in Europe' freely available and already adopted in the context of several projects and for official monitoring; iv) training to teachers and students (e.g., the MOOC course "Have you seen an alien?"), among others. EASIN collaborates with the Member States competent authorities and national experts to update and validate the datasets relevant in the EU policy context, to fulfill the EU mandate of setting up national monitoring networks for IAS; however, the extent to which this has been implemented is not fully known. For example, in the freshwater realm an evaluation of 2010 showed

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

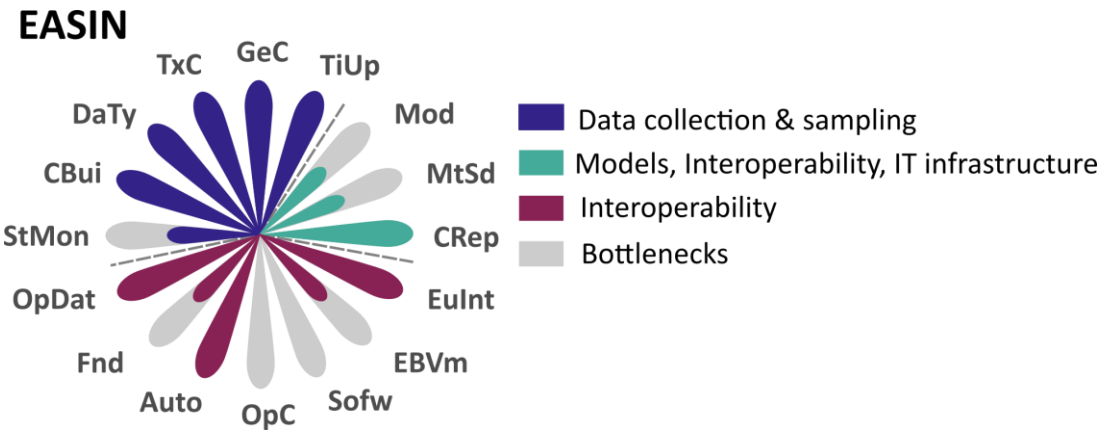
that the large majority of Member States did not directly account for the effects of IAS in the ecological status values reported to the Water Framework Directive (Vandekerkhove and Cardoso 2010) (see also Boon et al. 2020 and Vandekerkhove et al. 2013). Some bottlenecks regarding data collection may relate to 1) lack of communication between national coordinators & EU funded projects working on IAS in a country (e.g. LIFE projects); 2) at national level people in charge of monitoring AS belong to different admin bodies among which there is little communication and coordination and 3) complex administrative structure in some countries (federal, regions, etc.) and how the environmental competences are distributed among them, making it difficult to standardize and coordinate data collection tasks and to facilitate data flows and integration. Currently, the EASIN AS Geo Database contains occurrence records for more than 14,000 species, across 40 different countries (EU & beyond) (including data for the 88 species in the consolidated list of IAS of Union concern- European Union (EU) Regulation 1143/2014). Because the data is georeferenced, it could potentially be used to build species distribution models. The EASIN GeoDatabase is updated on average 3 times/ year, but some information sources are updated less regularly depending on effort for the update and the periodicity of their own updates. Recent EASIN database backend developments data updating will facilitate more regular updates.

2) Models, interoperability, and IT infrastructure. Habitat suitability models have been already fit for the freshwater invasive species *Elodea nuttalli* from EASIN data across EU (Steen et al. 2019); however, this is not a product regularly generated by this initiative, but a case study (**Mod** partial bottleneck), although there is ongoing work in this direction for IAS of Union concern: models are used to evaluate the range of expansion at both European and national levels, to understand whether niche of the species is at equilibrium in its invaded range or to predict the probability of occurrence of the species in Europe considering bioclimatic variables. The AS GeoDatabase (the central repository of EASIN curated by the Joint Research Centre of the European Commission) has developed a protocol to frequently retrieve data from the data partners; data are subsequently transformed by converting the harvested data to the EASIN Data Model through the following steps: validation, cleansing and standardization, geocoding, mapping, application of quality rules and finally loaded on the Geodatabase (Datawarehouse). However, while standards have been defined to harmonize data across source databases, these do not necessarily apply to the underlying data (**MtSd** partial bottleneck).

3) Data integration. The products generated by this initiative partially match the EBV, especially in terms of spatial resolution: the EASIN species mapping tool shows the distribution of species at the country level, river basin districts or at 10 x 10 km grid cells but not at the river catchment level (**EBVm** partial bottleneck). The test run on modeling species distributions has used Maxent, a machine-learning presence-background model that has a user-friendly interface facilitating its use by non-modellers (Phillips et al. 2006, 2008). However, this model is not routinely used by EASIN to model the data in the Alien species geodatabase, so there is no user-friendly software or open code to report in this regard (**Soft** and **OpC** bottlenecks). Data flows are automatized at different levels, from data collection to data integration: the EASIN has developed the "IAS Europe'' smartphone App to promote the report of sightings of IAS of Union Concern by citizens and their integration into the EASIN GeoDatabase (and its consequent harmonization with other data retrieved from Data Partners). The process of retrieving the data from

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

the Data Partners is done through the EASIN Data Broker system, which is able to retrieve the species occurrences and related information (date, source) from different kinds of data sources and store them in a normalized database structure. The EASIN has also developed and manages the NOTSYS platform as the official tool for EU Member States to notify the Commission and inform the other Member States as required by Reg. 1143/2014 on Invasive Alien Species (IAS). In particular, the tool has been designed to facilitate a timely comprehensive notification of new detections of IAS of Union concern and related eradication measures. Funding is a key bottleneck to the generation of this EBV: while the European Commission has set up EASIN as the official information system supporting Member States in the implementation of the Regulation (EU) No 1143/2014 , it does not directly co-participate in the expenses of either monitoring or eradication actions by national nodes which solely relies on national budgets; funding from the European Commission for monitoring of IAS is available via existing financial instruments such as LIFE, H2020, Cohesion or Regional Development funds; however, funding is warrant in the mid - long - term to maintain the EASIN data infrastructure and the development of new data collection tools (e.g., a web base platform for reporting citizen science data) (**Fnd** partial bottleneck). Data in the EASIN GeoDatabase can be easily accessed and downloaded from the website. The openness of the data is key to redirect surveillance and trigger early warning systems.



Monitoring data of invasive species to generate this EBV can also be retrieved from other initiatives described in detail in other EBVs , for example the WFD (freshwater fishes, macrophytes, and benthic invertebrates), the NA2RE (amphibians and freshwater reptiles) or the EMMA2 (freshwater mammals).

References:

Boon, P. J., Clarke, S. A., and Copp, G. H. (2020). Alien species and the EU Water Framework Directive: A comparative assessment of European approaches. *Biological Invasions*, *22*(4), 1497-1512. URL.

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological modelling*, *190*(3-4), 231-259.

Phillips, S. J. and Dudík, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31(2), 161-175.

Steen, B, Cristina Cardoso, A.C., Tsiamis, K., Nieto, K.,Engel J. and Gervasini, E. (2019) Modelling hot spot areas for the invasive alien plant Elodea nuttallii in the EU. *Management of Biological Invasions* 10(1): 151–170. URL .

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Vandekerkhove, J. and Cardoso, A.C. (2010) Alien species and the Water Framework Directive. Questionnaire results. European Commission - Joint Research Centre. URL.

Vandekerkhove, J., Cardoso, A. C., and Boon, P. J. (2013). Is there a need for a more explicit accounting of invasive alien species under the Water Framework Directive?. *Management of Biological Invasions*, *4*(1), 25. URL.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Freshwater species traits EBVs

EBV: Phenology of migration of wetland birds	
ID	58c
Realm	Freshwater
EBV class	Species traits
EBV name	Phenology of migration of terrestrial birds
Step in identification process	Expert workshop
Definition	The annual timing of arrival and departure of European wetland migratory bird species at breeding, staging and wintering sites over time.
Metric	Migration phenology metrics such as: - Day of arrival - Day of departure - Length of stay
Spatial resolution unit	10 × 10 km
Temporal resolution unit	1 week
Taxonomic/ ecosystem focus group	Migratory bird species defined as full migrants in the European Red List

Main European initiatives and description of current bottlenecks

There are two European-wide coordinated initiatives that collect data that could potentially be used to generate this EBV: The EuroBirdPortal (EBP) and the EURING.

The **EuroBirdPortal (EBP)** combines the data collected by the online bird recording portals operating in Europe to describe the large-scale spatiotemporal patterns of bird distributions and their changes over time. Online bird portals obtain year-round data from the intensive and widespread activities of birdwatchers. However, data is, essentially, collected using simple standardized protocols (complete lists) or no protocol at all (casual observations). As the EBP collects data from the entire year, the project has the potential to obtain at least some of the metrics required for this EBV (e.g. arrival and departure timing) for several migratory bird species.

Possible bottlenecks in the generation of this EBV from existing data collected by the EBP initiative relate to:

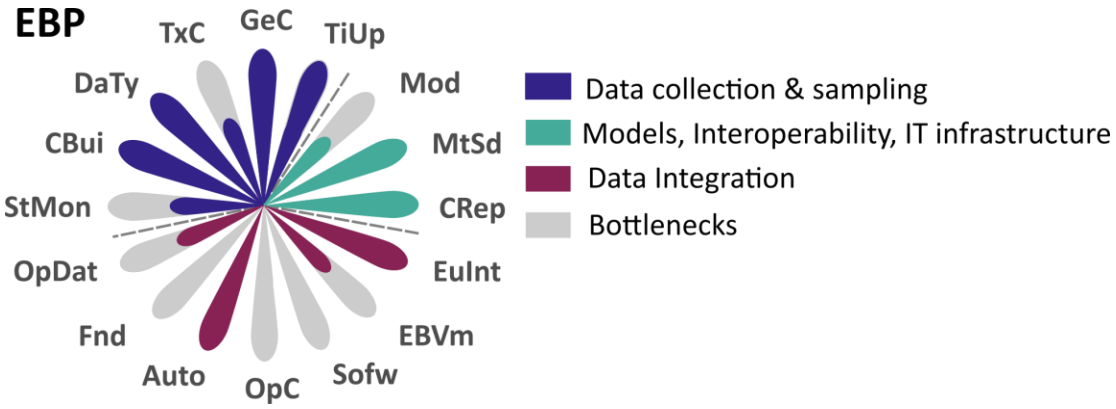
1) Data collection & sampling. The EBP has the potential to collect data on all bird species occurring in Europe but, currently, only the data for 137 species are stored in the EBP central data repository (**TxC** partial bottleneck). Data is collected across all European countries, but the degree of coverage is poorer in the South and, particularly, in the East and South-East. Data (species observations & counts) are gathered from online bird recording portals (e.g., ornitho) and include, essentially, both casual observations and data collected following simple systematic protocols (species lists). Only a small part of the data is collected following standardized monitoring protocols (**StMon** partial bottleneck). Training and capacity building for the network is provided during regular annual meetings or direct support to specific partners. Data is updated daily with information up to the previous day (species lists).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

2) Models, data interoperability and IT infrastructure. Despite the EBP/online bird portals data have been used to obtain some of the kind of metrics required to generate this EBV in some specific contexts (e.g. to predict the start of spring migration in the context of delimiting hunting seasons or study changes in migratory phenology; see, for example, Newton et al. 2016), the modelling of phenological traits has not been generalized nor automated within the framework of this initiative (**Mod** partial bottleneck). The EBP follows metadata standards to ensure data harmonization among its different data sources. The EBP data is harmonized, managed, and stored in a central repository curated by the European Bird Census Council.

3) Data Integration. The maps featured in the EBP viewer (www.eurobordportal.org) represent aggregated observations at the weekly basis and at 30 x 30 km resolution (**EBVm** partial bottleneck); however, in the EBP central data repository the casual data are aggregated at 10 x 10 km and date and the lists are in raw format (i.e. not aggregated) with date/timing and location given as precise location or at 10 x 10 km; therefore, the data available at the EBP repository would allow the generation of this EBV at the desired spatio-temporal resolution. Although this integration initiative can produce metrics of the kind required for this EBV, there is no open code nor user-friendly software yet to report in this regard as models are not routinely used to generate phenology products from EPB data yet (**Sofw** and **OpC** bottlenecks). Data streams are automated: the local online portals collect most of their data through mobile apps in near-real time or shortly after it has been recorded in the field; data collected by the data portals is then automatically transferred to the EBP daily. The EBP main developments have been possible thanks to the support of the LIFE programme (a new LIFE project proposal has been recently submitted), but funding stability in the mid- long- term is not warranted (**Fnd** bottleneck). EBP data is available upon request and subject to agreement by National coordinators who hold the ownership of data (**OpDat** partial bottleneck), but just one centralized data request should be done as data is already centralized in the EBP databank (authorizations by national owners are coordinated by EBP).



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EURING coordinates bird ringing in Europe and centralizes the data collected by the different bird ringing schemes operating in the continent to promote and encourage its use for bird management and conservation. Ringing data are collected year-round following standardized protocols (e.g., EUROCES Constant Effort Sites) or no protocol at all. The data consist of bird ringings (when a ring was first added to a bird), recaptures (retraps of ringed birds by ringers) and recoveries/resightings (ringed bird reported by the public (e.g., dead birds, rings/marks read at a distance) which form the bulk of the data currently stored in the EURING databank (EDB). The recently launched Eurasian African Migration Atlas (Spina et al. 2022) highlights the value of the EURING data to understand bird movements in time and space while the new Migration Mapping Tool is a good example of the applied value of combining the connectivity information of EURING with the observational data from the EuroBirdPortal in single tool (Gargallo et al. 2022).

The data collected by EURING and the bird ringing schemes in general have the potential to obtain the metrics required for this EBV and, more relevantly, are probably the best suited to estimate length of stay (e.g. though mark-recapture analysis).

Possible bottlenecks in the generation of this EBV from existing data collected by EURING relate to:

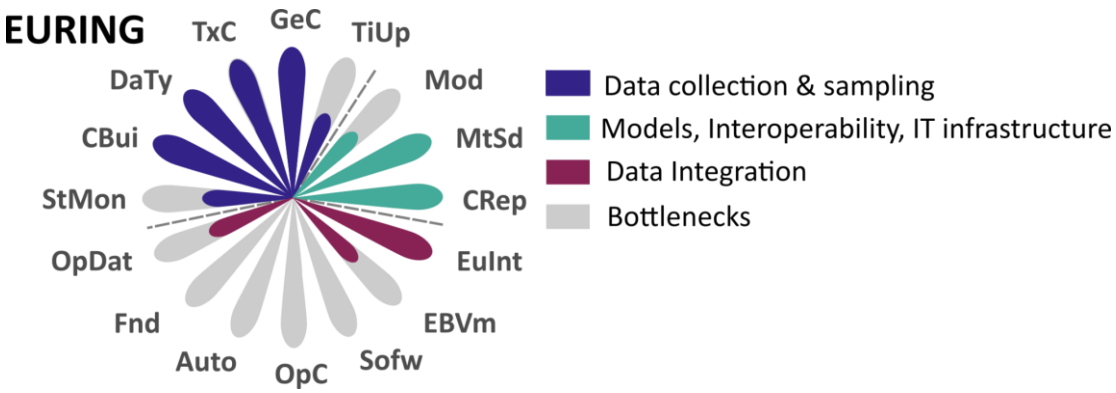
1) Data collection & sampling. EURING collects data of most bird species occurring in Europe and, though the amount of data (particularly recaptures/recoveries) varies greatly among species, figures are quite high for several of them. Data is collected across all European countries, but the degree of coverage is poorer in the South and, particularly, in the East and South-East. Only part of the data is collected following standardized monitoring protocols (**StMon** partial bottleneck). Training and capacity building for the network is provided during regular annual meetings or direct support to specific partners. Data in the EURINGdataback (EDB) are mostly updated once every year (**TiUp** partial bottleneck), although each record retains its temporal resolution (the date of data collection).

2) Models, data interoperability and IT infrastructure. Ringing data has been widely used to study phenology and length of stay and to obtain some of the metrics homologous to those required by this EBV for some specific uses, for example. to estimate the start of spring migration to delimit hunting seasons (e.g. the Migration seasons of hunted species research module of the Eurasian African Migration Atlas)(see also Ambrosini et al. 2014); however, so far, this has not been generalized nor automated within the framework of the initiative (**Mod** partial bottleneck). The EURING data flow takes place using the EURING Exchange Code standard, a pioneer of this kind among biodiversity data hubs in Europe, to ensure data harmonization and optimize its value. Data is centralized in the EURING databank.

3) Data Integration. EURING has already generated products that partially match the definition of this EBV (the Migration seasons of hunted species research module of the Eurasian African Migration Atlas, with predictions of prenuptial migration of different species at 10-day intervals - **EBVm** partial bottleneck). The data in the EDB have the required spatial and temporal resolution needed for the generation of this EBV. However, as the data is mostly updated once every year and therefore, this EBV could only be generated, at most, on a yearly basis. Though this integration initiative can produce metrics of the kind required for this EBV, there is no open code nor user-friendly software yet to report in this regard, as this modelling is not routinely integrated in this initiative (**Sofw** and **OpC** bottlenecks). Data streams are not

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

automated yet: ringing centers across Europe send their ringing observations (data exports) once a year to the central repository on a manual basis (**Auto** bottleneck). Funding stability in the mid- long- term is not warranted (**Fnd** bottleneck). EDB data is available upon request and subject to agreement by National bird ringing schemes who hold the ownership of data (**OpDat** partial bottleneck), but just one centralized data request should be done as data is already centralized in the EDB (authorizations by national owners are coordinated by EBP).



References:

Ambrosini, R., Borgoni, R., Rubolini, ... & Saino, N. (2014). Modelling the progression of bird migration with conditional autoregressive models applied to ringing data. *PloS one*, 9(7), e102440.

Gargallo, G., Davies, J.G., Faverjon, C., Kampichler C., Baillie, S.R., Cameron, A., Robinson, R.A. and Sierdsema, H. (2022) Development of a prototype early warning system for avian influenza in the EU based on risk-mapping. EFSA supporting publication 2022: EN-7762, 46pp. URL.

Newson, S. E., Moran, N. J., Musgrove, A. J., Pearce-Higgins, J. W., Gillings, S., Atkinson, P. W., ... & Baillie, S. R. (2016). Long-term change in spring and autumn migration phenology of common migrant breeding birds in Britain: results from large-scale citizen science bird recording schemes. *Ibis (Lond. 1859)*, *158*, 481-495.

Spina, F., Baillie, S.R., Bairlein, F, Fiedler, W. and Thorup, K. (Eds) 2022. *The Eurasian African Bird Migration Atlas*. https://migrationatlas.org. EURING/CMS.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Freshwater community composition EBVs

EBV: Ecological Quality Ratio (EQR) of phytoplankton in lakes	
ID	9
Realm	Freshwater
EBV class	Community composition
EBV name	Ecological Quality Ratio (EQR) of phytoplankton in lakes
Step in identification process	User & Policy Needs Assessment
Definition	Community composition and total biomass of phytoplankton in lakes (Ecological Quality Ratio) based on total abundance (biovolume), taxonomic composition index across all species based on biovolume per indicator species, and bloom intensity, e.g., maximum biomass of cyanobacteria or percentage of cyanobacteria of the total biomass for all taxa.
Metric	The Ecological Quality Ratio (EQR) of phytoplankton in European lakes, expressed as a numerical value between zero (bad) and one (high), quantifying the ecological status of phytoplankton community composition and its deviation from a reference condition. The metric describes the deviation from natural phytoplankton communities.
Spatial resolution unit	Lakes as delineated in ECRINS (European catchments and rivers network system)
Temporal resolution unit	1 year, weekly-monthly during growing season
Taxonomic/ ecosystem focus group	Phytoplankton iIndicator taxa and reference taxa as described in the Water Framework Directive Intercalibration Technical Reports (Part 2, Lakes)

Main European initiatives and description of current bottlenecks

The **WISE-2** (Water Information System for Europe - Biology data), is an ongoing European monitoring data collection scheme established by the European Environment Agency (EEA) to obtain a harmonized flow of biology monitoring data from all EEA member states reported as Ecological Quality Ratios (EQRs) from all surface water categories: rivers, lakes, transitional and coastal waters. This data flow is tightly linked to the reporting duties of Member States within the regulations of the Water Framework Directive.

Possible bottlenecks in the generation of this EBV from existing data collected by WISE-2 data relate to:

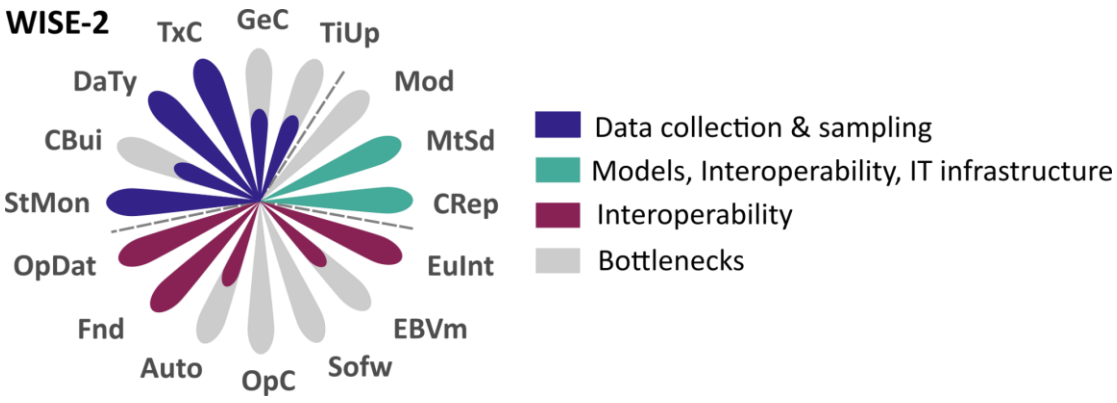
1) Data collection and sampling. Currently, only 13 European countries report phytoplankton in lakes to the WISE-2 data flows so there is a geographic coverage gap; however, phytoplankton species are quite generalist and widespread so probably most if not all taxa within this group are covered by the monitoring of those 13 countries (**Gec** partial bottleneck). However, 26 countries monitor phytoplankton in lakes on an annual basis for the Water Framework Directive flows, so probably the geographic gaps are smaller than expected by WISE-2 data flows. The WISE-2 data flow, as an integration initiative, organizes yearly webinars for all member states with information and Q&A sessions on data reporting, but does not have capacity building on monitoring. If this exists, it is the capacity building of the member states (**CBui** bottleneck); this EBV requires highly specialized taxonomic expertise, and capacity building is indeed done by the countries but not all countries have the same capacity. Data collected for this taxonomic group is composition, abundance and biomass. Monitoring is carried out every year in some water bodies,

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

but not in all, but EQRs are reported to the WISE system on an annual basis in all countries for water bodies monitored two or more years earlier than the year of reporting (there is not weekly -monthly reporting during the growing season; **TiUp** partial bottleneck).

2) Models, interoperability, and IT infrastructure. The use of models does not apply in EQR EBVs (**Mod** bottleneck): EQRs are estimated as a ratio between the value of the observed biological parameter for a given type of surface water body and the expected value under reference conditions. The ratio shall be expressed as a numerical value between 0 and 1, with high ecological status represented by values close to one (EQR > 0.8) and bad ecological status by values close to zero (EQR< 0.21). This ecological status value is the one that gets reported to the EEA. The EEA requires the countries to report their data to the CDR - Central Data Repository. The EQRs reported by member states are aggregated to various levels by EEA (ETC) to obtain overall EU values or by regions or groups of water bodies within the same status class or classes, e.g., WBs with EQR values corresponding to high and good status (EQR > 0.6), moderate status (EQR between 0.4 and 0.59) or poor+bad status (EQR > 0.21). Reporting by Member States follows metadata standards (the structure and contents of the data are described in a Data Dictionary).

3) Data integration. The phytoplankton EQRs reported under WISE-2 data flows partially match the definition of the EBV (although not in terms of temporal resolution as weekly-monthly data are not available; **EBVm** partial bottleneck). As there are not currently models in place to predict phytoplankton EQRs across European lakes, there is no open code or user-friendly software already being used to generate this EBV from WISE-2 data (**OpC** and **SoFw** bottlenecks). Data flows are not fully automated: the EEA provides empty templates for download from the data dictionary as Excel file and XML Schema. There is a reporting tool (ReportNet) where member states can upload these data. An automatic quality control (QR) process is implemented on the Central Data Repository to check the structure and content of the data file(s) uploaded by each member state. This first step of QC is followed by three more QC steps that involve: (2) cross-checking against previously reported data, (3) statistical tools and other algorithms for identifying suspicious values, and (4) expert-based checking and communication with data reporters. Automation of data flows between field data collection and national coordinators yet to be explored (some tests in Norway are underway) (**Auto** partial bottleneck). Funding has to be secured by member states to comply with reporting obligations. While phytoplankton EQR ratios at the waterbody level are openly available.



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Ecological Quality Ratio (EQR) of freshwater macrophytes	
ID	10
Realm	Freshwater
EBV class	Community composition
EBV name	Ecological Quality Ratio (EQR) of freshwater macrophytes
Step in identification process	User & Policy Needs Assessment
Definition	Community composition of macrophytes (Ecological Quality Ratio) based on presence-absence data.
Metric	The Ecological Quality Ratio (EQR) of macrophytes in European lakes, expressed as a numerical value between zero (bad) and one (very good), quantifying the ecological status of macrophyte community composition and its deviation from a reference condition.
Spatial resolution unit	Lakes as delineated in ECRINS (European catchments and rivers network system)
Temporal resolution unit	6 years
Taxonomic/ ecosystem focus group	All macrophytes species with indicator values as defined in the Water Framework Directive Intercalibration Technical Reports (Part 2, Lakes)

Main European initiatives and description of current bottlenecks

The **WISE-2** (Water Information System for Europe - Biology data), is an ongoing European monitoring data collection scheme established by the European Environment Agency (EEA) to obtain a harmonized flow of biology monitoring data from all EEA member states reported as Ecological Quality Ratios (EQRs) from all surface water categories: rivers, lakes, transitional and coastal waters. This data flow is tightly linked to the reporting duties of Member States within the regulations of the Water Framework Directive.

Possible bottlenecks in the generation of this EBV from existing data collected by WISE-2 data relate to:

1) Data collection and sampling. Currently, only 15 European countries report Macrophytes in lakes to the WISE-2 data flows so there is a geographic coverage gap that probably translates into a taxonomic coverage gap (some countries only report EQRs for this taxonomic group intermittently) (**Gec**, **TxC** partial bottlenecks). However, 19 countries monitor macrophytes in lakes on an annual basis for the Water Framework Directive flows, so probably the geographic gaps are smaller than expected by WISE-2 data flows. The WISE-2 data flow, as an integration initiative, organizes yearly webinars for all member states with information and Q&A sessions on data reporting, but does not have capacity building on monitoring. If this exists, it is the capacity building of the member states (CBui bottleneck). Data collected for this taxonomic group is mostly presence- absence with some measures of relative abundance. Monitoring is carried out every year in some water bodies, but not in all, but EQRs are reported to the WISE system on an annual basis in all countries for water bodies monitored two or more years earlier than the year of reporting.

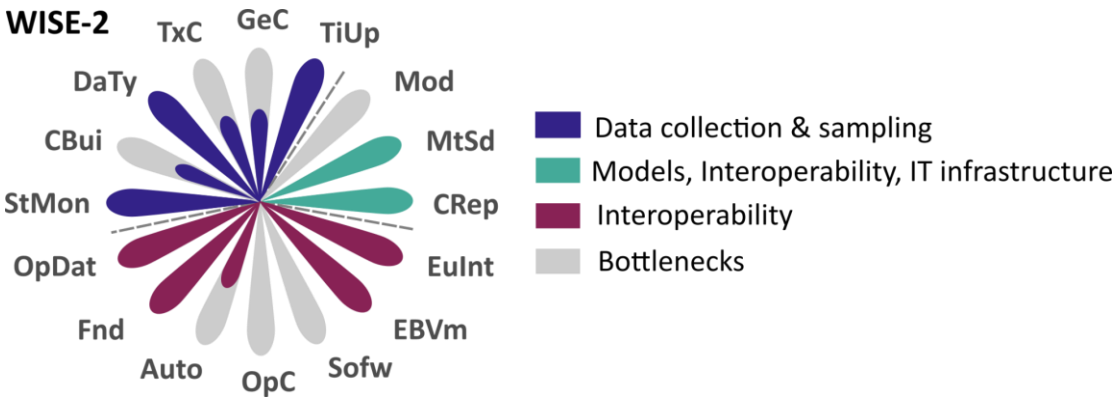
2) Models, interoperability, and IT infrastructure. The use of models does not apply in EQR EBVs (**Mod** bottleneck): EQRs are estimated as a ratio between the value of the observed biological parameter for a given type of surface water body and the expected value under reference conditions. The ratio shall be expressed as a numerical value between 0 and 1, with high ecological status represented by values

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023.* DOI: https://doi.org/10.3897/arphapreprints.e103765

close to one (EQR > 0.8) and bad ecological status by values close to zero (EQR< 0.21). This ecological status value is the one that gets reported to the EEA. The EEA requires the countries to report their data to the CDR - Central Data Repository. The EQRs reported by member states are aggregated to various levels by EEA (ETC) to obtain overall EU values or by regions or groups of water bodies within the same status class or classes, e.g., WBs with EQR values corresponding to high and good status (EQR > 0.6), moderate status (EQR between 0.4 and 0.59) or poor+bad status (EQR > 0.21). Reporting by Member States follows metadata standards (the structure and contents of the data are described in a Data Dictionary).

3) Data integration. The Macrophyte EQRs reported under WISE-2 data flows perfectly match the definition of the EBV. As there are not currently models in place to predict Macrophyte EQRs across European lakes, there is no open code or user-friendly software already being used to generate this EBV from WISE-2 data (**OpC** and **SoFw** bottlenecks). Data flows are not fully automated: the EEA provides empty templates for download from the data dictionary as Excel file and XML Schema. There is a reporting tool (ReportNet) where Member States can upload these data. An automatic quality control (QC) process is implemented on the Central Data Repository, to check the structure and content of the data file(s) uploaded by each member state. This first step of QC is followed by three more QC steps that involve: (2) cross-checking against previously reported data, (3) statistical tools and other algorithms for identifying suspicious values, and (4) expert-based checking and communication with data reporters. Automation of data flows between field data collection and national coordinators yet to be explored (some tests in Norway are underway) (**Auto** partial bottleneck). Funding must be secured by member states to comply with reporting obligations. While Macrophytes EQR ratios at the waterbody level are openly available, but raw data at the site level t (composition and abundance values) are not.



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Ecological Quality Ratio (EQR) of freshwater phytobenthos	
ID	11
Realm	Freshwater
EBV class	Community composition
EBV name	Ecological Quality Ratio (EQR) of freshwater phytobenthos
Step in identification process	User & Policy Needs Assessment
Definition	The ecological status of phytobenthos in European rivers, measured as Ecological Quality Ratio (EQR).
Metric	The Ecological Quality Ratio (EQR) of phytobenthos in European rivers, expressed as a numerical value between zero (bad) and one (very good), quantifying the ecological status of phytobenthos community composition and its deviation from a reference condition.
Spatial resolution unit	River catchments as delineated in ECRINS (European catchments and Rivers network system)
Temporal resolution unit	1-3 years
Taxonomic/ ecosystem focus group	Phytobenthic species with indicator values as defined in the Water Framework Directive Intercalibration Technical Reports (Part 1, Rivers)

Main European initiatives and description of current bottlenecks

The **WISE-2** (Water Information System for Europe - Biology data), is an ongoing European monitoring data collection scheme established by the European Environment Agency (EEA) to obtain a harmonized flow of biology monitoring data from all EEA member states reported as Ecological Quality Ratios (EQRs) from all surface water categories: rivers, lakes, transitional and coastal waters. This data flow is tightly linked to the reporting duties of Member States within the regulations of the Water Framework Directive.

Possible bottlenecks in the generation of this EBV from existing data collected by WISE-2 data relate to:

1) Data collection and sampling. Currently, only 22 European countries report phytobenthos in rivers to the WISE-2 data flows so there is a geographic coverage gap that probably translates into a taxonomic coverage gap (some countries only report EQRs for this taxonomic group intermittently) (**GeC**, **TxC partial** bottlenecks). The WISE-2 data flow, as an integration initiative, organises yearly webinars for all member states with information and Q&A sessions on data reporting, but does not have capacity building on monitoring. If this exists, it is the capacity building of the member states (**CBui** bottleneck); Data collected for this taxonomic group is mostly presence- absence with some measures of relative abundance. Monitoring is carried out every year in some water bodies, but not in all, but EQRs are reported to the WISE system on an annual basis in all countries for water bodies monitored two or more years earlier than the year of reporting.

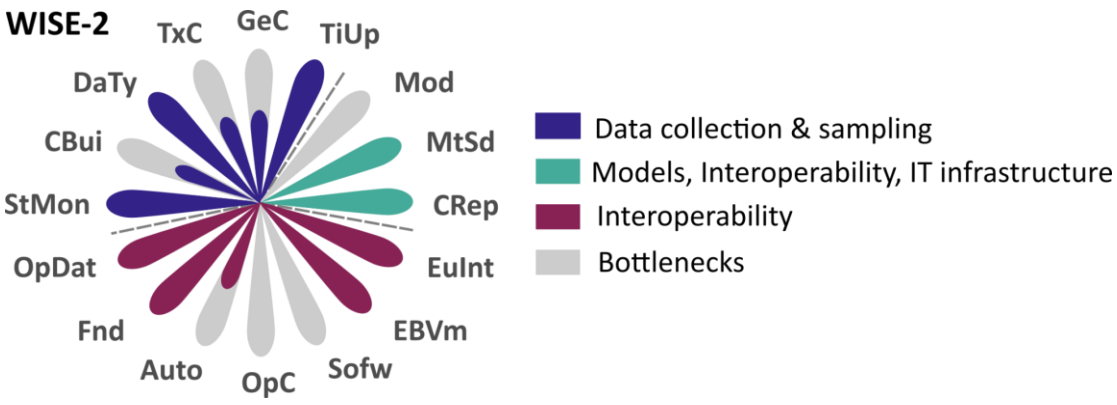
2) Models, interoperability, and IT infrastructure. The use of models does not apply in EQR EBVs (**Mod** bottleneck): EQRs are estimated as a ratio between the value of the observed biological parameter for a given type of surface water body and the expected value under reference conditions. The ratio shall be expressed as a numerical value between 0 and 1, with high ecological status represented by values close to one (EQR > 0.8) and bad ecological status by values close to zero (EQR< 0.21). This ecological status value is the one that gets reported to the EEA. The EEA

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

requires the countries to report their data to the CDR - Central Data Repository. The EQRs reported by member states are aggregated to various levels by EEA (ETC) to obtain overall EU values or by regions or groups of water bodies within the same status class or classes, e.g., WBs with EQR values corresponding to high and good status (EQR > 0.6), moderate status (EQR between 0.4 and 0.59) or poor+bad status (EQR > 0.21). Reporting by Member States follows metadata standards (the structure and contents of the data are described in a Data Dictionary).

3) Data integration. The phytobenthos EQRs reported under WISE-2 data flows perfectly match the definition of the EBV. As there are not currently models in place to predict phytobentos EQRs across European rivers there is no open code or user-friendly software already being used to generate this EBV from WISE-2 data (**OpC** and **SoFw** bottlenecks). Data flows are not fully automated: the EEA provides empty templates for download from the data dictionary as Excel file and XML Schema.There is a reporting tool (ReportNet) where to upload these data.
An automatic quality control (QC) process is implemented on the Central Data Repository, to check the structure and content of the data file(s) uploaded by each member state. This first step of QC is followed by three more QC steps that involve: (2) cross-checking against previously reported data, (3) statistical tools and other algorithms for identifying suspicious values, and (4) expert-based checking and communication with data reporters. Automation of data flows between field data collection and national coordinators yet to be explored (some tests in Norway are underway) (**Auto** partial bottleneck). Funding has to be secured by member states to comply with reporting obligations. Phytobentos EQR ratios at the waterbody level are openly available; however raw data (composition and abundance values are not.



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Ecological Quality Ratio (EQR) of benthic freshwater invertebrates	
ID	12
Realm	Freshwater
EBV class	Community composition
EBV name	Ecological Quality Ratio (EQR) of benthic freshwater invertebrates
Step in identification process	User & Policy Needs Assessment
Definition	The ecological status of benthic invertebrates in European rivers, measured as Ecological Quality Ratio (EQR).
Metric	The Ecological Quality Ratio (EQR) of benthic invertebrates in European rivers, expressed as a numerical value between zero (bad) and one (very good), quantifying the ecological status of benthic invertebrates community composition and its deviation from a reference condition.
Spatial resolution unit	River catchments as delineated in ECRINS (European catchments and rivers network system)
Temporal resolution unit	2-3 years
Taxonomic/ ecosystem focus group	Benthic invertebrate species with indicator values as defined in the Water Framework Directive Intercalibration Technical Reports (Part 1, Rivers)

Main European initiatives and description of current bottlenecks

The **WISE-2** (Water Information System for Europe - Biology data), is an ongoing European monitoring data collection scheme established by the European Environment Agency (EEA) to obtain a harmonized flow of biology monitoring data from all EEA member states reported as Ecological Quality Ratios (EQRs) from all surface water categories: rivers, lakes, transitional and coastal waters. This data flow is tightly linked to the reporting duties of Member States within the regulations of the Water Framework Directive.

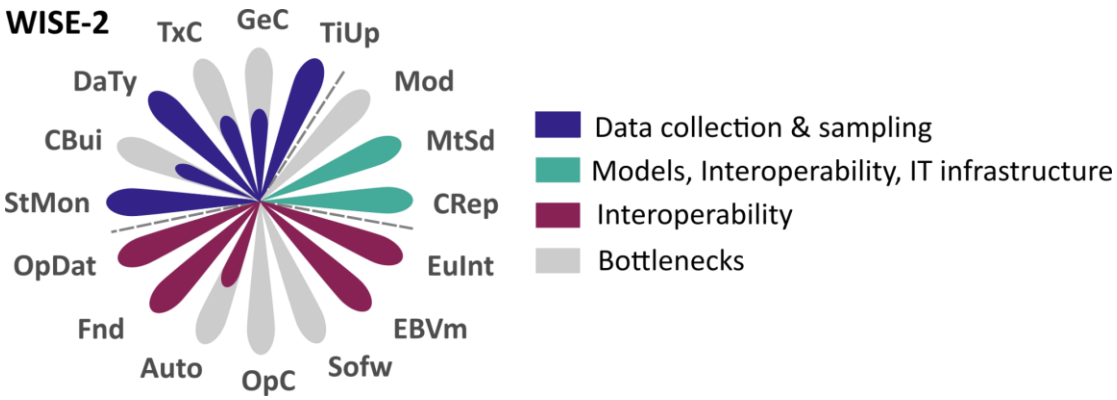
Possible bottlenecks in the generation of this EBV from existing data collected by WISE-2 data relate to:

1) Data collection and sampling. Currently, only 25 European countries report benthic invertebrates in rivers to the WISE-2 data flows so there is a geographic coverage gap that probably translates into a taxonomic coverage gap (some countries only report EQRs for this taxonomic group intermittently) (**Gec**, **TxC** partial bottlenecks). However, 28 countries monitor benthic invertebrates in rivers on an annual basis for the Water Framework Directive flows, so probably the geographic gaps are smaller than expected by WISE-2 data flows. The WISE-2 data flow, as an integration initiative, organizes yearly webinars for all member states with information and Q&A sessions on data reporting, but does not have capacity building on monitoring. If this exists, it is the capacity building of the member states (**CBui** partial bottleneck); Data collected for this taxonomic group is composition and abundance. Monitoring is carried out every year in some water bodies, but not in all, but EQRs are reported to the WISE system on an annual basis in all countries for water bodies monitored two or more years earlier than the year of reporting.
2) Models, interoperability, and IT infrastructure. Currently, the use of models does not apply in EQR EBVs (**Mod** bottleneck): EQRs are estimated as a ratio between the value of the observed biological parameter for a given type of surface water body and the expected value under reference conditions. The ratio shall be expressed as

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

a numerical value between 0 and 1, with high ecological status represented by values close to one (EQR > 0.8) and bad ecological status by values close to zero (EQR< 0.21). This ecological status value is the one that gets reported to the EEA. The EEA requires the countries to report their data to the CDR - Central Data Repository. The EQRs reported by member states are aggregated to various levels by EEA (ETC) to obtain overall EU values or by regions or groups of water bodies within the same status class or classes, e.g., WBs with EQR values corresponding to high and good status (EQR > 0.6), moderate status (EQR between 0.4 and 0.59) or poor+bad status (EQR > 0.21). However, for these taxa, there are some modelling initiatives underway to extrapolate EQR values from monitoring sites to non-monitored water bodies, using as reference the EQR values reported to WISE-2 data flows and considering predictors such as land use and abiotic variables (e.g., nutrient concentrations). Reporting by Member States follows metadata standards (the structure and contents of the data are described in a Data Dictionary).

3) Data integration. The benthic invertebrate EQRs reported under WISE-2 data flows perfectly match the definition of the EBV. As there are not currently models in place to predict freshwater invertebrates EQR values across European rivers, there is no open code or user-friendly software already being used to generate this EBV from WISE-2 data (**OpC** and **SoFw** bottlenecks). Data flows are not fully automated: the EEA provides empty templates for download from the data dictionary as Excel file and XML Schema. There is a reporting tool (ReportNet) where Member States can upload these data. An automatic quality control (QC) process is implemented on the Central Data Repository, to check the structure and content of the data file(s) uploaded by each member state. This first step of QC is followed by three more QC steps that involve: (2) cross-checking against previously reported data, (3) statistical tools and other algorithms for identifying suspicious values, and (4) expert-based checking and communication with data reporters. Automation of data flows between field data collection and national coordinators yet to be explored (some tests in Norway are underway) (**Auto** partial bottleneck). Funding must be secured by member states to comply with reporting obligations. Benthic invertebrate EQR ratios at the waterbody level are openly available, but raw data are not (composition and abundance values).



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Ecological Quality Ratio (EQR) of freshwater fish	
ID	13
Realm	Freshwater
EBV class	Community composition
EBV name	Ecological Quality Ratio (EQR) of freshwater fish
Step in identification process	User & Policy Needs Assessment
Definition	The ecological status of fish in European freshwater systems (lakes and rivers), measured as Ecological Quality Ratio (EQR).
Metric	The Ecological Quality Ratio (EQR) of fish in European lakes and rivers, expressed as a numerical value between zero (bad) and one (very good), quantifying the ecological status of fish community composition and its deviation from a reference condition.
Spatial resolution unit	Lakes and rivers catchments as delineated in ECRINS (European catchments and rivers network system)
Temporal resolution unit	3 – 6 years
Taxonomic/ ecosystem focus group	Freshwater fish species with indicator values defined in the Water Framework Directive Intercalibration Technical Reports (Part 1, Rivers; Part2, lakes)

Main European initiatives and description of current bottlenecks

The **WISE-2** (Water Information System for Europe), is an ongoing European monitoring scheme established by the European Environment Agency (EEA) to obtain a harmonized flow of biology data from all EEA member states reported as Ecological Quality Ratios (EQRs) from all surface water categories: rivers, lakes, transitional and coastal waters. This data flow is tightly linked to the reporting duties of Member States within the regulations of the Water Framework Directive.

Possible bottlenecks in the generation of this EBV from existing data collected by WISE-2 data relate to:

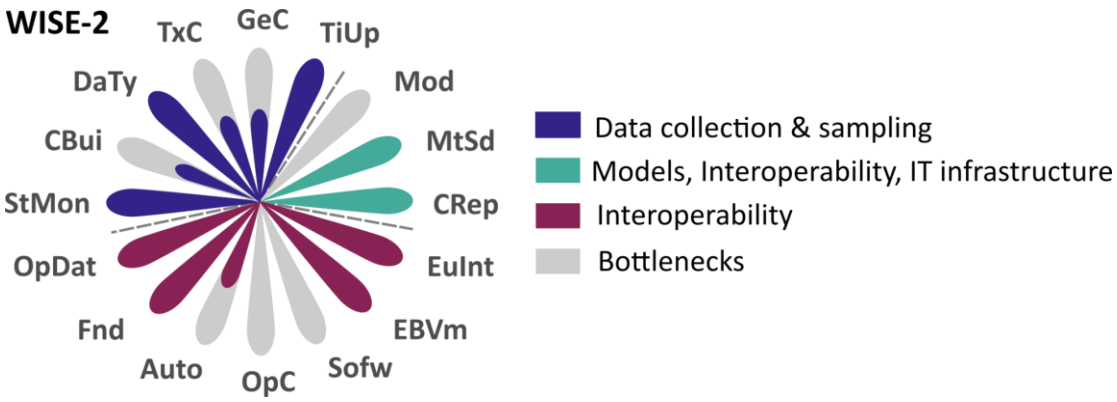
1) Data collection and sampling. Currently, only 10 European countries report FISH_EQRs to the WISE-2 so there is a gap in the reporting of fish data to the European Commission, that probably translates into a gap in taxonomic coverage given many fish species are endemic of specific river catchments (**GeC, TxC** partial bottlenecks). However, 27 countries collect data on an annual basis on fish species for the Water Framework Directive flows, so probably the geographic and taxonomic monitoring gaps are smaller than expected by WISE-2 data flows and Fish EQRs could be reported by a larger number of countries. The WISE-2 data flow, as an integration initiative, organises yearly webinars for all member states with information and Q&A sessions on data reporting, but does not have capacity building on monitoring. If this exists, it is the capacity building of the member states (**CBui** partial bottleneck). Data collected for this taxonomic group is composition, abundance and age structure. Monitoring is carried out every year in some water bodies, but not in all, but EQRs are reported to the WISE system on an annual basis in all countries for water bodies monitored two or more years earlier than the year of reporting.

2) Models, interoperability, and IT infrastructure. The use of models does not apply in EQR EBVs (**Mod** bottleneck): EQRs are estimated as a ratio between the value of the observed biological parameter for a given type of surface water body and the

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

expected value under reference conditions. The ratio shall be expressed as a numerical value between 0 and 1, with high ecological status represented by values close to one (EQR > 0.8) and bad ecological status by values close to zero (EQR < 0.21). This ecological status value is the one that gets reported to the EEA. The EEA requires the countries to report their data to the CDR - Central Data Repository. The EQRs reported by member states are aggregated to various levels by EEA (ETC) to obtain overall EU values or by regions or groups of water bodies within the same status class or classes, e.g., WBs with EQR values corresponding to high and good status (EQR > 0.6), moderate status (EQR between 0.4 and 0.59) or poor+bad status (EQR > 0.21). Reporting by Member States follows metadata standards (the structure and contents of the data are described in a Data Dictionary).

3) Data integration. The fish EQRs reported under WISE-2 data flows perfectly match the definition of the EBV. As there are not currently models in place to predict fish EQR values across European lakes and rivers, there is no open code or user-friendly software already being used to generate this EBV from WISE-2 data (**OpC** and **SoFw** bottlenecks). Data flows are not fully automated: the EEA provides empty templates for download from the data dictionary as Excel file and XML Schema. There is a reporting tool (ReportNet) where to upload these data. An automatic quality control (QC) process is implemented on the Central Data
Repository, to check the structure and content of the data file(s) uploaded by each member state. This first step of QC is followed by three more QC steps that involve: (2) cross-checking against previously reported data, (3) statistical tools and other algorithms for identifying suspicious values, and (4) expert-based checking and communication with data reporters. Automation of data flows between field data collection and national coordinators yet to be explored (some tests in Norway are underway) (**Auto** partial bottleneck). Funding must be secured by member states to comply with reporting obligations. FISH_EQR ratios at the waterbody level are openly available, but raw data are not (composition, abundance and age structure measures).



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Freshwater ecosystem function EBVs

EBV: Harmful freshwater algal blooms	
ID	19
Realm	Freshwater
EBV class	Ecosystem function
EBV name	Harmful freshwater algal blooms
Step in identification process	User & Policy Needs Assessment
Definition	Distribution, intensity, frequency and position of harmful algal blooms in European lakes which occur when cyanobacteria accumulate in water, with the potential to harm the health of humans, plants, and animals.
Metric	- Observed location and intensity of algal blooms as derived from satellite imagery (e.g. Copernicus Sentinel-3) or regulatory monitoring - Modeled cyanobacterial density based on hydrodynamic models and satellite imagery
Spatial resolution unit	Lakes as delineated in ECRINS (European catchments and Rrivers network system)
Temporal resolution unit	Real-time, weekly-monthly during the growing season
Taxonomic/ ecosystem focus group	ECRINS lakes

Main European initiatives and description of current bottlenecks

Regulatory monitoring of algal blooms (Cyanobacteria) is made as part of the Water Framework Directive Lake phytoplankton monitoring to assess phytoplankton status class. Cyanobacteria biovolume (in mg L-1 or mm3 L-1) or the % of total biovolume in lakes are reported to the Water Information System for Europe - Water Quality (WISE-6) - an ongoing European monitoring scheme established by the European Environment Agency (EEA) to obtain a harmonized flow of water quality data in groundwater, rivers, lakes, transitional, coastal and marine waters. Includes data on nutrients, organic matter, chlorophyll-a, hazardous substances and general physico-chemical parameters in water, sediment and biota. Observed presence and intensity of algal blooms derived from satellite imagery (e.g. Copernicus Sentinel-3) is also possible, but public services are still under-development (this is an area "under development" for GEO AQuaWatch and the Copernicus Global Land Service (CGLS) for lake water quality Lake Water Quality | Copernicus Global Land Service.

Possible bottlenecks in the generation of this EBV from regulatory monitoring collected by WISE-6 data relate to:

1) Data collection and sampling. Currently, while 26 countries monitor phytoplankton in lakes to report ecological status to the Water Framework Directly, only 5 European countries report cyanobacteria to the WISE-6 data flows, so there is a geographic coverage gap in WISE-6 data flows for these taxa (**GeC** partial bottlenecks). The WISE-6 data flow, as an integration initiative, organizes yearly webinars for all member states with information and Q&A sessions on data reporting, but does not have capacity building on monitoring. If this exists, it is the capacity building of the member states (**CBui** partial bottleneck). Data collected for

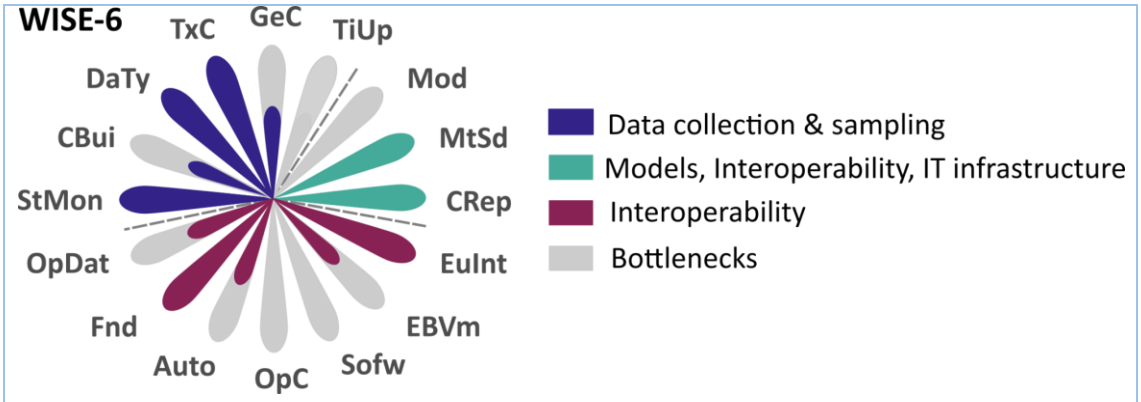
This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

this taxonomic group is Cyanobacteria biovolume (in mg L-1 or mm3 L-1) that could potentially be used for modelling purposes. Monitoring is carried out every year in some water bodies, but not in all, but Cyanobacteria biovolume is reported to the WISE system on an annual basis. Regulatory monitoring for the Water Framework Directive is currently available monthly for summer months (July to September) for one or two years every six years. However, the sample frequency is not enough to generate this EBV in real time (**TiUp** bottleneck). Near real-time or weekly monitoring may become available in future years through satellite EO (CGLS) and citizen science.

2) Models, interoperability, and IT infrastructure. Modeled cyanobacteria density in European lakes based on lake type, climate and water quality (total phosphorus) data is possible using published statistical models (Richardson et al., 2018) but it is still not available through WISE-6 data flows (**Mod** bottleneck). Reporting by Member States follows metadata standards (the structure and contents of the data are described in a Data Dictionary) and stored in the WISE data repository.

3) Data integration. The data collected and reported via WISE-6 does not perfectly match the EBV description, especially in terms of temporal resolution (**EBVm** partial bottleneck). Since models are not currently in place to predict the cyanobacterial density at WISE-6, there is no open code or user-friendly software to report here (**OpC** and **SoFw** bottlenecks). Data flows are not fully automated: the EEA provides empty templates for download from the data dictionary as Excel file and XML Schema. There is a reporting tool (ReportNet) where Member States can upload these data. An automatic quality control (QC) process is implemented on the Central Data Repository, to check the structure and content of the data file(s) uploaded by each member state. This first step of QC is followed by three more QC steps that involve: (2) cross-checking against previously reported data, (3) statistical tools and other algorithms for identifying suspicious values, and (4) expert-based checking and communication with data reporters. Automation of data flows between field data collection and national coordinators is yet to be explored (some tests in Norway are underway); moreover, citizen science data are available for this taxa in 5 countries in Europe (Belgium, Ireland, Netherlands, Norway and UK) using the Bloomin' Algae app. Data provide high spatial resolution (10-20 m) and potentially high frequency (daily data), but data are qualitative (presence/absence of blooms judged by experts on photographic evidence) and do not provide quantitative information on biovolume or % abundance. The data are viewable and raw data (location, date) can be made available real-time through an API from iRecord (**Auto** partial bottleneck). Funding must be secured by member states to comply with reporting obligations. Cyanobacteria biovolume is available upon request for all the countries reporting for WISE-6 (**OpDat** partial bottleneck). Cyanobacteria is not reported as a separate metric to the WFD but is part of the metrics used to assess lake phytoplankton status class.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.



References:

Richardson, J., Miller, C., Maberly, S. C., Taylor, P., Globevnik, L., Hunter, P., ... and Carvalho, L. (2018). Effects of multiple stressors on cyanobacteria abundance vary with lake type. *Global Change Biology,* 24(11), 5044-5055.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ANNEX II: Marine EBV fact sheets

Marine species populations

EBV: Species Distributions of Marine Fishes	
ID	23
Realm	Marine
EBV class	Species Populations
EBV name	Species distributions of marine fishes
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of European Marine fish species in EU's marine waters within contiguous spatial units (grid cells) over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	50 × 50 km – 200 × 200 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Marine fish species indicated in the Review and Analyses of Member States 2018 reports Descriptor 1: Species biological diversity
Main European initiatives and description of current bottlenecks	

There is not a single monitoring network integrating data of marine fishes across the whole of Europe. However, there are different regional integration initiatives collecting data on presence and abundance of marine fishes that could potentially provide data to generate this EBV. Among them, we highlight here the fish data gathered by the International Council for the Exploitation of the Sea (ICES), corresponding to surveys undertaken through trawls (main source of fish monitoring data in Europe - Jessop et al. 2022). While these are mostly directed to commercial species (see fact sheet of EBV 24 - *Species abundances of marine commercial fish species and long-distance migratory fishes*), data from other species gets also recorded and individual species get recorded via a unique FishID code.

Possible bottlenecks in the generation of this EBV from ICES data relate to:

1) Data collection and sampling. The International Council for the Exploitation of the Sea (ICES) holds a database on pelagic and bottom fish trawl surveys collected in the North-east Atlantic and Baltic seas (DATRAS online database). Survey data comes from standardized survey protocols but these differ across regions (e.g. between the Baltic and the Western and Southern Areas) (**StMon** partial bottleneck). Data reported to DATRAS includes the coordinates of the shooting and hauling locations, a species ID and information on the age disaggregated abundance of fish species that could potentially be used to generate this EBV. The survey data currently covers the Baltic Sea, Skagerrak, Kattegat, North Sea, English Channel, Celtic Sea, Irish Sea, Bay of Biscay and the eastern Atlantic from the Shetlands to Gibraltar; therefore, the geographic coverage of data is not complete as it does not include data from the Mediterranean sea (**GeC** partial bottleneck); this translates in the lack of representation of Mediterranean species in the database (**TxC** partial bottleneck). There are more than 45 years of continuous time series data in DATRAS, and survey

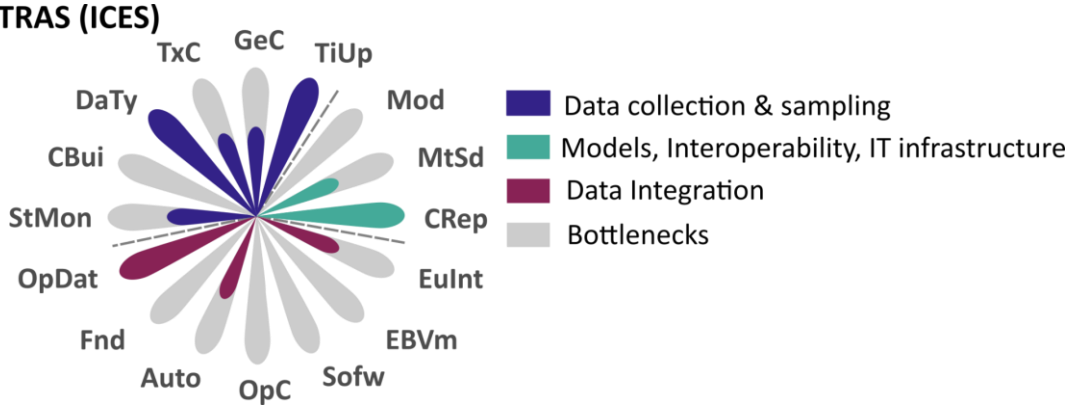
This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

data are continuously updated by national institutions, so the data availability matches the temporal resolution sought by this EBV.

2) Models, interoperability, and IT infrastructure. ICES does not use models to generate spatially explicit distribution maps of fish species. Data are used to estimate different species abundance indices (**Mod** bottleneck). The DATRAS dataset has defined a vocabulary and a set of standards for data reporting to ease data handling, integration and availability but there are not common metadata standards followed across all surveys (**MtSd** partial bottleneck). Data is centralized and stored in the ICES data portal.

3) Data integration. There is not a single integration initiative that collects and processes data of marine fishes' abundance and distributions across Europe (only regional and national ones, being the Mediterranean Sea clearly underrepresented) (**EuInt** partial bottleneck). The products generated from DATRAS (i.e., indices of fish stock abundance estimates of commercial species) do not match this EBV (**EBVm** bottleneck). Since there are not currently models in place to map the distribution of marine fish species, there is no open code or user-friendly software to report here (**OpC** and **SoFw** bottlenecks). Data streams are not fully automated: worksheets in .csv format are submitted to the DATRAS via the ICES platform but there is little information about the level of automatization of the data flows of the underlying datasets (**Auto** partial bottleneck). If funding for adequate survey effort delivering the data exists, it is that of the countries reporting data to DATRAS and ICES (**Fnd** bottleneck). DATRAS data products (such as Catch per area of fishing effort estimates - CPUE- or indices) and raw data, can be freely downloaded according to the ICES Data policy.

DATRAS (ICES)



See Jessop et al. (2022) for a summary of other national and regional fish monitoring efforts available across Europe.

References:

Jessop, A., Chow, C., Dornelas, M., Pereira, P., Sousa-Pinto, I., Hernández Chan, S., Junker, J., Soares, J., Ratnarajah, L., Fernández, M., Mendo, T. (2022) MarBioME. Overview and assessment of the current state of Biodiversity Monitoring in the European Union and adjacent marine waters. European Commission. Directorate General for Research and Innovation.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species abundances of marine commercial fish species and long-distance migratory fishes	
ID	24
Realm	Marine
EBV class	Species populations
EBV name	Species abundances of marine commercial fish species and long-distance migratory fishes
Step in identification process	User & Policy Needs Assessment
Definition	The estimated count of individuals of commercially relevant marine fish species and long-distance migratory fishes in EU's marine waters within contiguous spatial units (grid cells) over time.
Metric	- Estimated count of individuals - Modeled relative abundance
Spatial resolution unit	50 × 50 km – 200 × 200 km
Temporal resolution unit	1 year
Taxonomic/ ecosystem focus group	Commercial marine fish species listed in the Common Fisheries Policy
Main European initiatives and description of current bottlenecks	

There is not a single monitoring network integrating data of marine commercial fishes across the whole of Europe. However, there are different regional integration initiatives collecting data on the relative abundance of marine fishes that could potentially provide data to generate this EBV. Among them, we highlight here the fish data gathered by the International Council for the Exploitation of the Sea (ICES), corresponding to surveys undertaken through trawls (main source of fish monitoring data in Europe - Jessop et al. 2022). These are mostly directed to commercial species and as such it will prove a good source of data to generate this EBV.

Possible bottlenecks in the generation of this EBV from ICES data relate to:

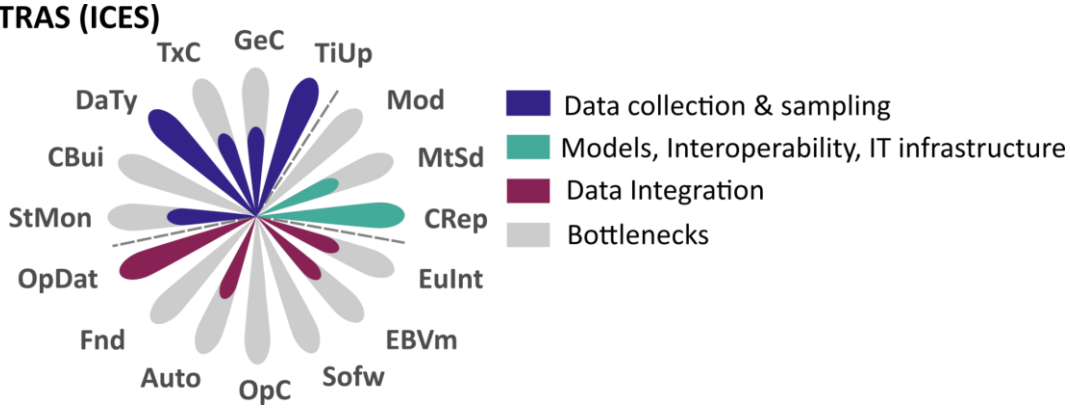
1) Data collection and sampling. The International Council for the Exploitation of the Sea (ICES) holds a database on pelagic and bottom fish trawl surveys collected in the North-east Atlantic and Baltic seas (DATRAS online database). Survey data comes from standardized survey protocols, but these differ across regions (e.g. between the Baltic and the Western and Southern Areas) (**StMon** partial bottleneck). Data reported to DATRAS includes the age disaggregated abundance of captured fish species that could potentially be used to generate this EBV. The survey data currently covers the Baltic Sea, Skagerrak, Kattegat, North Sea, English Channel, Celtic Sea, Irish Sea, Bay of Biscay and the eastern Atlantic from the Shetlands to Gibraltar; therefore, the geographic coverage of data is not complete as it does not include data from the Mediterranean sea (**GeC** partial bottleneck); this translates in the lack of representation of Mediterranean species in the database (**TxC** partial bottleneck). There are more than 45 years of continuous time series data in DATRAS, and survey data are continuously updated by national institutions (more than once a year in some cases), so the data availability matches the temporal resolution sought by this EBV.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

2) Models, interoperability, and IT infrastructure. The ICES does not use models to estimate global and age disaggregated abundance estimates (indices of Catch per Unit Effort) (**Mod** bottleneck). The DATRAS dataset has defined a vocabulary and a set of standards for data reporting to ease data handling, integration, and availability but there are not common metadata standards followed across all surveys (**MtSd** partial bottleneck). Data is centralized and stored in the ICES data portal.

3) Data integration. There is not a single integration initiative that collects and processes data of marine commercial fishes abundance and distributions across Europe (only regional and national ones, being the Mediterranean sea clearly underrepresented) (**EuInt** partial bottleneck). The products generated from DATRAS (i.e. indices of fish stock abundance and estimates of commercial species) do partially match the EBV (DATRAS does not model relative abundance of commercial fish stocks but we could not assess whether this covers all long distance migratory species; **EBVm** partial bottleneck). Since there are not currently models in place to map the distribution of marine fish species, there is no open code or user-friendly software to report here (**OpC** and **SoFw** bottlenecks). Data streams are not fully automated: worksheets in .csv format are submitted to the DATRAS via the ICES platform but there is little information about the level of automatization of the data flows of the underlying datasets (**Auto** partial bottleneck). If funding for adequate survey effort delivering the data exists, it is that of the countries reporting data to DATRAS and ICES (**Fnd** bottleneck). DATRAS data products (such as Catch per area of fishing effort estimates - CPUE- or indices) and raw data, can be freely downloaded according to the ICES Data policy.

DATRAS (ICES)



See Jessop et al. (2022) for a summary of other national and regional fish monitoring efforts available across Europe.

References:

Jessop, A., Chow, C., Dornelas, M., Pereira, P., Sousa-Pinto, I., Hernández Chan, S., Junker, J., Soares, J., Ratnarajah, L., Fernández, M., Mendo, T. (2022) MarBioME. Overview and assessment of the current state of Biodiversity Monitoring in the European Union and adjacent marine waters. European Commission. Directorate General for Research and Innovation.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species distributions of marine birds	
ID	25
Realm	Marine
EBV class	Species populations
EBV name	Species distributions of marine birds
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of European marine bird species at their breedin sites over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	10 x 10 km - 50 × 50 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Marine bird species indicated in the Review and analysis of Member States' 2018 reports Descriptor 1: Species biological diversity

Main European initiatives and description of current bottlenecks

There is not a single monitoring network integrating data of marine birds across the whole of Europe. However, there are different regional integration initiatives collecting data on marine bird species presence and abundance across Europe that could potentially provide data to generate this EBV. Among them, we highlight here the OSPAR and HELCOM conventions and the EBBA2.

The OSPAR convention already collects data on breeding seabird colonies (incl. gulls and terns), breeding waterbirds (incl. waders) and wintering and passage water birds (incl. waders) in countries of the Northeast Atlantic, that gets integrated to generate a Marine Bird Abundance indicator. Similarly, the HELCOM generates a Marine Breeding birds abundance indicator integrating data for 6 marine bird species surveyed across the Baltic Sea. There is currently a joint working group on Marine Birds between OSPAR-HELCOM and ICES, to enable long-term planning and delivery of significant products across both regions. The European Seabirds At Sea (ESAS) ensembles offshore monitoring data on seabirds (e.g., from ship-based surveys) across the North Sea using a standardized data collection methods and complements coastal surveys of breeding and wintering marine birds within the framework of the OSPAR and HELCOM conventions. For example, ESAS data gets integrated to coastal data to generate two indicators "Abundance of waterbirds in the breeding season" and "Abundance of waterbirds in the wintering season".

There is also a newly developed initiative to integrate biodiversity (including marine birds) and environmental monitoring at the Mediterranean level but it is still in its preliminary stages. However, this initiative envisages the generation of indicators following those used by the OSPAR commission.

Possible bottlenecks in the generation of this EBV from existing data collected by these networks (OSPAR, HELCOM) relate to:

1) Data collection and sampling. In both OSPAR and HELCOM Marine Birds programs, systematic surveys are currently carried out and funded by national monitoring schemes. Regional/European monitoring standards have not been developed yet, making it difficult data integration at these scales (**StMon** partial bottleneck).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

HELCOM reports difficulties to integrate data from coastal surveys across countries because of the use of different sampling methods. Data collected are species counts (breeding pairs) but, because the geographic location of colonies sampled is known, this information could be used to generate species distribution maps. OSPAR collects data for up to 100 species, while HELCOM only for 6, so the taxonomic coverage is not complete (**TxC** bottleneck). Geographic coverage is also incomplete, as on the one hand, these two programs focus on North-East Atlantic and Baltic countries, leaving the Mediterranean Sea completely uncovered; on the other hand, the OSPAR Marine Bird Program reports data gaps for some countries which have limited seabird colony monitoring, or which are not contributing any data to the regional indicator (**GeC** bottleneck). Capacity building (if exist) is that of national monitoring schemes (**CBui** bottleneck).

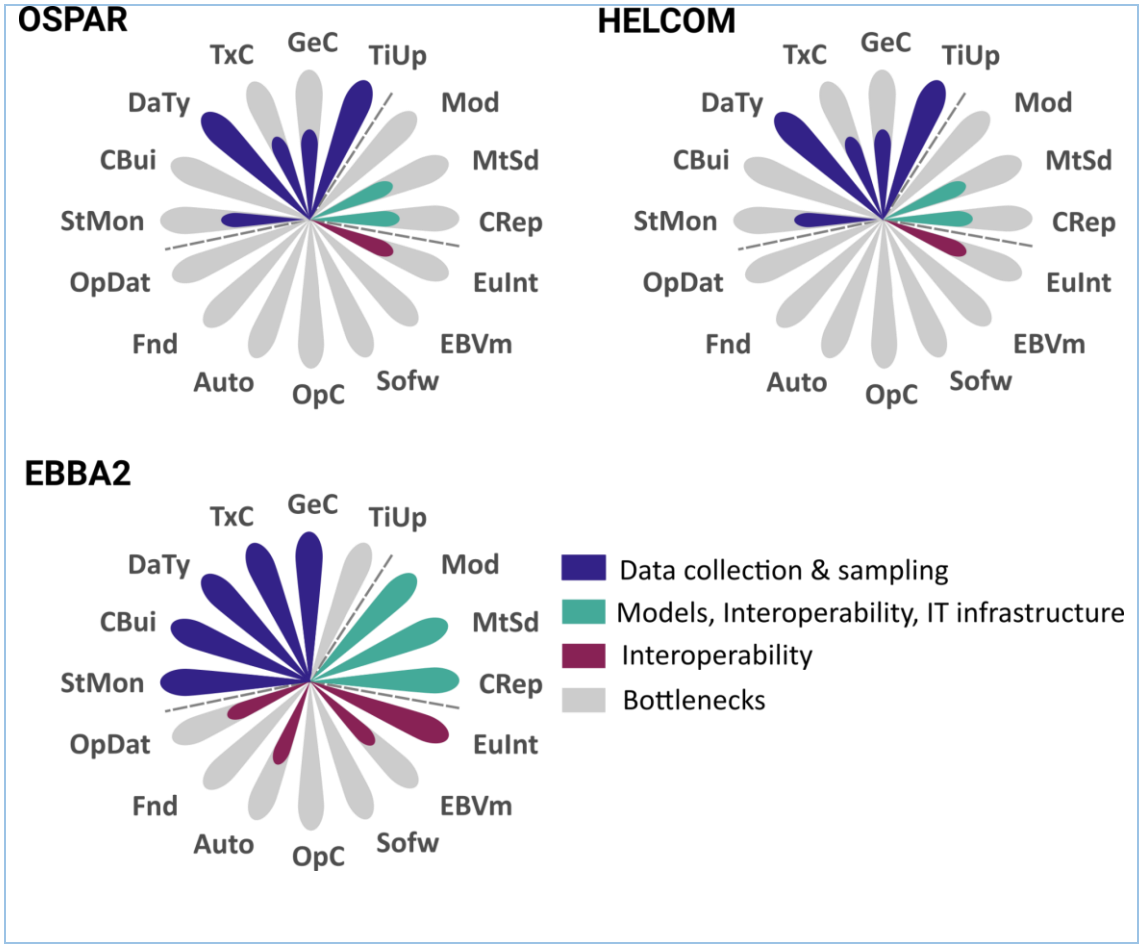
2) Models, interoperability and IT infrastructure. These initiatives do not use models to generate spatially explicit estimates of species distributions (**Mod** bottleneck). The Marine Bird Abundance indicator integrates species trends (as estimated from the software TRIM). Each contributing country has its own data Quality Assurance and Quality Control Protocols, with European standards to be developed (**MtSd** partial bottleneck). Similarly, each contributing country has its own storage mechanism. Ultimately, data is stored in the OSPAR and HELCOM Marine Bird Databases hosted by ICES Data Centre via the ICES Biodiversity Data Portal (this data is currently not available at ICES) (**CRep** partial bottleneck).

3) Data integration. There is not a single integration initiative that collects and processes data of marine birds' abundance and distributions across Europe (only regional ones, with Mediterranean Sea being clearly underrepresented) (**EUInt** partial bottleneck). Products from OSPAR and HELCOM (indicators of abundance trends) do not match the specifications of this EBV (**EBVm** bottleneck). As there are not currently models in place to map the distribution of marine fish species, there is no open code or user-friendly software to report here (**OpC** and **SoFw** bottlenecks). Data streams are not automated with, for example, OSPAR data being submitted to the coordination node via written procedure into a standard set of excel worksheets (**Auto** bottleneck). Currently, funding for adequate survey effort delivering the data basis needed for OSPAR and HELCOM indicator work (as well as other requirements by conservation directives) is still lacking in several countries (**Fnd** bottleneck). Raw data is currently not available (**OpDat** bottleneck).

The **Second European Breeding Bird Atlas (EBBA2)** (Keller et al. 2020) mapped the distribution of 63 marine and coastal species at 50 x 50 km resolution for the 2013-2017 period (distribution of nesting places on land), but it also presents bottlenecks when it comes to generating this EBV including insufficient temporal and spatial resolution (see a detailed description of bottlenecks regarding the generation of EBV *Species distributions of all birds* in Annex III).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023.* DOI: https://doi.org/10.3897/arphapreprints.e103765



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species distributions of marine mammals	
ID	26
Realm	Marine
EBV class	Species populations
EBV name	Species distributions of marine mammals
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of European marine mammal species within contiguous spatial units (grid cells) in EU's marine waters over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	10 x 10 km - 50 × 50 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Marine mammal species indicated in the Review and analysis of Member States' 2018 reports Descriptor 1: Species biological diversity
Main European initiatives and description of current bottlenecks	

There is not a single monitoring network integrating data of marine mammals across the whole of Europe. For cetaceans, there are currently not even internationally coordinated monitoring schemes at the relevant spatial scales needed for the assessment of these wide-ranging species. The available integration initiatives collecting data on marine mammal species presence and abundance data are national or regional (e.g. OSPAR - North-East Atlantic or HELCOM - Baltic sea) and generally focus on a single or a reduced set of species, making it difficult to assess the overall gaps and bottlenecks regarding the monitoring of this taxonomic group across European seas (but see Jessop et al. 2022 for a review of monitoring gaps of marine mammals across all scales).

For example, the abundance and distribution of cetaceans is one of the agreed common OSPAR indicators (M4), integrating data from different large-scale international surveys such as SCANS (Small Cetaceans in European Atlantic waters and the North Sea) and CODA (Cetacean Offshore Distribution and Abundance in the European Atlantic) plus other unknown number of smaller-scale surveys. It estimates species trends and generates maps of predicted abundances across the North-East Atlantic (modelled density surfaces). Therefore, data used for the generation of M4 indicators could potentially serve the generation of this EBV, but only for a limited number of species. Moreover, OSPAR recognizes there is a general lack of data to fulfil the six-yearly reporting requirements under the Habitats Directive and the Marine Strategy Framework Directive for most cetacean species/area units (also the temporal resolution sought for this EBV), since the frequency of these large-scale surveys is > 10 years.

Another OSPAR indicator on marine mammals is the indicator on *Seal abundance and distributio*n (M3). This indicator integrates estimates of seal numbers from monitoring programmes that count seals on land when they are moulting or breeding and are run on a regular basis. The frequency and timing of seal surveys varies among OSPAR Contracting Parties and take place during one or more seal key life stages, such as moulting or breeding. The monitoring is conducted in the Great North Sea and the UK part of the Celtic Sea, so its geographic coverage is very restricted. A similar seal monitoring is carried out across the

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Baltic sea under the umbrella of HELCOM convention; data collected gets integrated to estimate population trends and abundance of seals and seals' distribution maps across the Baltic.

In the Mediterranean region we highlight the 'Fixed Line Transect Mediterranean monitoring Network' that coordinates a continuous monitoring of mega and macro marine fauna (cetaceans, sea turtles, seabirds, and other macro marine fauna) since 2007. However, data from this network has not been integrated yet for example to generate reference indicators across the Mediterranean Sea. These data have been used to investigate species presence and distribution for conservation purposes, habitat use, long-term trends, correlation with environmental features and influence of the main threats of maritime traffic and marine litter. All these studies have for the moment been restricted to scientific publications.

References:

Jessop, A., Chow, C., Dornelas, M., Pereira, P., Sousa-Pinto, I., Hernández Chan, S., Junker, J., Soares, J., Ratnarajah, L., Fernández, M., Mendo, T. (2022) MarBioME. Overview and assessment of the current state of Biodiversity Monitoring in the European Union and adjacent marine waters. European Commission. Directorate General for Research and Innovation.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Distributions of marine turtle species nesting grounds	
ID	27
Realm	Marine
EBV class	Species populations
EBV name	Distributions of marine turtle species nesting grounds
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of marine turtle species nesting ground in EU's coastline over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	10 x 10 km - 50 × 50 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Turtle species indicated in the Habitats Directive and the Review and analysis of Member States' 2018 reports Descriptor 1: Species biological diversity: *Caretta caretta* (Loggerhead turtle), *Chelonia mydas* (Green turtle), *Dermochelys coriacea* (Leatherback turtle), *Eretmochelys imbricata* (Hawksbill turtle), *Lepidochelys kempii* (Kemp's Ridley turtle)

Main European initiatives and description of current bottlenecks

There is not a single initiative integrating data on marine turtle nesting sites over Europe. There are however a few regional short- term regional or national initiatives monitoring turtle nesting sites such as those framed within the EuroTurtles or the MedTurtle LIFE projects; moreover, guidelines have been developed by the UNEP to standardize to the maximum extent monitoring protocols for nesting areas of *Careta careta* and *Chelonia mydas* across the Mediterranean, but to what extent this methodology has been applied it is unknown.

In the MarBioMe review, Jessop et al. (2022) found sea turtles to be among the marine taxonomic groups less monitored across Europe with most sea turtle monitoring programmes being conducted by the Republic of Ireland (i.e. large Geographic Coverage gaps) and/or focused on recording offshore observations rather than nesting grounds (e.g., the 'Fixed Line Transect Mediterranean monitoring Network').

The lack of information on exactly what these monitoring programs are, what is the temporal resolution of data collection or to what extent data gets shared and integrated among them, complicates the evaluation of the current bottlenecks on data flows for the generation of this EBV at a European level.

References:

Jessop, A., Chow, C., Dornelas, M., Pereira, P., Sousa-Pinto, I., Hernández Chan, S., Junker, J., Soares, J., Ratnarajah, L., Fernández, M., Mendo, T. (2022) MarBioME. Overview and assessment of the current state of Biodiversity Monitoring in the European Union and adjacent marine waters. European Commission. Directorate General for Research and Innovation.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species distributions of invasive alien marine taxa of European concern	
ID	29
Realm	Marine
EBV class	Species Populations
EBV name	Species distributions of invasive alien marine taxa of European concern
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of invasive freshwater species (as specified in the Consolidated List of Invasive Alien Species of Union Concern) in Eu's marine waters within contiguous spatial units (grid cells) over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	1 x 1 km - 10 x 10 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Marine species as specified in the Consolidate List of Invasive Alien Species of Union Concern
Main European initiatives and description of current bottlenecks	

The European Alien Species Information Network (EASIN) is an integration initiative coordinated at the European level (Joint Research Centre, European Commission) that aggregates, integrates, and harmonizes spatio-temporal data for alien species (AS) and IAS across Europe. The species catalogue and geodatabase are regularly updated following the continuous revision of species reports in the literature, data published by EASIN data partners and the official reports by Member States competent authorities. The EASIN catalogue (v9.0- 19.07.22) includes 1,417 marine alien species, of which 3 are Invasive Alien Species of Union concern (IAS UC) (EU Regulation 1143/2014). EASIN facilitates the exploration of data and information from existing monitoring networks and programs available from a variety of distributed information sources by providing tools and interoperable web services, compliant with internationally recognized standards.

Possible bottlenecks in the generation of this EBV from existing data collected by EASIN data relate to:

1) Data collection and sampling. EASIN carries out the systematic collection of data on IAS in Europe, does not come exclusively from standardized monitoring programs but from a variety of sources, including literature review, occasional observations and data portals (**StyMon** partial bottleneck); EASIN has demonstrated capacity building through different activities including: i) ad hoc assessments, such as the Baseline Distribution of IAS of Union concern; ii) support to the establishment of surveillance systems in compliance with EU AS policy; iii) citizen science programs, (e.g. the development of the smartphone app 'IAS in Europe' freely available and already adopted in the context of several projects and for official monitoring; iv) training to teachers and students (e.g., the MOOC course "Have you seen an alien?"), among others. EASIN collaborates with the Member States competent authorities and national experts to update and validate the datasets relevant in the EU policy context, to fulfil the EU mandate of setting up national monitoring networks for IAS; however, the extent to which this has been implemented is not fully known. For example, a recent European evaluation on the reporting of non-

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

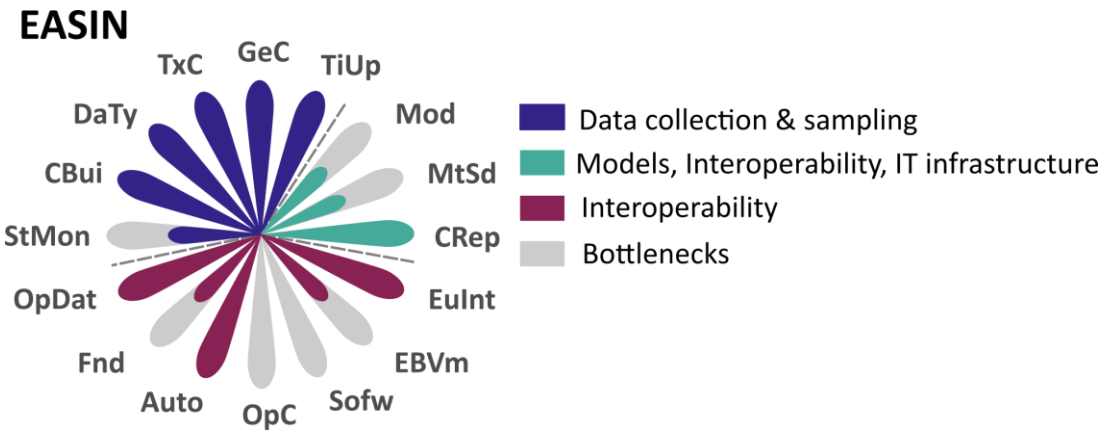
indigenous species under the Marine Strategy Framework Directive (Tsiamis et al. 2021) showed that only 8 Member States reported information regarding the abundance and spatial distribution of invasive alien species and only 3 MS reported information of the impacts of these species on natural habitats and species groups. Some bottlenecks regarding data collection by MS may relate to 1) lack of communication between national coordinators & EU funded projects working on IAS in a country (e.g. LIFE projects); 2) at the national level, people in charge of monitoring AS belong to different admin bodies among which there is little communication and coordination and 3) complex administrative structure in some countries (federal, regions, etc.) and how the environmental competences are distributed among them, making it difficult to standardize and coordinate data collection tasks and to facilitate data flows and integration. Currently, the EASIN AS Geo Database contains occurrence records for more than 14,000 species, across 40 different countries (EU & beyond) (including data for the 88 species in the consolidated list of IAS of Union concern- European Union (EU) Regulation 1143/2014). Because the data is georeferenced, it could potentially be used to build species distribution models. The EASIN GeoDatabase is updated on average 3 times/ year, but some information sources are updated less regularly depending on effort for the update and the periodicity of their own updates. Recent EASIN database backend developments data updating will be facilitate more regular updates.

2) Models, interoperability, and IT infrastructure. Habitat suitability models have been already fit within the EASIN framework for the freshwater species *Elodea nuttallii* from EASIN data across EU (Steen et al. 2019); however, this is not a product regularly generated by this initiative, but a case study (**Mod** partial bottleneck), although there is ongoing work in this direction for IAS of Union concern: models are used to evaluate the range of expansion at both European and national levels, to understand whether niche of the species is at equilibrium in its invaded range or to predict the probability of occurrence of the species in Europe considering bioclimatic variables. The AS GeoDatabase (the central repository of EASIN curated by the Joint Research Centre of the European Commission) has developed a protocol to frequently retrieve data from the data partners; data are subsequently transformed by converting the harvested data to the EASIN Data Model through the following steps: validation, cleansing and standardization, geocoding, mapping, application of quality rules and finally loaded on the Geodatabase (Datawarehouse). However, while standards have been defined to harmonize data across source databases, these do not necessarily apply to the underlying data (**MtSd** partial bottleneck).

3) Data integration. The products generated by this initiative partially match the EBV, especially in terms of spatial resolution: the EASIN species mapping tool shows the distribution of species at the country level or at 10 x 10 km grid cells (**EBVm** partial bottleneck). The test run on modelling the distribution of *Elodea nuttalli* has used Maxent, a machine-learning presence-background model that has a user-friendly interface facilitating its use by non-modellers (Phillips et al. 2006, 2008). However, this model is not routinely used by EASIN to model the data in the Alien species geodatabase, so there is no user-friendly software or open code to report in this regard (**Soft** and **OpC** bottlenecks). Data flows are automatized at different levels, from data collection to data integration: the EASIN has developed the "IAS Europe" smartphone App to promote the report of sightings of IAS of Union Concern by citizens and their integration into the EASIN GeoDatabase (and its consequent harmonization with other data retrieved from Data Partners). The process of

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

retrieving the data from the Data Partners is done through the EASIN Data Broker system, which can retrieve the species occurrences and related information (date, source) from different kinds of data sources and store them in a normalized database structure. The EASIN has also developed and manages the NOTSYS platform as the official tool for EU Member States to notify the Commission and inform the other Member States as required by Reg. 1143/2014 on Invasive Alien Species (IAS). In particular, the tool has been designed to facilitate a timely comprehensive notification of new detections of IAS of Union concern and related eradication measures. Funding is a key bottleneck to the generation of this EBV: while the European Commission has set up EASIN as the official information system supporting Member States in the implementation of the Regulation (EU) No 1143/2014 , it does not directly co-participate in the expenses of either monitoring or eradication actions by national nodes which solely relies on national budgets; funding from the European Commission for monitoring of IAS is available via existing financial instruments such as LIFE, H2020, Cohesion or Regional Development funds; however, funding is warrant in the mid - long - term to maintain the EASIN data infrastructure and the development of new data collection tools (e.g., a web base platform for reporting citizen science data) (**Fnd** partial bottleneck). Data in the EASIN GeoDatabase can be easily accessed and downloaded from the website. The openness of the data is key to redirect surveillance and trigger early warning systems.



References:

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological modelling*, *190*(3-4), 231-259.

Phillips, S. J. and Dudík, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31(2), 161-175.

Steen, B, Cristina Cardoso, A.C., Tsiamis, K., Nieto, K.,Engel J. and Gervasini, E. (2019) Modelling hot spot areas for the invasive alien plant Elodea nuttallii in the EU. *Management of Biological Invasions* 10(1): 151–170. URL .

Tsiamis, K., Boschetti, S., Palialexis, A., Somma, F. and Cardoso, A.C. (2021) Marine Strategy Framework Directive - Review and analysis of EU Member States' 2018 reports - Descriptor 2: Non-indigenous species; Assessment (Art. 8), Good Environmental Status (Art. 9) and Targets (Art. 10), EUR 30520 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-79-27700-2, doi:10.2760/7897, JRC123179. URL.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Marine ecosystem structure

EBV: Ecosystem distribution of hard coral habitats	
ID	32
Realm	Marine
EBV class	Ecosystem structure
EBV name	Ecosystem distribution of hard coral habitats
Step in identification process	Internal review process
Definition	Presence/absence or probability of occurrence of hard corals habitats in EU's marine waters within contiguous spatial units (grid cells) over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	10 x 10 m - 300 x 300 m
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Live hard coral covers defined by EMODnet Seabed Habitats

EBV: Ecosystem distribution of macroalgae canopy cover	
ID	33
Realm	Marine
EBV class	Ecosystem structure
EBV name	Ecosystem distribution of macroalgae canopy cover
Step in identification process	Internal review process
Definition	The estimated percentage of macroalgae canopy cover in EU's marine waters within contiguous spatial units (grid cells) over time.
Metric	Estimated canopy cover
Spatial resolution unit	10 x 10 m - 300 x 300 m
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Macroalgae canopy covers defined by EMODnet Seabed Habitats

EBV: Ecosystem distribution of seagrass habitats	
ID	34
Realm	Marine
EBV class	Ecosystem structure
EBV name	Ecosystem distribution of seagrass habitats
Step in identification process	Internal review process
Definition	Presence/absence or probability of occurrence of seagrass habitats in EU's marine waters within contiguous spatial units (grid cells) over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	10 x 10 m - 300 x 300 m

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Seagrass covers defined by EMODnet Seabed Habitats

The EBVs *Ecosystem distribution of hard coral habitats, Ecosystem distribution of marine macroalgae canopy cover* and *Ecosystem distribution of marine seagrass habitats*, are also defined as Essential Ocean Variables (EOV), and their distribution around Europe has been recently mapped by EMODnet (vector format) for the Atlas of Marine Life in Europe. The generation of distribution maps of these three EBVs is framed within the EMODnet broad-scale seabed habitat map for Europe (EUSeaMap). Maps were derived from a compilation of different data sources used including data from the Habitats Directive reporting, specific habitat locations collected in different monitoring programs (e.g. maps of threatened and/or declining habitats hard corals developed by the OSPAR convention) and others. The EuSeaMap project is under continuous development and improvement, and to date has been updated three times 2012, 2016, 2021. If the mapping of these habitats will continue the same frequency of update, these EMODnet products will match the temporal resolution sought by these EBVs. The resolution of the EuSeaMap and the habitat maps is of 100 x 100 m, so, within the range of the spatial resolution defined for this EBV. Given the diversity of data sources and methods involved in the mapping of these habitats, it is difficult to make a thorough assessment of the bottlenecks in data flows to the generation of these EBVs as it has been done for other EBVs.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ANNEX III: Terrestrial EBV fact sheets

Terrestrial species populations

EBV: Species distributions terrestrial birds	
ID	44
Realm	Terrestrial
EBV class	Species populations
EBV name	Species distributions of terrestrial birds
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of each European terrestrial bird species within contiguous spatial units (grid cells) during the breeding season across the EU over time.
Metric	- Binary presence/absence during breeding season - Probability of occurrence during breeding season
Spatial resolution unit	1 × 1 km – 10 × 10 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	All terrestrial birds of the EU (taxonomy based on the HBW and BirdLife Taxonomic Checklist, with focus on those bird species that are officially recognized in the List of birds of the European Union.
Main European initiatives and description of current bottlenecks	

The **Second European Breeding Bird Atlas (EBBA2)** (Keller et al. 2020) is the most recent and comprehensive European-wide integration initiative mapping the species distributions of all European terrestrial birds. The EBBA2 has mapped the distribution, abundance and breeding likelihood of 596 breeding birds (both native and non-native) across Europe at the 50 x 50 km grid resolution. For this purpose, an extensive collection of data from multiple sources was conducted at national level and then integrated at European level. Moreover, species distribution maps at 10 km resolution (probability of occurrence) were developed for 222 species using statistical modelling.

Possible bottlenecks in the generation of this EBV from existing data collected by the EBBA network relate to:

1) Data collection and sampling: EBBA2 targeted surveys (breeding period; 10 km2 squares) followed a standardized protocol (time surveys 60 – 120 min, carried out between 2013 – 2017) which derived from different ongoing monitoring programs depending on each country (data from national atlases, national breeding surveys, etc.). Thanks to external funding (by the MAVA foundation and others, as there was no EU funding in support of this initiative), the EBBA2 project contributed to capacity building in different countries; however, it will be difficult to replicate the sampling intensity and effort dedicated to EBBA2 with a frequency lower than five years for all species (**TiUp** bottleneck); however, and because of the joint work between EuropaBON and EBCC, species distributions for a set of 50 European bird species (in particular farmland species) will be generated to test the capacities and limitations to update maps and analyse change with a frequency of 5 years (EBBA

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

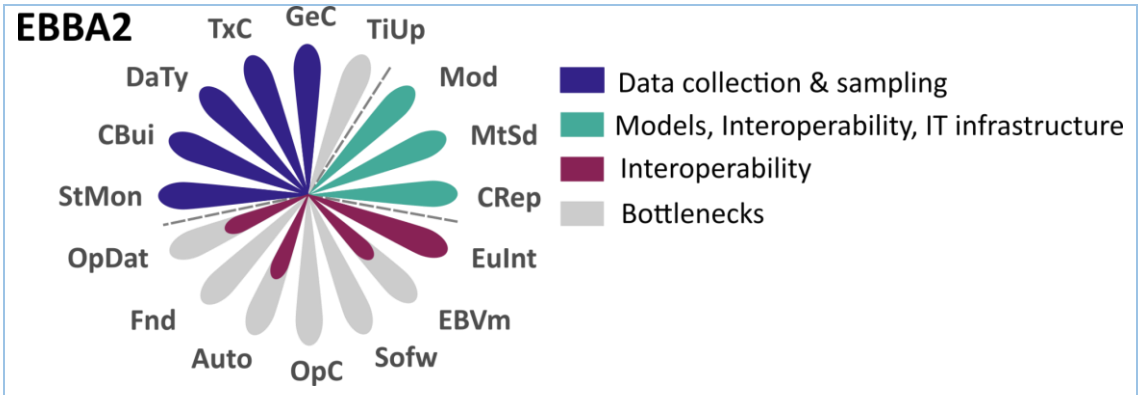
Live Farmland initiative; this initiative will be showcased in EuropaBON Deliverable 5.1).

2) Models, Interoperability and IT infrastructure: there are no bottlenecks to report in this regard: the EBBA2 used models to map the distribution of bird species (probability of occurrence), designed metadata standards for data reporting and aggregation, and developed tools to harmonize and centralize data at the European level; however, see below in "data integration" some potential bottlenecks relate to the complexity of the models used by this initiative (**OpC** and **Soft** bottlenecks).

3) Data integration: while EBBA2 has already generated continuous predictions of the distributions of all birds across Europe, EBBA2 products only partially match the definition of this EBV (with divergences in taxonomic completeness, temporal and spatial resolution demanded - only 222 species were modelled at 10 km resolution and none at a finer resolution; **EBVm** partial bottleneck). The 10 x 10 km grid species distribution models generated by the project were fit using R (a priori, a non-user-friendly software), and the code used to fit the models is not openly available (**OpC** and **Soft** bottlenecks). The models used to generate 10 x 10 km species distribution maps in EBBA2 (probability of occurrence) were complex (especially the site occupancy models) and require specific data from the observation process not always available because of the type of surveys carried out; there is a bottleneck due to the need of advanced technical programming skills and model knowledge to model species distributions. The EBBA2 was supported by external funding (the biggest donor being the MALVA foundation); the funding was directed to training, capacity building and bird sampling in countries not previously covered by any survey protocol but it also supported the overall coordination of the EBBA2 project. The funding support ended with the publication of the Atlas, hindering the continuity of sampling programs in some countries/regions and data integration tasks (**Fnd** bottleneck). While EBBA2 has made a huge effort to automate data flows (e.g., data aggregation at both the National and European levels has been automated through the use of MapViewers), the automatization of data flows from sampling plots to national coordinators was lacking (**Auto** partial bottleneck). All the data generated in EBBA2 is available either open access (50 x 50 observed occurrence data) or upon request (10 x 10 km modelled probability of occurrence data, 50 x 50 km breeding evidence data, 50 x 50 km abundance data and 50 x 50 km EBBA1 - EBBA2 change data). In the latter case, requests should be approved by an EBCC committee and data handling fees can be charged EBBA2 50 x 50 occurrence data are available under license CC BY 4.0, the rest of EBBA2 maps are © EBCC. Following EBCC data policy, raw 10 x 10 km sampling data is not owned by the EBCC but by its national partners, which could be asked for this data (**OpDat** partial bottleneck).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023.* DOI: https://doi.org/10.3897/arphapreprints.e103765



There are other European-coordinated initiatives collecting bird data in a systematic manner and on a yearly basis like the Pan-European Common Bird Monitoring Scheme (PECBMS), but the taxonomic coverage of this initiative (common birds, 170 bird species) falls short of achieving the mapping of all terrestrial birds across Europe (see a detailed description of this integration initiative in the EBV fact sheet below "Species abundances of terrestrial common birds").

The EuroBirdPortal (EBP) also compiles data on species distributions and breeding evidence at 1-10 km resolution across Europe on a daily basis, though data is mostly unstructured and semi-structured (see a detailed description of this integration initiative in the EBV fact sheet below "Species abundances of terrestrial migratory bird species"). The main bottleneck in this case is that, currently, the number of species covered by this integration initiatives is low (137 bird species in total).

The EBBA Live Farmland initiative (showcased in EuropaBON Deliverable 5.1), aims at combining the PECBMS and EBP datasets to produce maps of breeding distribution at a high temporal resolution (< 5 years), using the modelling experience gained by EBBA2.

References:

Keller, V., Herrando, S., Voříšek, P., Franch, M., Kipson, M., Milanesi, P., ... andFoppen, R. P. B. (2020). European breeding bird atlas 2: Distribution, abundance and change. European Bird Census Council and Lynx Editions, 967 pp.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species Abundances Terrestrial Birds: COMMON BIRDS	
ID	45 (a)
Realm	Terrestrial
EBV class	Species Populations
EBV name	Species abundances of terrestrial common birds
Step in identification process	User & Policy Needs Assessment
Definition	The estimated count of individuals of European common bird species within contiguous spatial units (grid cells) across the EU over time
Metric	- Estimated count of individuals - Modelled relative abundance
Spatial resolution unit	1 x 1 km, 10 x 10 km
Temporal resolution unit	1 year
Taxonomic/ ecosystem focus group	- Taxonomy based on the HBW and BirdLife Taxonomic Checklist - Common bird species as included in the Pan-European Common Bird Monitoring Scheme (PECBMS)

Main European initiatives and description of current bottlenecks

The **Pan-European Common Bird Monitoring Scheme (PECBMS)** integrates data on species occurrences and abundance of common birds across Europe following standardized methods. It generates species trends and multi-species trends indicators at the European level and on an annual basis. It also produces indicators for the four main European bioregions. All EU countries are included in the project. The data type collected in these standardized monitoring schemes could potentially be used to generate this EBV (spatially explicit abundances across Europe) but have not been used yet within the framework of this initiative for that purpose. PECBMS only covers the breeding season but some of the national Common Bird Monitoring Schemes participating in the project also collect data during winter.

Possible bottlenecks in the generation of this EBV from existing data collected by the PECBMS network relate to:

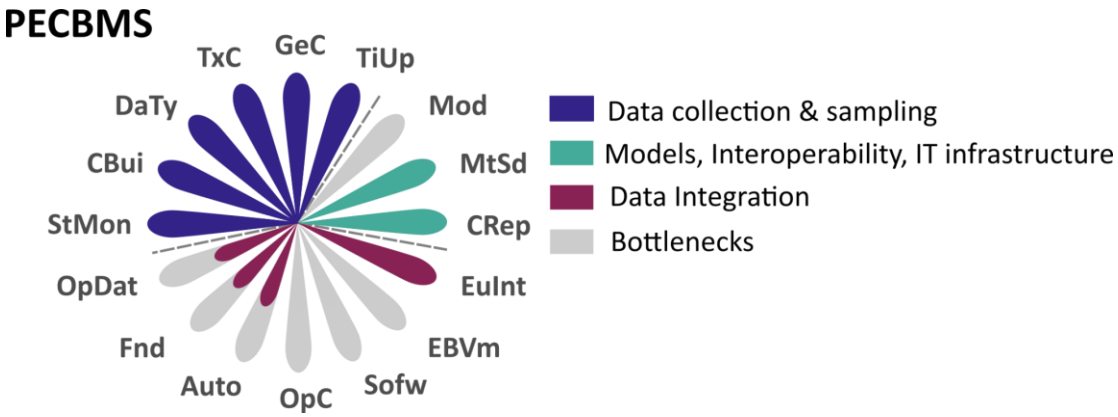
1) Data collection and sampling: PECBMS collects count data for all common bird species across Europe following standardized sampling protocols and on a seasonal basis (breeding counts) at a very fine spatial resolution (sampling transects, sampling point stations); while PECBMS transects are distributed across all European countries the degree of coverage of sampling transects in some countries is very low (especially in South and East European countries). Regular training for the network of national coordinators is provided by PECBMS (capacity building and training). Therefore, there are not bottlenecks to report regarding data collection for this monitoring initiative in relation to the generation of this EBV.

2) Models, Interoperability and IT infrastructure: The PECBMS integration initiative has demonstrated the capacity to use models to generate single- and multi-species indicator trends at the European level (TRIM model: TRends and Indices for Monitoring data). However, the program has not explored yet the use of species abundance models to obtain continuous spatially explicit estimates of species

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

abundances across Europe. The potential of using this type of surveys for mapping has been explored in scientific publications but not within the framework of the initiative (**Mod** bottleneck). PECBMS follows metadata standards to facilitate data harmonization and flows and counts on a central repository, curated by the Czech Society for Ornithology**.**

3) Data integration: There is a mismatch between the EBV description and its spatial resolution (1 x 1 – 10 x 10 km) and the products generated by this integration initiative (species trends reported at national and European level on an annual basis; **EBVm** bottleneck). The current software used for estimating population trends is user-friendly and has been made open to everyone through an R package (RTRIM-shell). While this same package cannot be used to map species abundances, it is to be expected that if this initiative were to derive spatially explicit abundance estimates, it would do so by relying on open software and code (**OpC** and **Sofw** bottlenecks). Data flows from national coordinators to the European integration node (the Czech Society for Ornithology) are automatized: national coordinators submit national species trends to an online platform using standardized formats and the system performs several automatic checks. Additional manual checks are carried out by the European coordination team; data integration at the national level is automatized in most cases but data collection still relies on field observations (**Auto** partial bottleneck). The PECBMS integration initiative has been supported by the European Commission since 2002 (3-year tender) but funding stability in the mid-long- term is not warranted; the same applies to both the national integration nodes and the survey programs (**Fnd** partial bottleneck). Raw data is only available upon request and subject to agreement by National coordinators who hold the ownership of data (**OpDat** partial bottleneck), but the PECBMS central coordination team has a database with all the raw data and just one centralized data request should be done (authorizations by national owners are coordinated by PECBMS).



The EuroBirdPortal (EBP) also compiles data on species distributions at 1-10 km resolution across Europe on a daily basis (see a detailed description of this integration initiative in the EBV fact sheet below "Species abundances of terrestrial migratory bird species"). The EBP data, though mostly consisting of unstructured and semi-structured data, could also be used to produce models of relative abundance (e.g., in combination with PECBMS data), however, the number of species covered by this integration initiative is, currently, quite low (137 bird species in total). For example, the EBBA Live Farmland initiative (showcased in EuropaBON Deliverable 5.1), will combine the PECBMS and EBP datasets to produce maps of breeding distribution of farmland birds at a high temporal resolution (< 5 years).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species abundances terrestrial birds: PRIORITY AND RARE BIRDS	
ID	45 (a)
Realm	Terrestrial
EBV class	Species Populations
EBV name	Species abundances of terrestrial priority and rare birds
Step in identification process	User & Policy Needs Assessment
Definition	The estimated count of individuals of European rare and priority bird species within contiguous spatial units (grid cells) across the EU over time.
Metric	- Estimated count of individuals - Modelled relative abundance
Spatial resolution unit	1 x 1 km, 10 x 10 km
Temporal resolution unit	1 year
Taxonomic/ ecosystem focus group	Taxonomy based on the HBW and BirdLife Taxonomic Checklist - Rare bird species as included in the Annex I of the Birds Directive (i.e., species with small populations or restricted local distribution) - Priority bird species as included in Annex I of the Birds Directive (i.e., in danger of extinction, vulnerable to specific changes in their habitat or requiring particular attention for reasons of the specific nature of the habitat)

Main European initiatives and description of current bottlenecks

The **Second European Breeding Bird Atlas (EBBA2)** (Keller et al. 2020) is the most recent and comprehensive European-wide integration initiative mapping the species distributions of all European terrestrial birds. The EBBA2 has mapped the distribution, abundance and breeding likelihood of 596 breeding birds (both native and non-native) across Europe at the 50 x 50 km grid resolution. For this purpose, an extensive collection of data from multiple sources was conducted at national level and then integrated at European level. Moreover, species distribution maps at 10 km resolution (probability of occurrence) were developed for 222 species using statistical modelling.

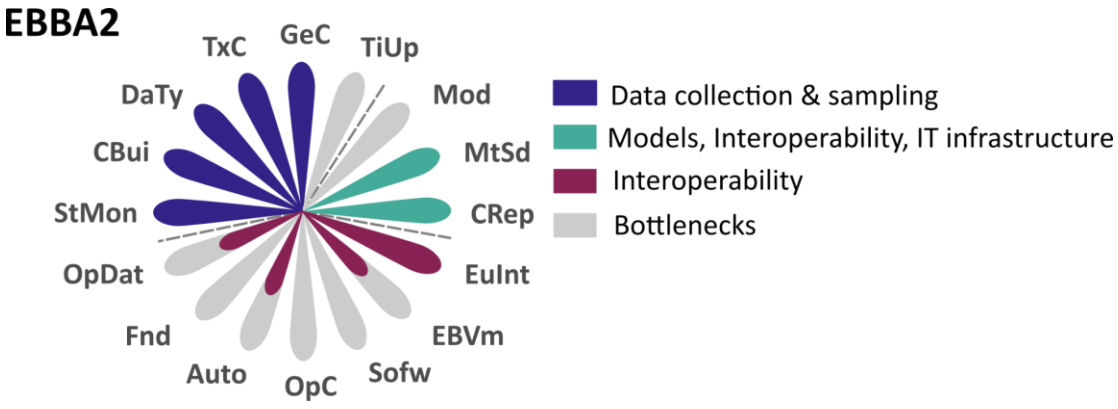
Possible bottlenecks in the generation of this EBV from existing data collected by the EBBA network relate to:

1) Data collection and sampling: EBBA2 targeted surveys (breeding period; 10 km2 squares) followed a standardized protocol (time surveys 60 – 120 min, carried out between 2013 – 2017) which derived from different ongoing monitoring programs depending on each country (data from national atlases, national breeding surveys, etc.). Thanks to external funding (by the MAVA foundation and others, as there was no EU funding in support of this initiative), the EBBA2 project contributed to capacity building in different countries; however, it will be difficult to replicate the sampling intensity and effort dedicated to EBBA2 with the annual frequency required to generate this EBV (**TiUp** bottleneck); however, and because of the joint work between EuropaBON and EBCC, species distributions for a set of 50 European bird species (in particular farmland species) will be generated to test the capacities and limitations to update maps and analyse change with a frequency of 5 years (EBBA Live Farmland initiative; this initiative will be showcased in EuropaBON Deliverable 5.1).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

2) Models, Interoperability and IT infrastructure: the EBBA2 used models to map the distribution of bird species (probability of occurrence) but it did not build spatially-explicit models of species abundance (50 x 50 km abundance maps resulted from the simple aggregation of the data collected in each pixel)(**Mod** bottleneck); The EBBA2 designed metadata standards for data reporting and aggregation, and developed tools to harmonize and centralize data at the European level.

3) Data integration: while EBBA2 has already generated continuous maps of the abundance of all birds across Europe, EBBA2 products only partially match the definition of this EBV (with divergences in temporal and spatial resolution demanded - species abundance maps are only available at 50 x 50 km; **EBVm** partial bottleneck). The 10 x 10 km grid species distribution models generated by the project were fit using R (a priori, a non-user-friendly software), and the code used to fit the models is not openly available (**OpC** and **Soft** bottlenecks; in any case, these two criteria do not apply because abundance models were not used within the framework of the EBBA2 project). The EBBA2 was supported by external funding (the biggest donor being the MALVA foundation); the funding was directed to training, capacity building and bird sampling in countries not previously covered by any survey protocol but it also supported the overall coordination of the EBBA2 project. The funding support ended with the publication of the Atlas, hindering the continuity of sampling programs in some countries/regions and data integration tasks (**Fnd** bottleneck). While EBBA2 has made a huge effort to automate data streams (e.g., data aggregation at both the National and European levels has been automated through the use of MapViewers), the automatization of data streams from sampling plots to national coordinators was lacking (**Auto** partial bottleneck). All the data generated in EBBA2 is available either open access (50 x 50 observed occurrence data) or upon request (10 x 10 km modelled probability of occurrence data, 50 x 50 km breeding evidence data, 50 x 50 km abundance data and 50 x 50 km EBBA1 - EBBA2 change data). In the latter case, requests should be approved by an EBCC committee and data handling fees can be charged EBBA2 50 x 50 occurrence data are available under license CC BY 4.0, the rest of EBBA2 maps are © EBCC. Following EBCC data policy, raw 10 x 10 km sampling data is not owned by the EBCC but by its national partners, which could be asked for these data (**OpDat** partial bottleneck).



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Birdlife international compiles and regularly updates data on bird populations at the European scale for the *Red List of European Birds* and as such, it could represent an interesting data source for the generation of this EBV (two publications to date in 2015 and 2021). This Red List estimates the regional extinction risk of all species of birds occurring regularly and naturally in Europe. This assessment is especially relevant for rare and priority species and has a special focus on bird conservation and environmental legislation. Data sources used for this assessment include data reported by the 28 Member States of the European Union (EU)8 under Article 12 of the EU Birds Directive, national reports from 22 countries and territories in Europe which are not part of the EU and additional data sources (e.g. data from the European Atlas of Breeding Birds and other bibliographic sources). The main bottleneck here is that data compiled (mostly number of breeding pairs of a given species at the national country) are not spatially explicit and therefore could not be used to generate this EBV.

The EuroBird Portal(EBP) also compiles data on bird species at 1-10 km resolution across Europe on a daily basis (see a detailed description of this integration initiative in the EBV fact sheet below “Species abundances of terrestrial migratory bird species”). The main bottleneck in this case is that, currently, the number of species covered by this integration initiatives is low (137 bird species in total) and only some priority and rare species are included.

The Pan-European Common Bird Monitoring Scheme (PECBMS) collects bird data in a systematic manner and on a yearly basis but as it is focussed on common birds species it could provide data for a few species (see a detailed description of this integration initiative in the EBV fact sheet below “Species abundances of terrestrial common birds”).

References:

Keller, V., Herrando, S., Voříšek, P., Franch, M., Kipson, M., Milanesi, P., ... and Foppen, R. P. B. (2020). European breeding bird atlas 2: Distribution, abundance and change. European Bird Census Council and Lynx Editions, 967 pp.

This project receives funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species Abundances Terrestrial Birds: MIGRATORY SPECIES	
ID	45 (b)
Realm	Terrestrial
EBV class	Species Populations
EBV name	Species abundances of terrestrial migratory bird species
Step in identification process	User & Policy Needs Assessment
Definition	The estimated count of individuals of European migratory bird species within contiguous spatial units (grid cells) across the EU over time.
Metric	- Estimated count of individuals - Modeled relative abundance
Spatial resolution unit	1 x 1 km, 10 x 10 km
Temporal resolution unit	Real-time
Taxonomic/ ecosystem focus group	- Taxonomy based on the HBW and BirdLife Taxonomic Checklist - Migratory bird species defined as full migrants in the European Red List
Main European initiatives and description of current bottlenecks	

The **EuroBirdPortal (EBP)** is the only integration initiative that can currently produce distribution maps for migratory species during the migratory periods, but also for other seasons (e.g. winter), as the data is collected year-round. The data collected by the EBP has already been used to generate the kind of metrics required for this EBV (spatially explicit abundances across Europe; cf. Gargallo *et al.* 2022) for a few species but this process has not been automated yet nor protocolized for that purpose within the framework of the EBP.

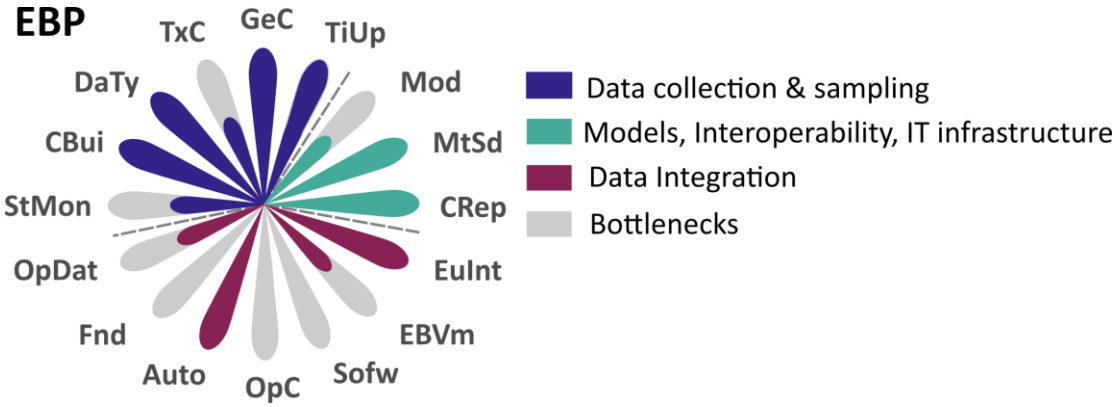
Possible bottlenecks in the generation of this EBV from existing data collected by the EBP initiative relate to:

1) Data collection & sampling. The EBP has the potential to collect data on all bird species occurring in Europe (including migratory species) but, currently, only the data for 137 species is stored in the EBP central data repository (**TxC** partial bottleneck). Data is collected across all European countries, but the degree of coverage is poorer in the South and, particularly, in the East and South-East. Data (species observations & counts) are gathered from online bird recording portals (e.g., ornitho) and include, essentially, both casual observations and data collected following simple systematic protocols (species lists). Only a small part of the data is collected following standardized monitoring protocols (**StMon** partial bottleneck). Training and capacity building for the network is provided during regular annual meetings or direct support to specific EPB partners. Data is updated daily with information up to the previous day, which for the purpose of the current EBV could be considered nearly as in real time.

2) Models, data interoperability and IT infrastructure. Although EBP data has already been used to produce weekly models of relative abundance at the necessary spatial resolution for a few species, the process has not been automated yet as a EBP data product nor extended to all migratory species (**Mod** partial bottleneck). The EBP follows metadata standards to ensure data harmonization among its different data sources. The EBP data is harmonized, managed, and stored in a central repository curated by the European Bird Census Council.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

3) Data Integration. The maps featured in the EBP viewer (www.eurobordportal.org) are aggregated by week and 30 x 30 km but in the EBP central data repository the casual data is aggregated at 10 x 10 km and date and the lists are in raw format (i.e. not aggregated) with date/timing and location given as precise location or at 10 x 10 km; therefore, the EBP maps partially match the specifications of this EBV (**EBVm** partial bottleneck: data available would not allow the mapping of species abundances at 1 x 1 km resolution). Though this integration initiative has already produced models of species relative abundances over space, these are not routinely done to generate EBP products and therefore there is no open code nor user-friendly software yet to report in this regard (**Sofw** and **OpC** bottlenecks). Data streams are automated: the local online portals collect most of their data through mobile apps in near-real time or shortly after it has been recorded in the field; data collected in data portals is then automatically transferred to the EBP daily. The EBP main developments have been possible thanks to the support of the LIFE programme (a new LIFE project proposal has been recently submitted), but funding stability in the mid- long- term is not warranted (**Fnd** bottleneck). EBP data is available upon request and subject to agreement by National coordinators who hold the ownership of data (**OpDat** partial bottleneck), but just one centralized data request to should be done as data is already centralized in the EBP databank (authorizations by national owners are coordinated by EBP).



References:

Gargallo, G., Davies, J.G., Faverjon, C., Kampichler C., Baillie, S.R., Cameron, A., Robinson, R.A. and Sierdsema, H. (2022) Development of a prototype early warning system for avian influenza in the EU based on risk-mapping. EFSA supporting publication 2022: EN-7762, 46pp. URL.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species abundance of selected terrestrial mammals	
ID	46
Realm	Terrestrial
EBV class	Species Populations
EBV name	Species abundance of selected terrestrial mammals
Step in identification process	User & Policy Needs Assessment
Definition	The estimated count of individuals of European terrestrial Carnivora, Artiodactyla and Chiroptera species within contiguous spatial units (grid cells) across the EU over time.
Metric	- Estimated count of individuals - Modelled relative abundance - Estimated counts of individuals in key underground sites (hibernation, breeding and transitional roost-sites) as defined by EUROBATS - Estimated counts of individuals in key overground sites (hibernation, breeding and transitional roost-sites) as defined by EUROBATS
Spatial resolution unit	1 x 1 km - 10 x 10 km
Temporal resolution unit	1 year
Taxonomic/ ecosystem focus group	Terrestrial Carnivora, Artiodactyla and Chiroptera species included in the European Red List.

Main European initiatives and description of current bottlenecks

The Atlas of European Mammals is the first coordinated initiative that has attempted to map the distribution of all mammals at 50 x 50 km resolution across Europe. The First Atlas of European Mammals was published in 1999 (Mitchell-Jones et al. 1999) but the Second Atlas of European Mammals (EMMA2) will be released in 2023.

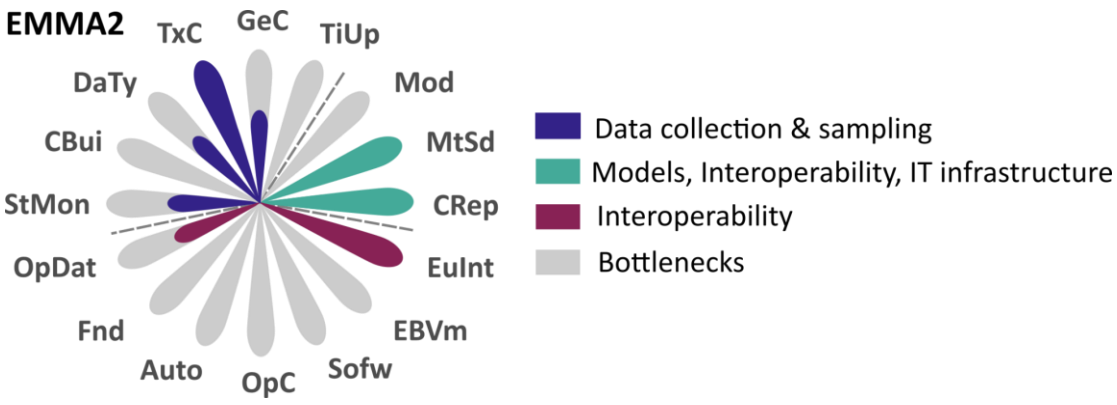
Possible bottlenecks in the generation of this EBV from existing data collected by EMMA2 data relate to:

1) Data collection and sampling. Most of the data used for the EMMA2 compilation will not come exclusively from systematic monitoring programs (**StMon** partial bottleneck). Data sources include national and regional databases, literature records and data portals. While the products generated by this initiative are binary (presence/absence of species within each 50 x 50 km grid cell), data contributors should report the number of individuals observed (counts); that raw data could potentially be used to estimate spatially-explicit species abundances across Europe, although probably not at 1 x 1 km or 10 x 10 km resolution (**DaTy** partial bottleneck). The integration initiative does not have capacity building (if it exists, it depends on national initiatives) (**Cbui** bottleneck). The EMMA2 gathers data from 21 countries but the coverage of data is uneven across them, so gaps of information are expected particularly in some areas (**GeC** partial bottleneck). The spatial resolution of EMMA2 maps does not match the finer spatial resolution sought in this EBV; moreover, the temporal resolution (frequency) of atlas updates is too low (> 10 years) (**TiUp** bottleneck), therefore the products generated by this integration initiative do not match the EBV definition (**EBVm** bottleneck in *Data Integration*).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

2) Models, interoperability, and IT infrastructure. EMMA2 maps will reflect the confirmed presence of the species aggregated at 50 x 50 km and there is no modelling involved in predicting species abundances across areas not covered by data (**Mod** bottleneck). Metadata standards are defined (with the intention of following Darwin Core) and data get integrated into a central repository by the European Mammal Foundation.

3) Data integration. Products generated by EMMA2 reflect species distributions rather than abundances; moreover, the spatial resolution of EMMA2 maps does not match the finer spatial resolution sought in this EBV, neither the temporal resolution (the frequency of atlas updates is too low - > 10 years) (**EVBm** bottleneck). Currently, there are no models in place to map the abundance of terrestrial mammals across Europe within the framework of the EMMA2, therefore, there is no open code or user-friendly software to report here (**OpC** and **SoFw** bottlenecks). While metadata standards are defined, data flows are not automated (**Auto** bottleneck): records are to be collected into CSV files or Access database (or similar) and submitted via email to coordinators who harmonize and integrate the data from the different sources. Funding has been secured for the publication of the Atlas but there is no funding to support the long-term continuation of this initiative (**Fnd** bottleneck). The EMMA2 original records will be held in a system that is accessible to researchers, with appropriate safeguards for sensitive records (rare species, vulnerable habitats or other reasons), though the great majority of records will be freely accessible at a higher resolution than in the Atlas, for example in national or regional atlases (**OpDat** partial bottleneck).



Also noteworthy to the generation of this EBV are emerging integration initiatives such as the European Observatory of Wildlife (EOW) or the Bat Monitoring Programme.

The EOW (a project by ENETWILD) aims to enhance collaboration among the stakeholders that monitor, conserve, and manage wildlife in Europe (with a special focus on mammals). It seeks to develop a framework where data will be comparable, interoperable, and openly accessed at the European level, providing guidelines for monitoring mammals in Europe (e.g., protocols to estimate wildlife density), training (e.g., on new tools for data processing and analysis) and compiling and harmonize existing databases on mammals. Data gathered through this portal following standardized monitoring protocols could eventually be used to obtain spatially-explicit estimates of mammal abundances/densities across Europe.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

The Bat Monitoring Programme's main goal is to use bats as indicators of the general state of nature using large-scale and long-term monitoring data on changes in bat populations across Europe. This programme has developed four standard protocols specifically designed for monitoring a certain group of bat species, either forest-dwelling, cave-dwelling or urban-dwelling species (ChiroRivers, ChiroHAbitats, ChiroBoxes and ChiroRoosts). The combination of all four monitoring programmes provides a complete image of the health status of all bat species populations. Data gets collected by professionals and volunteers (citizen science). Given data collection follows standardized protocols, it will be easy to harmonize it across countries and be potentially used for modelling the distribution of bats across Europe and generating this EBV. This monitoring programme has kicked off in Catalonia and is now being implemented at the Spanish scale, with plans to be expanded in all of Europe in the near future. The Bat Monitoring Programme's efforts are well-aligned with the Bat Monitoring and Surveillance Intersessional Working Group in the EUROBATS Commission. This group has worked on gathering data collected via systematic protocols by national surveillance groups, along with other data sources (e.g. literature review) to develop a prototype indicator of trends in European bat populations. These data get reported to EUROBATS (key underground and overgroud sites) by countries on a voluntary basis alongside the reporting to the Habitats Directive.

References:

Mitchell-Jones et al. (Eds) 1999 The Atlas of European Mammals. *Academic press*, p 484

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species distributions of all terrestrial mammals	
ID	47
Realm	Terrestrial
EBV class	Species Populations
EBV name	Species distributions of all terrestrial mammals
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of all European terrestrial mammal species within contiguous spatial units (grid cells) across the EU over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	10 x 10 km - 50 x 50 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	European terrestrial mammal species included in the European Red List.

Main European initiatives and description of current bottlenecks

The Atlas of European Mammals is the first coordinated initiative that has attempted to map the distribution of all mammals at 50 x 50 km resolution across Europe. The First Atlas of European Mammals was published in 1999 (Mitchell-Jones et al. 1999) but the Second Atlas of European Mammals (EMMA2) will be released in 2023.

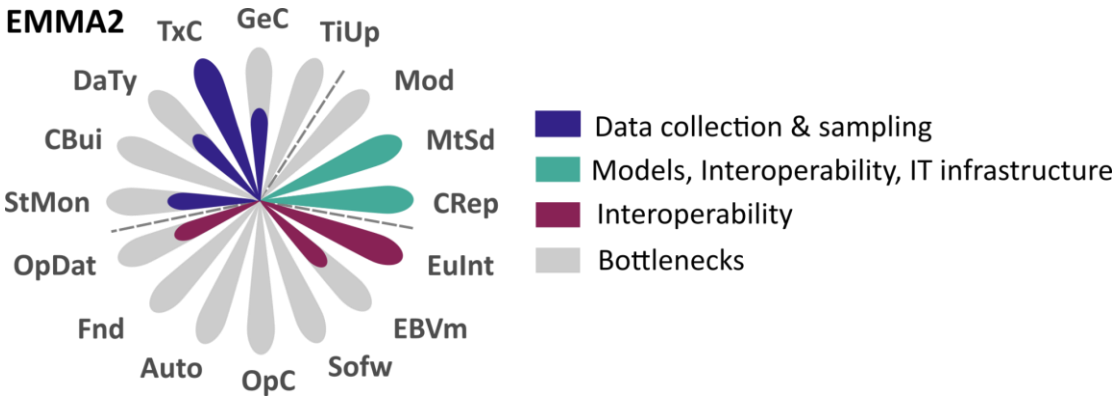
Possible bottlenecks in the generation of this EBV from existing data collected by EMMA2 data relate to:

1) Data collection and sampling. Most of the data used for the EMMA2 compilation will not come exclusively from systematic monitoring programs (**StMon** partial bottleneck). Data sources include national and regional databases, literature records and data portals. While the products generated by this initiative are binary (presence/absence of species within each 50 x 50 km grid cell), data contributors should report the number of individuals observed (counts); that raw data could potentially be used to make spatially-explicit predictions of species to unsampled areas, although probably not at 10 x 10 km resolution (**DaTy** partial bottleneck). The integration initiative does not have capacity building (if it exists, it depends on national initiatives) (**Cbui** bottleneck). The EMMA2 gathers data from 21 countries and the coverage of data is uneven across them, so gaps of information are expected particularly in some areas (**GeC** partial bottleneck). The spatial resolution of EMMA2 maps does not match the finer spatial resolution sought in this EBV (10 x 10 km) ; moreover, the temporal resolution (frequency) of atlas updates is too low (> 10 years) (**TiUp** bottleneck), therefore the products generated by this integration initiative do not match the EBV definition (**EBVm** bottleneck in *Data Integration*).

2) Models, interoperability, and IT infrastructure. EMMA2 maps will reflect the confirmed presence of the species aggregated at 50 x 50 km and there is no modelling involved in predicting species distributions across areas not covered by data (**Mod** bottlenecks). Metadata standards are defined (with the intention of following Darwin Core) and data get integrated into a central repository by the European Mammal Foundation.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

3) Data integration. Products generated by EMMA2 are maps of species distributions at 50 x 50 km resolution and as such they partially match the definition of this EBV; however, the frequency of atlas updates is too low (> 10 years) and it does not match the temporal resolution sought for this EBV (**EVBm** partial bottleneck). Currently, there are no models in place to map the distribution of terrestrial mammals across Europe within the framework of the EMMA2, therefore, there is no open code or user-friendly software to report here (**OpC** and **SoFw** bottlenecks). While metadata standards are defined, data flows are not automated (**Auto** bottleneck): records are to be collected into CSV files or Access database (or similar) and submitted via email to coordinators who harmonize and integrate the data from the different sources. Funding has been secured for the publication of the Atlas but there is no funding to support the long-term continuation of this initiative (**Fnd** bottleneck). The EMMA2 original records will be held in a system that is accessible to researchers, with appropriate safeguards for sensitive records (rare species, vulnerable habitats or other reasons), though the great majority of records will be freely accessible at a higher resolution than in the Atlas, for example in national or regional atlases (**OpDat** partial bottleneck).



Also noteworthy to the generation of this EBV are emerging integration initiatives such as the European Observatory of Wildlife (EOW) or the Bat Monitoring Programme.

The EOW (a project by ENETWILD) aims to enhance collaboration among the stakeholders that monitor, conserve, and manage wildlife in Europe (with a special focus on mammals). It seeks to develop a framework where data will be comparable, interoperable, and openly accessed at the European level, providing guidelines for monitoring mammals in Europe (e.g., protocols to estimate wildlife density), training (e.g., on new tools for data processing and analysis) and compiling and harmonize existing databases on mammals. Data gathered through this portal following standardized monitoring protocols could eventually be used to obtain spatially-explicit estimates of mammal distributions across Europe.

The Bat Monitoring Programme's main goal is to use bats as indicators of the general state of nature using large-scale and long-term monitoring data on changes in bat populations across Europe. This programme has developed four standard protocols specifically designed for monitoring a certain group of bat species, either forest-dwelling, cave-dwelling or urban-dwelling species (ChiroRivers, ChiroHAbitats, ChiroBoxes and ChiroRoosts). The combination of all four monitoring programmes provides a complete image of the health status of all bat species populations. Data gets collected by professionals and volunteers (citizen science). Given data collection follows standardized protocols, it will be easy to harmonize it across

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

countries and be potentially used for modelling the distribution of bats across Europe and generating this EBV. This monitoring programme has kicked off in Catalonia and is now being implemented at the Spanish scale, with plans to be expanded in all of Europe in the near future. The Bat Monitoring Programme's efforts are well-aligned with the Bat Monitoring and Surveillance Intersessional Working Group in the EUROBATS Commission. This group has worked on gathering data collected via systematic protocols by national surveillance groups, along with other data sources (e.g. literature review) to develop a prototype indicator of trends in European bat populations. These data get reported to EUROBATS (key underground and overground sites) by countries on a voluntary basis alongside the reporting to the Habitats Directive.

References:

Mitchell-Jones et al. (Eds) 1999 The Atlas of European Mammals. *Academic press*, p 484.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species distributions of terrestrial reptiles	
ID	48
Realm	Terrestrial
EBV class	Species Populations
EBV name	Species distributions of terrestrial reptiles
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of all European terrestrial reptile species within contiguous spatial units (grid cells) across the EU over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	1 x 1 km - 10 x 10 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	European terrestrial reptile species included in the European Red List.

Main European initiatives and description of current bottlenecks

The New Atlas of Amphibians and Reptiles of Europe (NA2RE, 2014) is the only Europe-wide coordinated initiative that has attempted to map the distribution of reptiles across the continent at 50 x 50 km grid. It generated maps for 218 taxa (it covers 145 species of reptiles) updated as of 2014 (no information about temporal dynamics).

Possible bottlenecks in the generation of this EBV from data flows in NA2RE relate to:

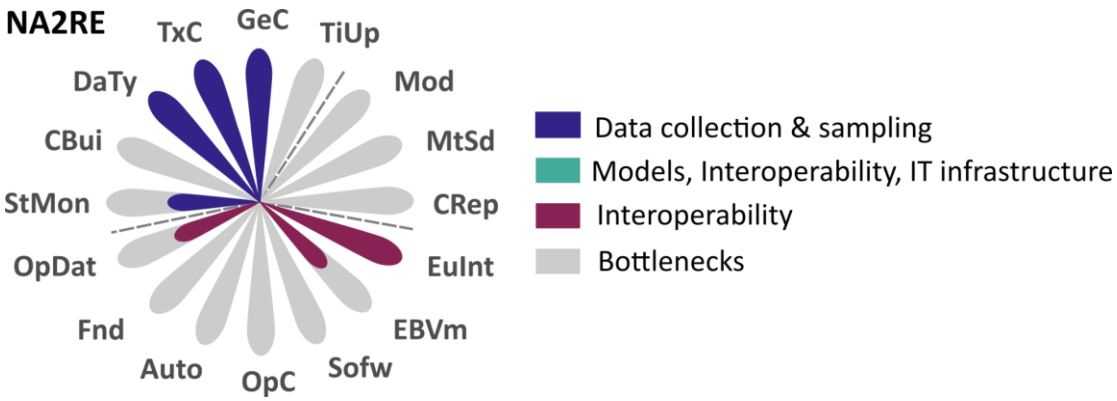
1) Data collection and sampling. Most of the data used for the Atlas compilation did not come from systematic monitoring programs (**StMon** partial bottleneck). Data sources included: (1) data published (in books or websites) or ongoing national atlases, (2) personal data kindly provided to the *Societas Europaea Herpetologica*, (3) the 1997 European Atlas, and (4) the Global Information Facility (GBIF). While data is available across most European countries (23; i.e., it does fulfil the criteria set to qualify for full Geographic Coverage in this assessment), the data coverage within each country was uneven and very scarce in many cases, with lack of funding and personnel being identified as limiting factors to set up national databases in many countries. The initiative did not have capacity building (if it existed, it was that of the national nodes) (**CBui** bottleneck). The Atlas is a snapshot of the distribution of amphibians and reptiles and therefore, does not match the temporal resolution criteria of this EBV (**TiUp** bottleneck).

2) Models, interoperability, and IT infrastructure. Maps in NA2RE reflect an aggregation of the raw data at 50 x 50 km and there is no modelling involved in predicting species distributions across areas not covered by data (**Mod** bottleneck) or at finer resolutions (e.g., 10 x 10 km). Data from different sources (collected using different standards) were centralized and harmonized for the atlas (**MtSd** bottleneck); the authors of the atlas recognize the challenge of finding better ways to gather species occurrence data across Europe and centralized it (**CRep** bottleneck) (Sillero et al. 2014).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

3) Data integration. The maps generated by this integration initiative do not fully match the EBV definition (they do not fulfil the criteria of spatial and temporal resolution; **EBVm** partial bottleneck). The NA2RE initiative did not use models to map the distribution of reptiles, therefore, there is no open code or user-friendly software to report here (**OpC** and **SoFw** bottlenecks). Data streams were not automated (**Auto** bottleneck). Funding supported the publication of the Atlas, but it is not available in the long-term to generate this EBV with the temporal frequency needed (**Fnd** bottleneck). The NA2RE maps are freely available for download but the raw data are not (**OpDat** partial bottleneck).



Another integration initiative at the European level retrieving information on reptiles distributions is the IUCN European Red List of Reptiles. However, most of the time IUCN maps are not linked to a monitoring program (but expert based) and represent a snapshot of the species distributions, so this initiative couldn´t contribute to the quantification of this EBV at the desired spatio-temporal scale.

References:

Sillero, N., Campos, J., Bonardi, A., Corti, C., Creemers, R., Crochet, P. A., ... and Vences, M. (2014). Updated distribution and biogeography of amphibians and reptiles of Europe. *Amphibia-reptilia*, 35(1), 1-31.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

EBV: Species abundance of butterflies	
ID	49
Realm	Terrestrial
EBV class	Species Populations
EBV name	Species abundance of butterflies
Step in identification process	User & Policy Needs Assessment
Definition	The estimated count of individuals of butterfly species within contiguous spatial units (grid cells) across the EU over time.
Metric	- Estimated count of individuals of grassland butterfly species - Modeled relative abundance of grassland butterfly species
Spatial resolution unit	10 x 10 km - 50 x 50 km
Temporal resolution unit	1 year
Taxonomic/ ecosystem focus group	Current list of butterfly species underlying the European grassland butterfly indicator, with extension to butterfly species from other habitats.

Main European initiatives and description of current bottlenecks

The **European Butterfly Monitoring Scheme (eBMS)** collects abundance data for > 312 butterfly and moth species, and a few bumblebees and dragonflies. The abundance data collected (counts per transect) gets mostly integrated to estimate population trends at the national and European levels (six indicators: grassland, forests, N2000, climate change, widespread species and urban butterflies).

Possible bottlenecks in the generation of this EBV from existing data collected by the eBMS data relate to:

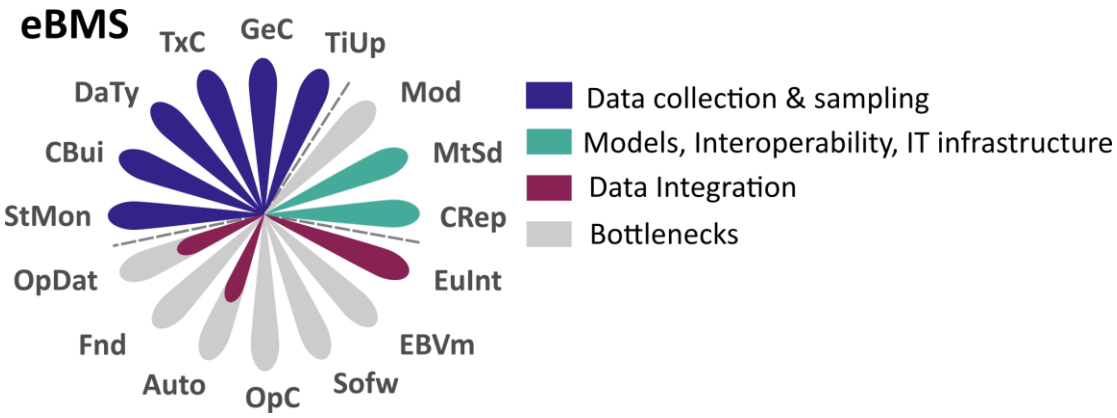
1) Data collection and sampling. There are no bottlenecks to report in this regard to the generation of this EBV. eBMS focuses its efforts mostly on the collection of butterflies and moths' abundance data using systematic transect counts that are visited several times per year (the spatio-temporal resolution of raw data is high and could potentially be used to generate this EBV). The geographic coverage of this monitoring program has recently been enlarged thanks to the support of a service contract from the European Union Directorate General for the Environment (ABLE project), so it currently covers 30 countries. The program promotes capacity building through the free publication of divulgation informative material for the identification of butterflies in different languages and countries, the offering of open training courses on the use of modelling tools for species trend estimations, the development of Apps to ease data collection and identification, etc.

2) Models, interoperability, and IT infrastructure. This initiative uses the TRIM model to estimate population trends and calculation of trend indices/products at the national level (training courses and materials are available on the eBMS website), there is no model proposal from the network for the quantification of species abundance from the sampled data in a spatially-explicit way, neither a proposal of the model platform or open code for this modelling (**Mod** bottleneck). The eBMS

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

follows metadata standards that allow harmonization of the data collected across the different European countries. The eBMS database is stored and managed in a centralized repository under the custody of the Butterfly Conservation Europe and the UK Center for Ecology and Hydrology.

3) Data Integration. There is not a match between the products generated by this initiative (grassland indicator; trends at the national and European level) and the EBV definition (in terms of spatial resolution 1 x 1 km and 10 x 10 km) (**EBVm** bottleneck). The current software used for calculating butterfly trends, indexes and indicators is R (*rbms* package; non-user-friendly software; **SoFw** bottleneck). However, the code is openly shared and training is provided to users of eBMS data (workshop contents are available on the eBMS website). While this same package cannot be used to map species distributions, it could be expected that if this initiative were to derive spatially explicit estimates of species occurrences, it would do so by relying on the same software and open code. However, currently there is no open code for spatially-explicit abundance models within the framework of the eBMS program (**OpC** bottleneck). eBMS has recently developed a mobile application (eBMS App) that allows volunteers to record species observations and abundances and directly upload them to the eBMS database (also to review the uploaded data). However, data collected (species lists, tables with data observations) also get reported via email to eBMS coordinators, so data streams are not fully automated (**Auto** partial bottleneck). Funding for coordination and integration tasks is not secured both at the European and national levels; for example, the last geographic expansion of the eBMS network was possible thanks to the ABLE project, funded by the European Union Directorate General for the Environment, for a period of two years from 2018-2020 (**Fnd** bottlenecks). Data is only available upon request (**OpDat** partial bottleneck).



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023.* DOI: https://doi.org/10.3897/arphapreprints.e103765

EBV: Species distributions of terrestrial priority invertebrates and key pollinators	
ID	50
Realm	Terrestrial
EBV class	Species Populations
EBV name	Species distributions of terrestrial priority invertebrates and key pollinators
Step in identification process	Internal review process
Definition	The presence/absence or probability of occurrence of priority invertebrates and key pollinator species within contiguous spatial units (grid cells) across the EU over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	10 x 10 km - 50 x 50 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	- Priority invertebrates as listed in the Annex II and Annex IV of the Habitats Directive - Key pollinator species as specified by the EU Pollinator Monitoring Scheme (EUPoMS)

Main European initiatives and description of current bottlenecks

Here we describe two European-wide integration initiatives that could provide data to the generation of species distribution maps of key pollinators: The **European Butterfly Monitoring Scheme (eBMS)** and the **European Pollinator Monitoring scheme (EU PoMS)**.

The **European Butterfly Monitoring Scheme (eBMS)** collects abundance data for > 312 butterfly and & moth species. It also collects data of other pollinator species (e.g. bumblebees) but only in a small proportion of the transects and the quality of the data is highly variable. The abundance data collected (counts per transect) gets mostly integrated to estimate population trends at the national and European levels (six indicators: grassland, forests, N2000, climate change, widespread species and urban butterflies).

Possible bottlenecks in the generation of this EBV from existing data collected by the eBMS data relate to:

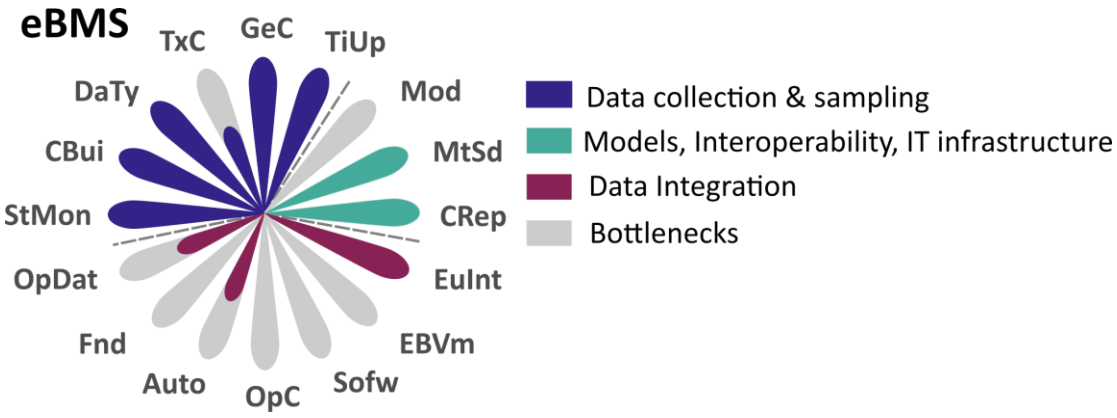
1) Data collection and sampling. The eBMS focuses its efforts mostly in the collection of butterflies and moths' abundance data using systematic transect counts that are visited several times per year (the resolution of raw data is high and could potentially be used to generate this EBV: species distributions from transect counts/occurrences). However, these two taxonomic groups represent only a small fraction of all European pollinators (lack of taxonomic completeness) (**TxC** bottleneck) and play a small role in the pollination of European crops. The geographic coverage of this monitoring program has recently been enlarged thanks to the support of a service contract from the European Union Directorate General for the Environment (ABLE project), so it currently covers 30 countries. The program promotes capacity building through the free publication of divulgation informative material for the identification of butterflies in different languages and countries, the offering of open training courses on the use of modelling tools for species trend estimations, the development of Apps to ease data collection and identification, etc.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

2) Models, interoperability, and IT infrastructure. This initiative uses the TRIM model to estimate population trends and calculation of trend indices/products at the national level (training courses and materials are available on the eBMS website); however, there is not a model proposal from the network for the generation of species distribution maps from the sampled data, neither a proposal of the model platform or open code for this modelling (**Mod** bottleneck). The eBMS follows metadata standards that allow harmonization of the data collected across the different European countries. The eBMS database is stored and managed in a centralized repository under the custody of the Butterfly Conservation Europe and the UK Center for Ecology and Hydrology.

3) Data Integration. There is not a perfect match between the products generated by this initiative (indicator trends at the national and European level) and the EBV definition (in terms of product and spatial resolution 10 x 10 km, 50 x 50 km) (**EBVm** bottleneck). The current software used for calculating butterfly trends, indexes and indicators is R (*rbms* package; non-user-friendly software; **SoFw** bottleneck). However, the code is openly shared and training is provided to users of eBMS data (workshop contents are available on the eBMS website). While this same package cannot be used to map species distributions, it could be expected that if this initiative were to derive spatially explicit estimates of species occurrences, it would do so by relying on the same software and open code. However, currently there is no open code for spatially-explicit distribution models in place within the framework of the eBMS program (**OpC** bottleneck). The eBMS has recently developed a mobile application (eBMS App) that allows volunteers to record species observations and abundances and directly upload them to the eBMS database (also to review the uploaded data). However, data collected (species lists, tables with data observations) also get reported via email to eBMS coordinators, so data streams are not fully automated (**Auto** partial bottleneck). Funding for coordination and integration tasks is not secured both at the European and national levels; for example, the last geographic expansion of the eBMS network was possible thanks to the ABLE project, funded by the European Union Directorate General for the Environment, for a period of two years from 2018-2020 (**Fnd** bottlenecks). Data is only available upon request (**OpDat** partial bottleneck).



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

The 2021 report on the design of an European Pollinator Monitoring scheme (EU PoMS) (Potts et al. 2021), when implemented in practice, will collect the data to generate this EBV since the abundance estimates, along with the location/spatial coordinates of systematic surveys will be recorded and made available for this purpose. The EU PoMS design has already identified current bottlenecks that could hinder the estimation of pollinator abundances across Europe which also apply to species distribution models.

These relate to:

1) Data collection & sampling. The EU PoMs proposal has identified more than 76 pollinator monitoring schemes already collecting pollinators' data across Europe; however, the variety of sampling methods used makes it difficult to combine the already existing data in these programs to generate pollinator indicators or to estimate abundance trends at the European level. Moreover, for some important pollinator taxonomic groups the data is limited (**TxC** partial bottleneck); this relates to the limited capacity building of some countries where there is lack of taxonomic resources and experts (**Cbui** and **GeC** partial bottlenecks). The EU PoMs proposal describes in detail how systematic surveys will be carried out for each taxonomic group (surveys are initially planned on an annual basis and assessments every three years). There are ongoing Preparatory Actions already working on overcoming these bottlenecks: the SPRING project seeks to strengthen taxonomic and citizen science capacity with regard to pollinating insects and the ORBIT and Taxo-FLY , seek to create a more centralized taxonomic EU facility for the identification of wild bees and to develop resources for European hoverfly inventory and taxonomy, respectively.

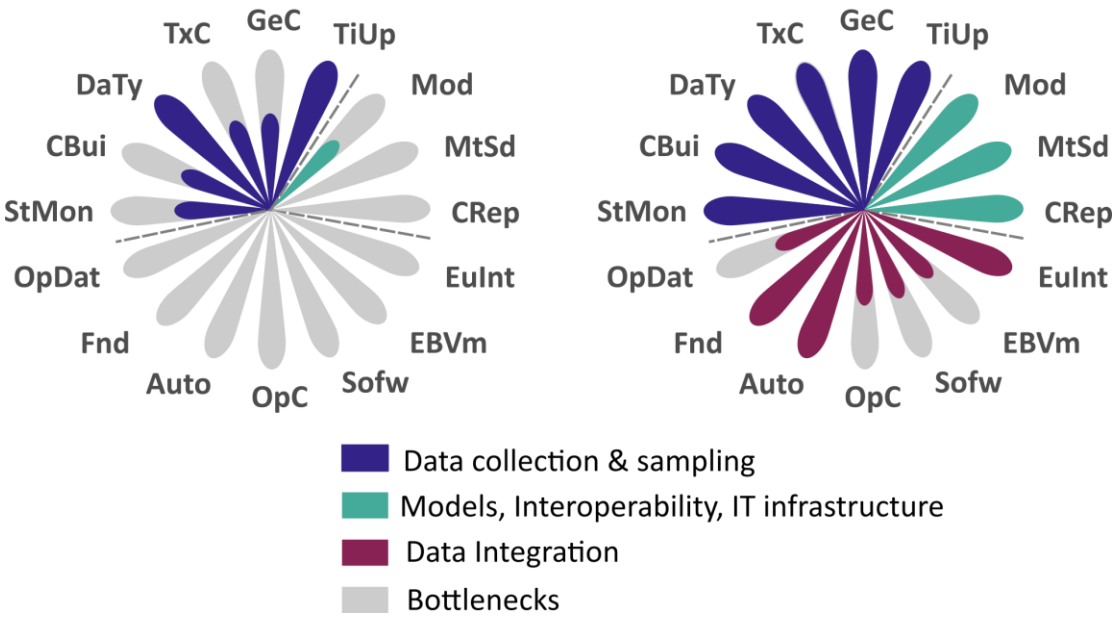
2) Models, interoperability, and IT infrastructure. Some tests have already been run on using models to develop species distributions from data collected across this monitoring network, including integrated distribution models accounting for occupancy and process-based models for predicting abundance and visitation (**Mod** partial bottleneck). The EU PoMs pilot proposal envisions data to be centralized in a repository at the EEA, European Commission (DG ENV), JRC or Eurostat. Data will be submitted or shared in a standardized form (following metadata standards) via an online platform to the European coordination facility, where pan-European analyses will be made. Currently, none of the latter two elements exist (**MtSd** and **CRep** bottlenecks), though they are specifically being co-developed between an expert working group, DG ENV, EEA and Member States with delivery due in late 2023.

3) Data integration. The EU PoMs proposal presents options to set up a Pan-European pollinator monitoring network, which does not exist so far (**EuInt** bottleneck). The proposal cites the European Butterfly Monitoring (eBMS) as the closest initiative to what the EU PoMs wants to set up at the European level. Initially, EU PoMs expects to be integrating data at the national level to estimate trends in pollinators abundances and taxonomic diversity and therefore there the integration product that will be generated by EU PoMS does not match well the definition of this EBV (**EBVm** bottleneck): however, the abundance estimates collected along with the location/spatial coordinates of systematic surveys could be made available for making spatially explicit predictions of pollinators distributions. There are not yet clear guidelines about whether the code used for data integration and potentially modelling will be openly shared (although tests have been run with R, which is not a user-friendly software; **OpC** and **Sofw** bottlenecks). The Eu PoMS will put in place

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

metadata standards to facilitate data integration and there is a plan to make the data openly available upon request, following the eBMS model (**OpDat** partial bottleneck). One of the actions proposed to automate data flows is the development of a pan-European internet identification platform for pollinators, which is constantly maintained and updated. The monitoring of pollinator populations by EU PoMS will prove key to tracking the goals set in the proposed Nature Restoration Law of halting the decline of pollinator populations by 2030 and achieve thereafter an increasing trend of pollinator populations. This regulation (Nature Restoration law) will impose an obligation to monitor pollinator communities and therefore member states will have to ensure that sufficient funding is allocated for this purpose (therefore the change in **Fnd** bottleneck between the two Eu PoMS figures below).

a) Eu PoMS (Potts et al. 2021) **b) Eu PoMS (set 2026)**



Besides pollinators and butterflies, there are no other European-level integrated standardized monitoring initiatives targeting priority invertebrates. There are examples of monitoring carried out by research networks, such as for example, the European Swag Beetle monitoring network . This network was set up in 2008 by researchers from 8 countries and currently collects citizen science observations of the European stag beetle (*Lucanus cervus*) across 14 European countries.

References:

Potts, S.G., Dauber, J., Hochkirch, A., Oteman, B., Roy, D.B., Ahrné, K., Biesmeijer, K., Breeze, T.D., Carvell, C., Ferreira, C., FitzPatrick, Ú., Isaac, N.J.B., Kuussaari, M., Ljubomirov, T., Maes, J., Ngo, H., Pardo, A., Polce, C., Quaranta, M., Settele, J., Sorg, M., Stefanescu, C. and Vujić, A. (2021) Proposal for an EU Pollinator Monitoring Scheme, EUR 30416 EN, Publications Office of the European Union, Ispra, 2021, ISBN 978-92-76-23859-1, doi:10.2760/881843, JRC122225.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

EBV: Species distributions terrestrial plants	
ID	51
Realm	Terrestrial
EBV class	Species Populations
EBV name	Species distributions of terrestrial plants
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of terrestrial vascular plant species within contiguous spatial units (grid cells) across the EU over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	- Priority species: 1 × 1 km – 10 × 10 km - All vascular plant species: 10 × 10 km – 50 × 50 km
Temporal resolution unit	- Priority species: 1 year - All vascular plant species: 3 or 6 years
Taxonomic/ ecosystem focus group	All European terrestrial vascular plants species included in the European Red List Priority terrestrial vascular plants as listed in Annex II and Annex IV of the Habitats Directive

Main European initiatives and description of current bottlenecks

The European Vegetation Archive (EVA) is an integrative database of vegetation plots across Europe. The purpose of EVA is to establish and maintain a single data repository of vegetation-plot observations (i.e. records of plant taxon co-occurrence at particular sites, also called phytosociological relevés) from Europe and adjacent areas and to facilitate the use of these data for non-commercial purposes, mainly academic research and applications in nature conservation and ecological restoration. The EVA is an initiative of the Working Group European Vegetation Survey (EVS) of the International Association for Vegetation Science (IAVS), and it is coordinated by a board of members distributed across different European institutions that gets renewed every 4 years. By April 2021, EVA comprised 99 national and supranational vegetation plots databases and contains 1,804,985 vegetation plots from 53 countries. At the end of 2021 the EVA launched ReSurveyEurope, an initiative that seeks to mobilize vegetation-plot resurvey data with repeated measurements over time and establish a collaborative initiative as a basis for nuanced and robust assessment of biodiversity trends on small spatial grains over longer periods in Europe.

Possible bottlenecks in the generation of this EBV from existing data collected by the EVA data relate to:

1) Data collection and sampling: The data gathered in ResurveyEurope comes from monitoring that follows systematic sampling protocols (data from the general EVA database also follows systematic sampling protocols but plots are sampled just once and therefore, data could somehow be considered opportunistic observations). Despite the large geographic coverage of the EVA data (53 countries), and given the enormous diversity of this taxa, it is probable not all terrestrial vascular species targeted in this EBV are covered by the vegetation plots included in the database (e.g tree species are better cover by other initiatives such as the National Forest Inventories - see EBV "Species distributions of trees") (**TxC** bottleneck). Moreover,

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

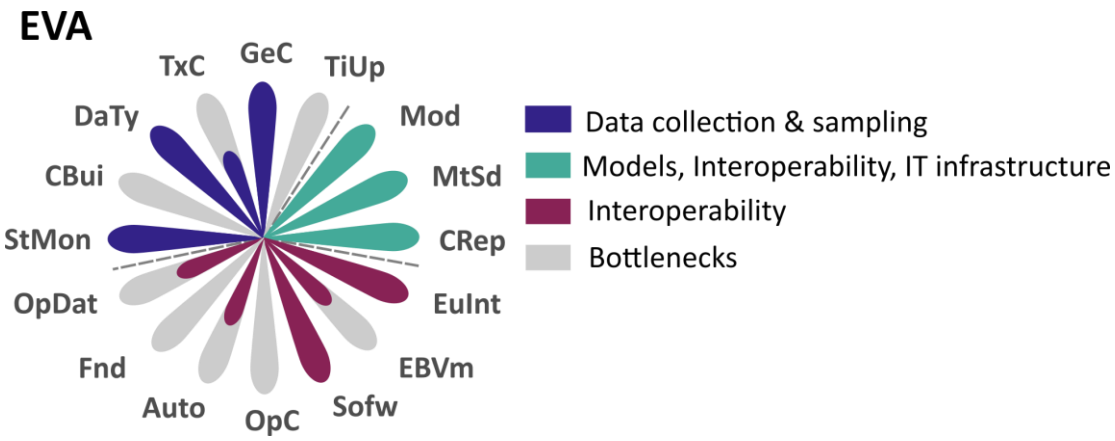
the database is mostly based on single surveys conducted over the last decades in Europe and therefore represents a snapshot of abundance of vascular plants in vegetation plots so it cannot be used to generate this EBV at the desired temporal resolution (**TiUp** bottleneck). The ReSurveryEurope initiative seeks to overcome this bottleneck, by compiling temporal series data that will allow to estimate changes in distribution and abundance of vascular plants over time (it includes plots, transects or relevès that have at least two repeated measures using the same of comparable sampling methods); however, the spatial and taxonomic coverage of ReSurveyEurope is smaller than that of the generalist EVA database. The ReSurveyEurope initiative will seek to overcome the current bottleneck in capacity building that EVA has because of its opportunistic character (**Cbui** bottleneck): it will promote the generation of time series data from EVA sampling plots in areas where they have been sampled only once.

2) Models, interoperability, and IT infrastructure. There are not bottlenecks to report in this regard: The EVA initiative has already published a series of reports commissioned by the European Environmental Agency where vegetation plots data is used to generate spatially-explicit predictions of phytosociological relevés across Europe using the Maxent model, as a measure of ecosystems distributions (Schaminée et al. 2014, 2016a,b): this method could perfectly be also applied to model individual single species (see more details of the modelling approach used below under “data integration”). To set up the EVA a software platform was developed (TurboVeg3) to facilitate data harmonization across the multiple databases on vegetation plot data collected across Europe (the software defines metadata standards for harmonization). The EVA database is centralized and curated by the EVA coordination board.

3) Data integration. There is a mismatch between the products generated by EVA and the description of this EBV, especially in terms of temporal resolution (**EBVm** partial bottleneck). The model used by EVA to model ecosystems distributions was a machine-learning presence-background model Maxent (Phillips et al. 2006, 2008). Maxent has a user-friendly interface facilitating its use by non-modellers. Code of the models has not been made publicly available (**OpC** bottleneck). Data flows are only partially automated: contributing national and subnational data bases have to upload their data using the TurboVeg3 that requires filling different tables and forms; there is no reference to the use of Apps or other software to automatically transfer data from the field to the EVA (**Auto** partial bottleneck). EVA data management has been partly funded by the Czech Science Foundation, and also partially by European-funded projects involving the staff of Masaryk University and the program developer S. Hennekens (ALTERRA, Wageningen, UR). No other information seems to be available related to current funding, but the project is ongoing and new initiatives are developed within the framework of EVA (e.g. the ReSurveyEurope was launched in 2022); however, funding for contributing parties is not warranted (**Fnd** bottleneck). The data in the EVA database is not fully open: at the time of data submission or update, the custodians assign one of the following data availability regimes to the data contributed by them, either for the whole database or its individual subsets: 1) Restricted-access data are available for data contributors only; 2) Semi-restricted-access data are available for data contributors only and 3) Free-access data are available to a wider community of users. These data can be released based on the proposal to the EVA Coordinating Board, with no need for special approval. The EVA initiative encourages a gradual transfer of data

This project receives funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003553.

contributed to EVA from regime 1 to regime 3, but the decision on the regime is entirely upon the custodian (**OpDat** partial bottleneck).



The Atlas of Flora Europea (AFE) is a long-running long-term programme for mapping the current and past distribution of vascular plants in Europe. To date it has mapped the distribution of nearly 25% of the vascular plants of Europe at a 50 x 50 km resolution in a series of book volumes starting in 1972 (**TxC** bottleneck). All the distribution maps published in the AFE vols. 1–12 were made manually. Later these maps were scanned into a digital database. For the last volumes (vol. 13–14), the distribution data have been entered directly to the database. A simple software called AFE Editor (its updated version 2010 is available from the AFE tools page) has been made and distributed among the AFE collaborators and assistants. This program facilitates electronic data recording and subsequent entering to the database. Data is centralized by LUMOUS, the Finnish Museum of Natural History, although the database containing observations per 50 km grill cell is under publishers' copyright and is not freely available to the public and only to researchers on demand (**OpDat** partial bottleneck). The low resolution of raw data - 50 km (**DaTy** bottleneck), the lack of taxonomic completeness (**TxC**), the lack of systematic data collection protocols linked to it (**StMon**), but especially the lack of temporal replication (**TiUp** bottleneck**)**, make the data of this initiative lowly suitable to the generation of this EBV.

References:

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological modelling*, *190*(3-4), 231-259.

Phillips, S. J. and Dudík, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31(2), 161-175.

Schaminée J.H.J., et al. (2014) Vegetation analysis and distribution maps for EUNIS habitats. Report EEA/NSV/14/006. EEA, Copenhagen. URL.

Schaminée J.H.J., et al. (2016a) Review of grassland habitats and development of distribution maps of heathland, scrub and tundra habitats of EUNIS habitats classification. Report EEA/NSV/15/005. European Environment Agency, Copenhagen. URL.

Schaminée J.H.J, et al. (2016b) Development of distribution maps of grassland habitats of EUNIS habitat classification. Report EEA/NSV/16/005. EEA, Copenhagen.URL.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

EBV: Species distributions of main trees	
ID	52
Realm	Terrestrial
EBV class	Species populations
EBV name	Species distributions of main trees
Step in identification process	Expert workshop
Definition	The presence/absence or probability of occurrence of European tree species within contiguous spatial units (grid cells) across the EU over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	10 × 10 km – 50 × 50 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Tree species included in the EU-Trees4F dataset (67 species)

Main European initiatives and description of current bottlenecks

The European National Forest Inventory Network (ENFIN) is a pan-European initiative set up in 2003 with the goal of harmonizing data from National Forest Inventories (NFI) across Europe (a network of NFI organizations). ENFIN envisions a modern and responsive European system for forest monitoring that makes the best use of both field-based sample plots and remote sensing products through their appropriate combination. The ENFIN is custodian of the most comprehensive data set on *in situ* tree species occurrences in Europe. It provides support and data to other European integration initiatives such as the European Forest Data Centre and more recently to the Forest Information System For Europe (FISE). The data harmonized by ENFIN has served as a basis to map the distribution of > 250 tress species across Europe, some of them collected in the European Atlas of Forest Tree species.

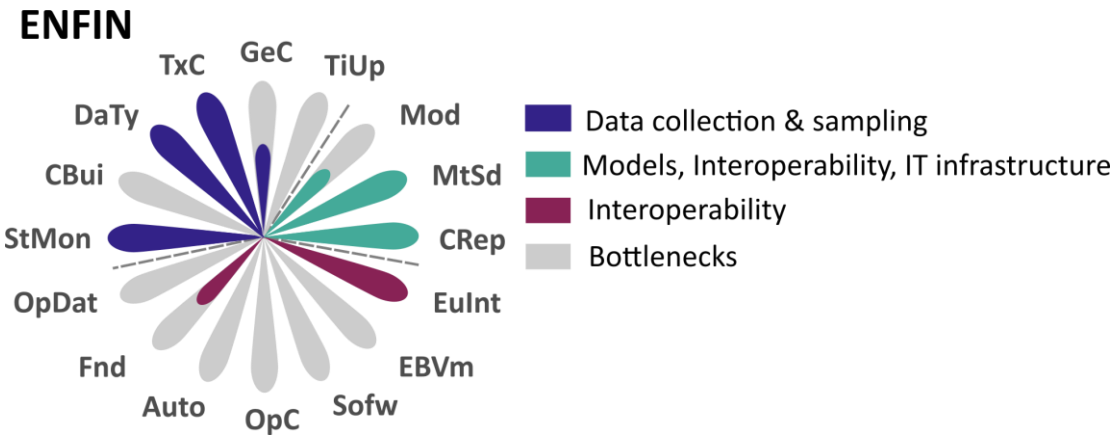
Possible bottlenecks in the generation of this EBV from the ENFIN data relate to:

1) Data collection and sampling. The collection of data in NFI follows standardized monitoring protocols. The data in EU-NFI are registered in a 1 km geographical grid that does not retain the precise location of the NFI plots but can potentially be used to generate species distribution maps across the continent. ENFIN does not have capacity building on its own; this depends on the capacity building of members of the network (**CBui** bottleneck). Currently, the ENFIN group is composed of 33 different organizations from 30 countries but the UK for example is not included (**GeC** partial bottleneck). While most common tree species are monitored by all the European NFIs, information on secondary species or taxa growing only in part of the continent may be missing from some NFIs; this is often also the case for smaller trees, alien and rare species, and species for agroforestry or short-rotation forestry. However, this EBV focuses on a small set of tree species, which data has already been harmonized and made public by members of ENFIN (Mauri et al. 2017). Moreover, the species sampled differ among countries making absences highly uncertain for some species in some countries. The frequency of data collection under NFI (time series) varies across countries but it is generally between 6 - 10 years, so the temporal resolution is smaller than the one desired for the generation of this EBV (**TiUp** bottleneck).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

2) Models, interoperability, and IT infrastructure. Data from ENFIN has been integrated with other datasets to model and map the distribution of trees across Europe but this is not done routinely within the framework of the ENFIN initiative (**Mod** partial bottleneck). The ENFIN has defined MetaData Standards to harmonize data across the different National Forest Inventories (not all countries collected the same parameters in their NFI). The basic prerequisite for the harmonization process is the common agreement on Reference Definitions that ENFIN has set.

3) Data integration. There is a mismatch between the integration achieved by ENFIN (integration of NFI databases across Europe) and the description of this EBV, especially in terms of temporal resolution (**EBVm** bottleneck). Since this integration initiative does not explicitly model species distributions over space, there is no open code or user-friendly software to report in this regard (**Sofw** and **OpC** bottlenecks). The funding of the ENFINprogram in the last few years has been secured through different European projects and framework contracts with the Joint Research Centre of the European Commision (**Fnd** partial bottleneck).; however, the funding for ENFINis not warranted in the long-term a, neither the funding of the underlying NFI depends on Member states and data partners. Data gathered at ENFIN is not available (**OpDat** bottleneck) but National Forest Inventory data is open or available upon request in many European countries contributing to ENFIN.



The International Co-operative Program on Assessment and Monitoring of Air Pollution Effects on Forests (ICP - Forest) has an established network of sample points that follow harmonized and standardized survey methods and an online platform for forest data storage and exchange.
Collected data in ICP-Forest database are designed in two monitoring levels:

- Level I for large scale monitoring of tree crown condition on a 16 x 16 km grid throughout Europe (approxomately 600 plots) ;
- Level II for intensive monitoring of around 800 plots in selected forest ecosystems with the aim of understanding the cause-effect relationships between natural stress factors (in particular, air pollution) and forest conditions.

The resulting geo-database covers 42 countries and includes165 tree species, through more than 18,000 geo-located sample plots (active and historical).
While ICP covers major forest ecosystems in Europe, its focus is more on processes than patterns and therefore it measures parameters such as tree crown condition, foliar

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

chemistry, tree growth, soil chemistry, etc. In this regard is more incomplete than the data available at ENFIN for the generation of this EBV.

References:

Mauri, A., Strona, G., and San-Miguel-Ayanz, J. (2017). EU-Forest, a high-resolution tree occurrence dataset for Europe. *Scientific data*, 4(1), 1-8.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Species distributions of invasive alien terrestrial taxa of European concern	
ID	54
Realm	Terrestrial
EBV class	Species populations
EBV name	Species distributions of invasive alien terrestrial taxa of European concern
Step in identification process	User & Policy Needs Assessment
Definition	The presence/absence or probability of occurrence of invasive terrestrial species within contiguous spatial units (grid cells) across the EU over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	1 x 1 km - 10 × 10 km
Temporal resolution unit	3 or 6 years
Taxonomic/ ecosystem focus group	Species specified in the List of Invasive Alien Species of Union Concern (88 species)

Main European initiatives and description of current bottlenecks

The European Alien Species Information Network (EASIN) is an integration initiative coordinated at the European level (Joint Research Centre, European Commission) that aggregates, integrates, and harmonizes spatio-temporal data for alien species (AS) and IAS across Europe. The species catalogue and geodatabase are regularly updated following the continuous revision of species reports in the literature, data published by EASIN data partners and the official reports by Member States competent authorities. The EASIN catalogue (v9.0- 19.07.22) includes 10,169 terrestrial alien species, of which 61 are IAS UC (EU Regulation 1143/2014). EASIN facilitates the exploration of data and information from existing monitoring networks and programs available from a variety of distributed information sources by providing tools and interoperable web services, compliant with internationally recognized standards.

Possible bottlenecks in the generation of this EBV from existing data collected by EASIN data relate to:

1) Data collection and sampling. EASIN carries out the systematic collection of data on IAS in Europe, does not come exclusively from standardized monitoring programs but from a variety of sources, including literature review, occasional observations and data portals (**StyMon** partial bottleneck); EASIN has demonstrated capacity building through different activities including: i) ad hoc assessments, such as the Baseline Distribution of IAS of Union concern; ii) support to the establishment of surveillance systems in compliance with EU AS policy; iii) citizen science programs, (e.g. the development of the smartphone app 'IAS in Europe' freely available and already adopted in the context of several projects and for official monitoring; iv) training to teachers and students (e.g., the MOOC course "Have you seen an alien?"), among others. EASIN collaborates with the Member States competent authorities and national experts to update and validate the datasets relevant in the EU policy context, to fulfil the EU mandate of setting up national monitoring networks for IAS; however, the extent to which this has been implemented is not fully known. Some bottlenecks regarding data collection may relate to 1) lack of communication between national coordinators & EU funded projects working on IAS

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

in a country (e.g. LIFE projects); 2) at national level, people in charge of monitoring AS belong to different admin bodies among which there is little communication and coordination and 3) complex administrative structure in some countries (federal, regions, etc.) and how the environmental competences are distributed among them, making it difficult to standardize and coordinate data collection tasks and to facilitate data flows and integration. Currently, the EASIN AS Geo Database contains occurrence records for more than 14,000 species, across 40 different countries (EU & beyond) (including data for the 88 species in the consolidated list of IAS of Union concern- European Union (EU) Regulation 1143/2014). Because the data is georeferenced, it could potentially be used to build species distribution models. The EASIN GeoDatabase is updated on average 3 times/ year, but some information sources are updated less regularly depending on effort for the update and the periodicity of their own updates. Recent EASIN database backend developments data updating will be facilitate more regular updates.

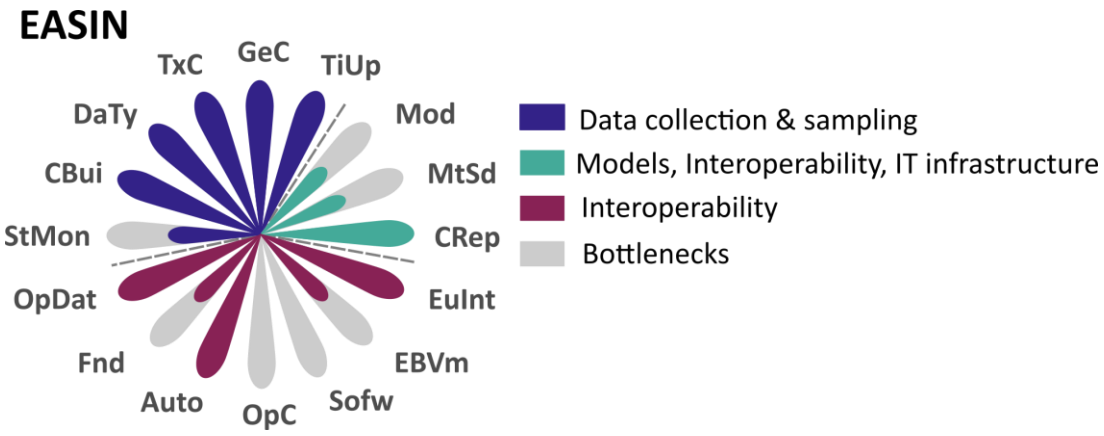
2) Models, interoperability, and IT infrastructure. Habitat suitability models have been already fit for the freshwater species *Elodea nuttallii* from EASIN data across EU (Steen et al. 2019); however, this is not a product regularly generated by this initiative, but a case study (**Mod** partial bottleneck), although there is ongoing work in this direction for IAS of Union concern: models are used to evaluate the range of expansion at both European and national levels, to understand whether niche of the species is at equilibrium in its invaded range or to predict the probability of occurrence of the species in Europe considering bioclimatic variables. The AS GeoDatabase (the central repository of EASIN curated by the Joint Research Centre of the European Commission) has developed a protocol to frequently retrieve data from the data partners; data are subsequently transformed by converting the harvested data to the EASIN Data Model through the following steps: validation, cleansing and standardization, geocoding, mapping, application of quality rules and finally loaded on the Geodatabase (Datawarehouse). However, while standards have been defined to harmonize data across source databases, these do not necessarily apply to the underlying data (**MtSd** partial bottleneck).

3) Data integration. The products generated by this initiative partially match the EBV, especially in terms of spatial resolution: the EASIN species mapping tool shows the distribution of species at the country level or at 10 x 10 km grid cells but not at finer spatial resolutions (**EBVm** partial bottleneck). The test run on modelling the distribution of *Elodea nuttalli* has used Maxent, a machine-learning presence-background model that has a user-friendly interface facilitating its use by non-modellers (Phillips et al. 2006, 2008). However, this model is not routinely used by EASIN to model the data in the Alien species geodatabase, so there is not user friendly software or open code to report in this regard (**Soft** and **OpC** bottlenecks). Data flows are automatized at different levels, from data collection to data integration: the EASIN has developed the "IAS Europe" smartphone App to promote the report of sightings of IAS of Union Concern by citizens and their integration into the EASIN GeoDatabase (and its consequent harmonization with other data retrieved from Data Partners). The process of retrieving the data from the Data Partners is done through the EASIN Data Broker system, which can retrieve the species occurrences and related information (date, source) from different kinds of data sources and store them in a normalized database structure. The EASIN has also developed and manages the NOTSYS platform as the official tool for EU Member States to notify the Commission and inform the other Member States as required by Reg. 1143/2014 on Invasive Alien Species (IAS). In particular, the tool has been

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

designed to facilitate a timely comprehensive notification of new detections of IAS of Union concern and related eradication measures. Funding is a key bottleneck to the generation of this EBV: while the European Commission has set up EASIN as the official information system supporting Member States in the implementation of the Regulation (EU) No 1143/2014 , it does not directly co-participate in the expenses of either monitoring or eradication actions by national nodes which solely relies on national budgets; funding from the European Commission for monitoring of IAS is available via existing financial instruments such as LIFE, H2020, Cohesion or Regional Development funds; however, funding is warrant in the mid - long - term to maintain the EASIN data infrastructure and the development of new data collection tools (e.g., a web base platform for reporting citizen science data) (**Fnd** partial bottleneck). Data in the EASIN GeoDatabase can be easily accessed and downloaded from the website. The openness of the data is key to redirect surveillance and trigger early warning systems.



Monitoring data of invasive species to generate this EBV can also be retrieved from other initiatives described in detail in other EBVs , for example the EBBA 2(birds), the NA2RE (amphibians and reptiles), EMMA2 (mammals) or the EVA (plants).

References:

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological modelling*, *190*(3-4), 231-259.

Phillips, S. J. and Dudík, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31(2), 161-175.

Steen, B, Cristina Cardoso, A.C., Tsiamis, K., Nieto, K.,Engel J., Gervasini, E. (2019) Modelling hot spot areas for the invasive alien plant Elodea nuttallii in the EU. *Management of Biological Invasions* 10(1): 151–170. URL .

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Terrestrial species traits

EBV: Phenology of migration of terrestrial birds	
ID	58a
Realm	Terrestrial
EBV class	Species traits
EBV name	Phenology of migration of terrestrial birds
Step in identification process	Expert workshop
Definition	The annual timing of arrival and departure of European terrestrial migratory bird species at breeding, staging and wintering sites over time.
Metric	Migration phenology metrics such as: - Day of arrival - Day of departure - Length of stay
Spatial resolution unit	10 × 10 km
Temporal resolution unit	1 week
Taxonomic/ ecosystem focus group	Migratory bird species defined as full migrants in the European Red List

Main European initiatives and description of current bottlenecks

There are two European-wide coordinated initiatives that collect data that could potentially be used to generate this EBV: The EuroBirdPortal (EBP) and the EURING.

The **EuroBirdPortal (EBP)** combines the data collected by the online bird recording portals operating in Europe to describe the large-scale spatiotemporal patterns of bird distributions and their changes over time. Online bird portals obtain year-round data from the intensive and widespread activities of birdwatchers. However, data is, essentially, collected using simple standardized protocols (complete lists) or no protocol at all (casual observations). As the EBP collects data from the entire year, the project has the potential to obtain at least some of the metrics required for this EBV (e.g. arrival and departure timing) for several migratory bird species.

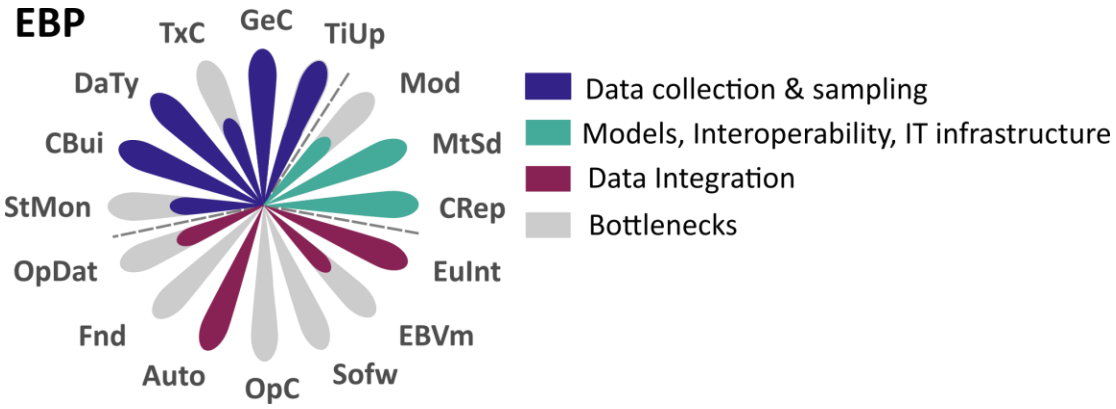
Possible bottlenecks in the generation of this EBV from existing data collected by the EBP initiative relate to:

1) Data collection & sampling. The EBP has the potential to collect data on all bird species occurring in Europe but, currently, only the data for 137 species are stored in the EBP central data repository (**TxC** partial bottleneck). Data is collected across all European countries but the degree of coverage is poorer in the South and, particularly, in the East and South-East. Data (species observations & counts) are gathered from online bird recording portals (e.g., ornitho) and include, essentially, both casual observations and data collected following simple systematic protocols (species lists). Only a small part of the data is collected following standardized monitoring protocols (**StMon** partial bottleneck). Training and capacity building for the network is provided during regular annual meetings or direct support to specific partners. Data is updated on a daily basis with information up to the previous day n (species lists).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

2) Models, data interoperability and IT infrastructure. Despite the EBP/online bird portals data have been used to obtain some of the kind of metrics required to generate this EBV in some specific contexts (e.g. to predict the start of spring migration in the context of delimiting hunting seasons or study changes in migratory phenology; see, for example, Newton et al. 2016), the modelling of phenological traits has not been generalized nor automated within the framework of this initiative (**Mod** partial bottleneck). The EBP follows metadata standards to ensure data harmonization among its different data sources. The EBP data is harmonized, managed and stored in a central repository curated by the European Bird Census Council.

3) Data Integration. The maps featured in the EBP viewer (www.eurobordportal.org) represent aggregated observations at the weekly basis and at 30 x 30 km resolution (**EBVm** partial bottleneck); however, in the EBP central data repository the casual data are aggregated at 10 x 10 km and date and the lists are in raw format (i.e. not aggregated) with date/timing and location given as precise location or at 10 x 10 km; therefore, the data available at the EBP repository would allow the generation of this EBV at the desired spatio-temporal resolution. Although this integration initiative can produce metrics of the kind required for this EBV, there is no open code nor user-friendly software yet to report in this regard as models are not routinely used to generate phenology products from EPB data yet (**Sofw** and **OpC** bottlenecks). Data streams are automated: the local online portals collect most of their data through mobile apps in near-real time or shortly after it has been recorded in the field; data collected in the data portals are then automatically transferred to the EBP daily. The EBP main developments have been possible thanks to the support of the LIFE programme (a new LIFE project proposal has been recently submitted), but funding stability in the mid- long- term is not warranted (**Fnd** bottleneck). EBP data is available upon request and subject to agreement by National coordinators who hold the ownership of data (**OpDat** partial bottleneck), but just one centralized data request should be done as data is already centralized in the EBP databank (authorizations by national owners are coordinated by EBP).



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EURING coordinates bird ringing in Europe and centralizes the data collected by the different bird ringing schemes operating in the continent in order to promote and encourage its use for bird management and conservation. Ringing data are collected year-round following standardized protocols (e.g. EUROCES Constant Effort Sites) or no protocol at all. The data consist of bird ringings (when a ring was first added to a bird), recaptures (retraps of ringed birds by ringers) and recoveries/resightings (ringed bird reported by the public (e.g. dead birds, rings/marks read at a distance) which form the bulk of the data currently stored in the EURING databank (EDB). The recently launched Eurasian African Migration Atlas (Spina et al. 2022) highlights the value of the EURING data to understand bird movements in time and space while the new Migration Mapping Tool is a good example of the applied value of combining the connectivity information of EURING with the observational data from the EuroBirdPortal in single tool (Gargallo et al. 2022).

The data collected by EURING and the bird ringing schemes in general have the potential to obtain the metrics required for this EBV and, more relevantly, are probably the best suited to estimate length of stay (e.g. though mark-recapture analysis).

Possible bottlenecks in the generation of this EBV from existing data collected by EURING relate to:

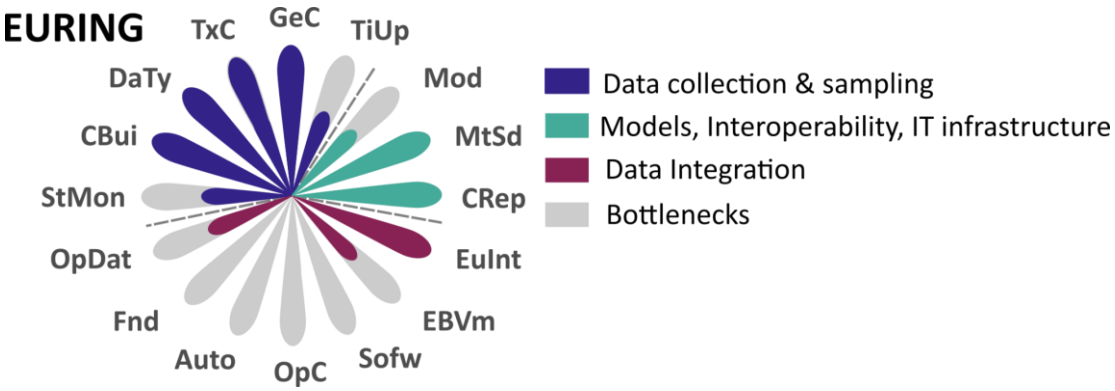
4) Data collection & sampling. EURING collects data of most bird species occurring in Europe and, though the amount of data (particularly recaptures/recoveries) varies greatly among species, figures are quite high for several of them. Data is collected across all European countries, but the degree of coverage is poorer in the South and, particularly, in the East and South-East. Only part of the data is collected following standardized monitoring protocols (**StMon** partial bottleneck). Training and capacity building for the network is provided during regular annual meetings or direct support to specific partners. Data in the EURINGdataback (EDB) are mostly updated once every year (**TiUp** partial bottleneck), although each record retains its temporal resolution (the date of data collection).

5) Models, data interoperability and IT infrastructure. Ringing data has been widely used to study phenology and length of stay and to obtain some of the metrics homologous to those required by this EBV for some specific uses, for example. to estimate the start of spring migration to delimit hunting seasons (e.g. the Migration seasons of hunted species research module of the Eurasian African Migration Atlas)(see also Ambrosini et al. 2014); however, so far, this has not been generalized nor automated within the framework of the initiative (**Mod** partial bottleneck). The EURING data flow takes place using the EURING Exchange Code standard, a pioneer of this kind among biodiversity data hubs in Europe, to ensure data harmonization and optimize its value. Data is centralized in the EURING databank.

6) Data Integration. EURING has already generated products that partially match the definition of this EBV (the Migration seasons of hunted species research module of the Eurasian African Migration Atlas, with predictions of prenuptial migration of different species at 10-day intervals - **EBVm** partial bottleneck). The data in the EDB have the required spatial and temporal resolution needed for the generation of this EBV. However, as the data is mostly updated once every year and therefore, this EBV could only be generated, at most, on a yearly basis. Though this integration initiative can produce metrics of the kind required for this EBV, there is no open code nor user-friendly software yet to report in this regard, as this modelling is not routinely integrated in this initiative (**Sofw** and **OpC** bottlenecks). Data streams are not

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

automated yet: ringing centres across Europe send their ringing observations (data exports) once a year to the central repository manually (**Auto** bottleneck). Funding stability in the mid- long- term is not warranted (**Fnd** bottleneck). EDB data is available upon request and subject to agreement by National bird ringing schemes who hold the ownership of data (**OpDat** partial bottleneck), but just one centralized data request should be done as data is already centralized in the EDB (authorizations by national owners are coordinated by EBP).



References:

Ambrosini, R., Borgoni, R., Rubolini, ... andSaino, N. (2014). Modelling the progression of bird migration with conditional autoregressive models applied to ringing data. *PloS one*, 9(7), e102440.

Gargallo, G., Davies, J.G., Faverjon, C., Kampichler C., Baillie, S.R., Cameron, A., Robinson, R.A. and Sierdsema, H. (2022) Development of a prototype early warning system for avian influenza in the EU based on risk-mapping. EFSA supporting publication 2022: EN-7762, 46 pp. URL.

Newson, S. E., Moran, N. J., Musgrove, A. J., Pearce-Higgins, J. W., Gillings, S., Atkinson, P. W., ... and Baillie, S. R. (2016). Long-term change in spring and autumn migration phenology of common migrant breeding birds in Britain: results from large-scale citizen science bird recording schemes. *Ibis (Lond. 1859)*, *158*, 481-495.

Spina, F., Baillie, S.R., Bairlein, F, Fiedler, W. and Thorup, K. (Eds) (2022) *The Eurasian African Bird Migration Atlas*. https://migrationatlas.org. EURING/CMS.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Phenology of the emergence of butterflies	
ID	58b
Realm	Terrestrial
EBV class	Species traits
EBV name	Phenology of the emergence of butterflies
Step in identification process	Expert workshop
Definition	The annual timing of seasonal emergence of butterflies within contiguous spatial units (grid cells) across the EU over time.
Metric	The day after which 5% of individuals have emerged
Spatial resolution unit	10 × 10 km
Temporal resolution unit	1 week (traits derived from weekly distribution data)
Taxonomic/ ecosystem focus group	Priority butterfly species listed in the Annex II and Annex IV of the Habitats Directive

Main European initiatives and description of current bottlenecks

The European Butterfly Monitoring Scheme (eBMS) collects abundance data for > 312 butterfly and moth species (and a few bumblebees and dragonflies) using systematic transect counts that are visited several times per year. The abundance data collected (counts per transect) gets mostly integrated to estimate population trends at the national and European levels (six indicators: grassland, forests, N2000, climate change, widespread species and urban butterflies).

Possible bottlenecks in the generation of this EBV from existing data collected by the eBMS data relate to:

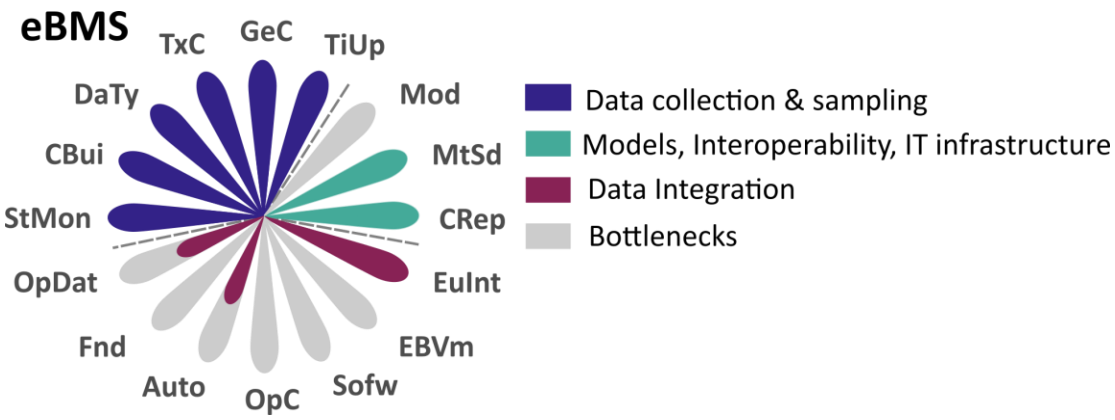
1) Data collection and sampling. There are no bottlenecks to report in this regard to the generation of this EBV. eBMS focuses its efforts mostly on the collection of butterflies and moths' abundance data. Data gets collected following standardized protocols with most sampling transects being visited at least 10 times per year. Each transect (approx. 1 km length) is ideally walked every week during the butterfly flight season. If it is not possible, volunteers are encouraged to count as often as possible, every two weeks or 10 days. Besides fixed transects, currently, the eBMS also collects one-off data (opportunistic 15 min-counts that seek to encourage data collection in areas currently uncovered by fixed transects, e.g. areas of difficult accessibility). The type of data collected (species counts), and the high frequency of data sampling (weekly temporal resolution in most transects) would allow the estimation of this EBV. The geographic coverage of this monitoring program has recently been enlarged thanks to the support of a service contract from the European Union Directorate General for the Environment (ABLE project), so it currently covers 30 countries. The program promotes capacity building through the free publication of divulgation informative material for the identification of butterflies in different languages and countries, the offering of open training courses on the use of modelling tools for species trend estimations, the development of Apps to ease data collection and identification, etc.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765

2) Models, interoperability, and IT infrastructure. This initiative uses the TRIM model to estimate population trends and calculation of trend indices/products at the national level on an annual basis (training courses and materials are available on the eBMS website); however, the modelling of phenological traits has not been generalized nor automated within the framework of this initiative (**Mod** bottleneck). The eBMS follows metadata standards that allow harmonization of the data collected across the different European countries. The eBMS database is stored and managed in a centralized repository under the custody of the Butterfly Conservation Europe and the UK Center for Ecology and Hydrology.

3) Data Integration. There is not a match between the products generated by this initiative (multi-species trend indicators; trends at the national and European level) and the EBV definition (date by which 5% of individuals have emerged) (**EBVm** bottleneck). Although this integration initiative can produce metrics of the kind required for this EBV, there is no open code nor user-friendly software yet to report in this regard as models are not routinely used to generate phenology products from eBMS data yet (**Sofw**, **OpC** bottlenecks). eBMS has recently developed a mobile application (eBMS App) that allows volunteers to record species observations and abundances and directly upload them to the eBMS database (also to review the uploaded data). However, data collected (species lists, tables with data observations) also get reported via email to eBMS coordinators, so data streams are not fully automated (**Auto** partial bottleneck). Funding for coordination and integration tasks is not secured both at the European and national levels; for example, the last geographic expansion of the eBMS network was possible thanks to the ABLE project, funded by the European Union Directorate General for the Environment, for a period of two years from 2018-2020 (**Fnd** bottlenecks). Data is only available upon request (**OpDat** partial bottleneck).



This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023.* DOI: https://doi.org/10.3897/arphapreprints.e103765

Terrestrial community composition

EBV: Community biomass of soil microbes	
ID	61
Realm	Terrestrial
EBV class	Community composition
EBV name	Community biomass of soil microbes
Step in identification process	Internal review process
Definition	Estimated biomass of the living component of soil organic matter (bacteria, fungi and protozoa) within contiguous spatial units (grid cells) across the EU over time.
Metric	- Mass of microbial carbon / mass of dry soil - Mass of microbial carbon / area
Spatial resolution unit	1 x 1 km
Temporal resolution unit	3 years
Taxonomic/ ecosystem focus group	Soil microbial species

Main European initiatives and description of current bottlenecks

The Land Use/Land Cover Area Frame Survey (LUCAS) collects information on land cover and land use change at the European scale. It is a project developed by EUROSTAT and it estimates the area occupied by different land use or land cover types based on observations taken at more than 250,000 sample points throughout the EU, which are visited on the ground and/or photo interpreted and classified over aerial photos and satellite images. This survey is repeated every 3 years, and it also collects a topsoil sample in a subset of all sampling points (18,000 – 26,000 points, depending on the year), which is analyzed for different chemical and biological properties. One of these properties is organic content and could potentially be used to generate this EBV.

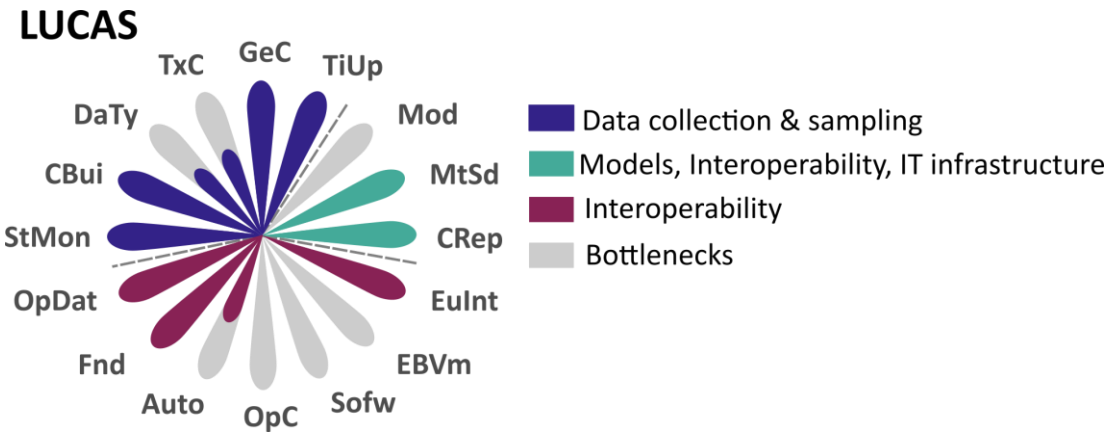
Possible bottlenecks in the generation of this EBV from existing data collected by the LUCAS data relate to:

1) Data collection & sampling. Soil data is collected following standardized protocols for a sample of approximately 0.5 kg of topsoil (0-20 cm) in each site. Data quality checks run in parallel to the collection of the data. The LUCAS initiative has capacity building and training, with surveyors receiving training before going into the field (a set of supporting documents, instructions on how to carry out the survey, and a set of quality control procedures). Since the first LUCAS survey (2009), organic content has been one of the measured properties of these samples. However, it does not differentiate between microbial carbon and general organic carbon, so it does not really match the data type required for generating this EBV (**DaTy** bottleneck). Additionally, although it is a very extensive survey, 26.000 points across Europe might not be enough to have every microbial species represented (**TxC** partial bottleneck). The LUCAS survey is meant to be carried out at three-year intervals, in synchronization with CORINE Land Cover and the update of the High-Resolution Layers, and as such, LUCAS data matches the temporal resolution sought by this EBV.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

2) Models, interoperability, and IT infrastructure. The data from the land cover and land use is corrected and run through various statistical models, but the topsoil data is not, only producing tables of organic content per location (**Mod** bottleneck). LUCAS follows metadata standards: surveyors use the same forms and instructions to integrate the data. Additionally, an ad-hoc IT tool, named Data Management Tool (DMT) is used to reinforce the standardization and integration of the LUCAS data into a central repository. The DMT records the data and analyses the quality of the recorded values through an automatic quality control.

3) Data integration. The LUCAS topsoil dataset and the derived products do not match well the definition of the EBV (overall organic content vs microbial organic content reported at the survey site level instead of in a 1 x 1 km continuous grid; **EBVm** bottleneck). Because there are no models in place for predicting microbial organic content in a spatially-explicit way within the framework of the LUCAS project, there is no user-friendly software or open code to report in this regard (**SoFw** and **OpC** bottlenecks). Besides LUCAS data flows and gets integrated using the Data Management Tool, data flows are not fully automated (e.g., lack of data collection via APIs, visual quality checks; **Auto** partial bottleneck). The funding for LUCAS survey is warranted in the long term (Copernicus services). The data are freely available and can be downloaded after prior registration through the Request Form.



Reference:

Eurostat. (2018). LUCAS - 2018 Quality Report. URL.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Community abundance and taxonomic diversity of pollinator insects	
ID	62
Realm	Terrestrial
EBV class	Community composition
EBV name	Community abundance and taxonomic diversity of pollinator insects
Step in identification process	User & Policy Needs Assessment
Definition	Total amount (abundance) of pollinator insects within spatial units over time.
Metric	Predicted number of individuals of pollinator insects
Spatial resolution unit	Small regions within countries based on nomenclature of territorial units for statistics (NUTS) from Eurostat (1166 regions at NUTS 3 level)
Temporal resolution unit	1 - 5 years (rotation across years)
Taxonomic/ ecosystem focus group	All pollinator insects as proposed in the species lists of butterflies, wild bees and hoverflies of the EU Pollinator Monitoring Scheme (EUPoMS)

Main European initiatives and description of current bottlenecks

The 2021 report on the design of an European Pollinator Monitoring scheme (EU PoMS) (Potts et al. 2021), when implemented in practice, will collect the data to generate this EBV since the abundance estimates, along with the location/spatial coordinates of systematic surveys will be recorded and made available for this purpose. The EU PoMS design has already identified current bottlenecks that could hinder the estimation of pollinator abundances across Europe. These relate to:

1) Data collection and sampling. The EU PoMs proposal has identified more than 76 pollinator monitoring schemes already collecting pollinators' data across Europe; however, the variety of sampling methods used makes it difficult to combine the already existing data in these programs to generate pollinator indicators or to estimate abundance trends at the European level. Moreover, for some important pollinator taxonomic groups the data is limited (**TxC** partial bottleneck); this relates to the limited capacity building of some countries where there is lack of taxonomic resources and experts (**Cbui** and **GeC** partial bottlenecks). The EU PoMs proposal describes in detail how systematic surveys will be carried out for each taxonomic group (surveys are initially planned on an annual basis and assessments every three years). There are ongoing Preparatory Actions already working on overcoming these bottlenecks: the SPRING project seeks to strengthen taxonomic and citizen science capacity with regard to pollinating insects and the ORBIT and Taxo-FLY seek to create a more centralized taxonomic EU facility for the identification of wild bees and to develop resources for European hoverfly inventory and taxonomy, respectively.

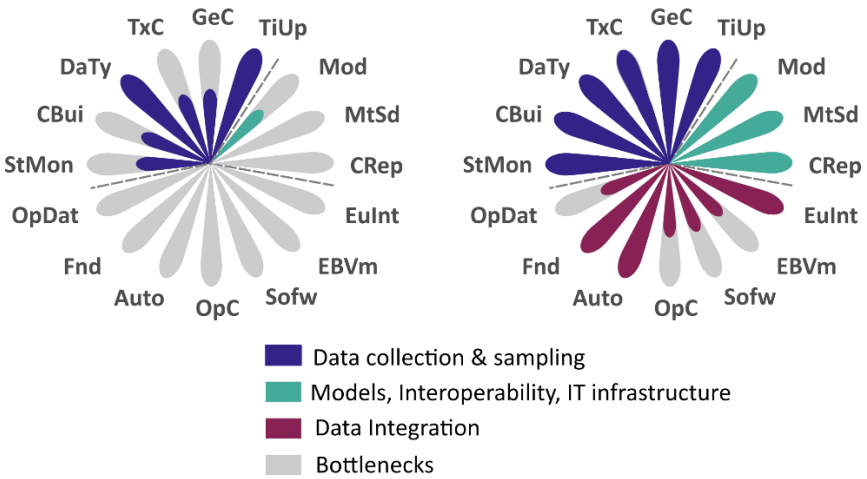
2) Models, interoperability, and IT infrastructure. The Eu PoMS initiative will aggregate community abundance and taxonomic diversity values at the national level. This will allow tracking progress on the target of reversing pollinators decline by 2030 in each Member State. The use of models to make spatially-explicit predictions of community abundance and taxonomic diversity across Europe is not directly

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

contemplated in the Eu PoMS proposal (**Mod** bottleneck) but this could be undertaken by organizations working with Eu PoMs data. The EU PoMs pilot proposal envisions data to be centralized in a repository at the EEA, European Commission (DG ENV), JRC or Eurostat. Data will be submitted or shared in a standardized form via an online platform (following metadata standards) to the European coordination facility, where pan-European analyses will be made. Currently, none of the latter two elements exist (**MtSd** and **CRep** bottlenecks), though they are specifically being co-developed between an expert working group, DG ENV, EEA and MS with delivery due in late 2023.

3) Data integration. The EuPoMs proposal presents options to set up a Pan-European pollinator monitoring network, which does not exist so far (**EuInt** bottleneck). The proposal cites the European Butterfly Monitoring (eBMS) as the closest initiative to what the EU PoMs wants to set up at the European level. The EU PoMS seeks to estimate overall abundances and trends of pollinator communities in each Member State and at the EU level. As such there is a good match between the main EU PoMS products and the definition of this EBV but a mismatch in the temporal resolution (Member State and European level vs subnational level; **EBVm** partial bottleneck). There are not yet clear guidelines about whether the code used for data integration and potentially modelling will be openly shared (although tests have been run with R, which is not a user-friendly software; **OpC** and **Sofw** bottlenecks). It will put in place metadata standards to facilitate data integration and there is a plan to make the data openly available upon request, following the eBMS model (**OpData** partial bottleneck). One of the actions proposed to automate data flows is the development of a pan-European internet identification platform for pollinators, which is constantly maintained and updated. The monitoring of pollinator populations by EU PoMS will prove key to tracking the goals set in the proposed Nature Restoration Law of halting the decline of pollinator populations by 2030 and achieve thereafter an increasing trend of pollinator populations. This regulation (Nature Restoration law) will impose an obligation to monitor pollinator communities and therefore member states will have to ensure that sufficient funding is allocated for this purpose (therefore the change in **Fnd** bottleneck between the two Eu PoMS figures below).

a) Eu PoMS (Potts et al. 2021) **b) Eu PoMS (set 2026)**



References:

Potts, S.G., Dauber, J., Hochkirch, A., et al. (2021) Proposal for an EU Pollinator Monitoring Scheme, EUR 30416 EN, Publications Office of the European Union, Ispra, 2021, ISBN 978-92-76-23859-1, doi:10.2760/881843, JRC122225.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Aerial biomass of migrating birds, bats and insects	
ID	63
Realm	Terrestrial
EBV class	Community composition
EBV name	Aerial biomass of migrating birds, bats and insects
Step in identification process	User & Policy Needs Assessment
Definition	Biomass flows of aerial migrants (birds, insects and bats) across Europe within contiguous spatial units (grid cells) over time.
Metric	Summary statistics of migration densities of birds, insects and bats derived from vertical profile time series of weather radar data (e.g., hourly averages of bird density and speed)
Spatial resolution unit	1 x 1 km - 10 x 10 km
Temporal resolution unit	1 day
Taxonomic/ ecosystem focus group	All migratory bird, bat and insect species (by size class)
Main policy targets	Habitats Directive; Birds Directive
Main European initiatives and description of current bottlenecks	

The European Network for the Radar surveillance of Animal Movement (ENRAM) is a research network focused on the use of data from operational weather radars for monitoring, understanding, and predicting aerial biomass flows (birds, bats and insects). The ENRAM network was established thanks to the supporting funds of a COST Action (2013-2017) and some of its members are currently active via other research projects (e.g. GloBAM, funded by BiodivERsA with partners from across Switzerland, Belgium, Finland, the Netherlands, the UK and the USA). The ENRAM network signed a data license agreement with the Operational Programme for the Exchange of weather Radar Information in Europe (OPERA) to use the weather radar data for ecological research (e.g. estimating volume and timing of species migration). Weather radar data belongs to the meteorological national institutes but in Europe many countries are organized via the OPERA and send their data to the OPERA repository.

The main bottlenecks to the generation of this EBV from meteorological data collected in OPERA data and the ENRAM and GloBAM networks relate to:

1) Data collection & sampling. The management of weather radar data across Europe is operated by OPERA. OPERA collects *clean* and *unclean* polar volume data with a very high temporal frequency (every 5 - 15 minutes): the former serves to generate meteorological and hydrological products (e.g., precipitation forecasts) and the latter can be used to extract biological information (so in this regard, the data collected serves the generation of this EBV). The ENRAM and GLoBAM projects have used this unclean polar volume data to produce summary statistics and maps depicting distributions of migration intensity, flight direction, altitude, and ground speed of avian migrants at various spatial and temporal scales (the use of this data to extract information about insects or bats is still under development; **TxC** partial bottleneck). However, OPERA has recently changed its data exchange policies to prioritize data with meteorological applications only so the data flows of unclean polar volume data are no longer available for most countries via the central repository, jeopardizing the long-term maintenance of the biological applications of the radar network (Shamoun-

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

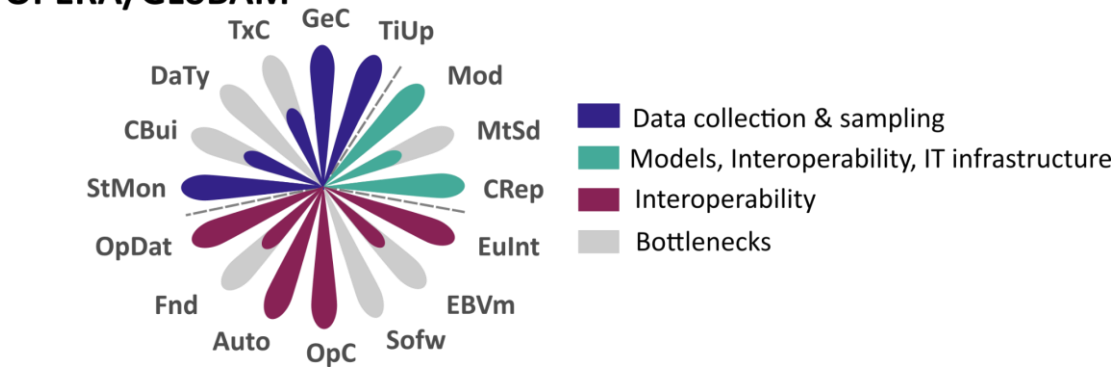
Baranes, J. et al. 2022) (**DaTy** bottleneck). These changes in data sharing policy by OPERA may limit the capacity building of ENRAM and GloBAM partners who had developed multiple initiatives since the setting of the ENRAM network (**CBui** partial bottleneck).

2) Models, interoperability, and IT infrastructure. Data from OPERA has been broadly used within the ENRAM and GLoBAM networks to model animal movement, estimate aerial biomass of birds and insects, forecast bird migration peaks, etc. OPERA centralizes data collected by national radar networks and standardizes it to develop, generate and distribute high-quality pan-European weather radar composite products on an operational basis. Not all data sent to OPERA is in the same format or the files are not always structured in the same way and thus there is need to do some extra work in order to process the weather radar data for biological products with current standard tools used by GloBAM (e.g. the R package *bioRAD;* Dokter et al. 2019); Similarly data is not always harmonized across radars, especially the uncleaned data so there is likely need to do more work to harmonize data across radars and countries to extract biological relevant data from radar images (**MtSd** partial bottleneck).

3) Data integration. The products generated by the ENRAM and GloBAM networks match well the definition of this EBV in terms of product type and temporal resolution, but a rasterized representation of biomass flows using weather radar data is currently available only for birds (**EBVm** partial bottleneck); moreover, currently the ENRAM/GloBAM summarize data around a buffer between 5 - 25 km from the radar or 5 - 40km from the radar. Yet with interpolation methods developed by Nussbaumer et al. (2019), it is possible to create products at a resolution of 10 km (value on the lower range of the spatial resolution sought for this EBV). To translate the radar data to bird metrics, one of the common tools used by ENRAM and GloBAM in Europe is to use the R package *bioRAD* (an R package; Dockter et al. 2019). In the following step, this data is used to develop models (for example predictive models of migration), which also often developed in R, which despite being open software, requires advanced technical knowledge for its use (**Sofw** bottleneck). One of the GloBAM project goals is to setup an automated data processing pipeline running on cloud infrastructure to retrieve biological information from European and US weather radar data as these become available, with the capacity to run over the long-term (i.e., to automatize the full data flow from retrieving radar data to providing estimates of the abundance of migrating species). While the funding for the OPERA infrastructure is secured in the long term given it is a meteorological facility, the funding for the GloBAM network is not secured in the long term and strongly dependent on the availability of research funding at the European scale (**Fnd** partial bottleneck). One of the goals of the GloBAM network is to make information on aerial migrations available as open data, supporting researchers in accessing, processing and analyzing weather radar data and vertical profiles of aerial migrants. The data on vertical profiles of birds from weather radar volume scans for over 100 radars (and from where aerial biomass can be inferred) are available via a GitHub repository. Some countries are now adopting open data policies and have their polar volume radar data now freely available (i.e., the raw weather radar data but not yet converted into biological data).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023.* DOI: https://doi.org/10.3897/arphapreprints.e103765



References:

Dokter, A. M., P. Desmet, J. H. Spaaks, S. van Hoey, L. Veen, L. Verlinden, C. Nilsson, G. Haase, H. Leijnse, A. Farnsworth, W. Bouten and J. Shamoun-Baranes (2019). "bioRad: biological analysis and visualization of weather radar data." *Ecography,* 42(5): 852-860.

Nussbaumer, R., Benoit, L., Mariethoz, G., Liechti, F., Bauer, S., & Schmid, B. (2019). A geostatistical approach to estimate high resolution nocturnal bird migration densities from a weather radar network. *Remote Sensing*, *11*(19), 2233.

Shamoun-Baranes, J., Bauer, S., Chapman, J. W., Desmet, P., Dokter, A. M., Farnsworth, A., ... & Leijnse, H. (2022). Meteorological data policies needed to support biodiversity monitoring with weather radar. *Bulletin of the American Meteorological Society*, 103(4), E1234-E1242.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EBV: Functional composition of soil biota	
ID	64
Realm	Terrestrial
EBV class	Community composition
EBV name	Functional composition of soil biota
Step in identification process	User & Policy Needs Assessment
Definition	The functional composition and diversity of soil biota based on morphological, physiological, phenological and behavioral traits or functional/taxonomic groups.
Metric	- Functional group richness - Functional diversity indices
Spatial resolution unit	1 x 1 m
Temporal resolution unit	1 year
Taxonomic/ ecosystem focus group	Bacteria, fungi, protozoa, collembola, mites, earthworms, larval and adult insects (e.g. Hymenoptera, Coleoptera and Diptera larvae), myriapods, spiders, mollusks and crustaceans

Main European initiatives and description of current bottlenecks

The Land Use/Land Cover Area Frame Survey (LUCAS) collects information on land cover and land use change at the European scale. It is a project developed by EUROSTAT and it estimates the area occupied by different land use or land cover types on the basis of observations taken at more than 250,000 sample points throughout the EU, which are visited on the ground and/or photo interpreted and classified over aerial photos and satellite images. This survey is repeated every 3 years, and it also collects a topsoil sample in a subset of all sampling points across the EU (18,000 – 26,000 points, depending on the year), which is analyzed for different chemical and biological properties. One of these properties, related to the EBV, is soil biodiversity through DNA sequentiation.

Possible bottlenecks in the generation of this EBV from existing data collected by the LUCAS topsoil data relate to:

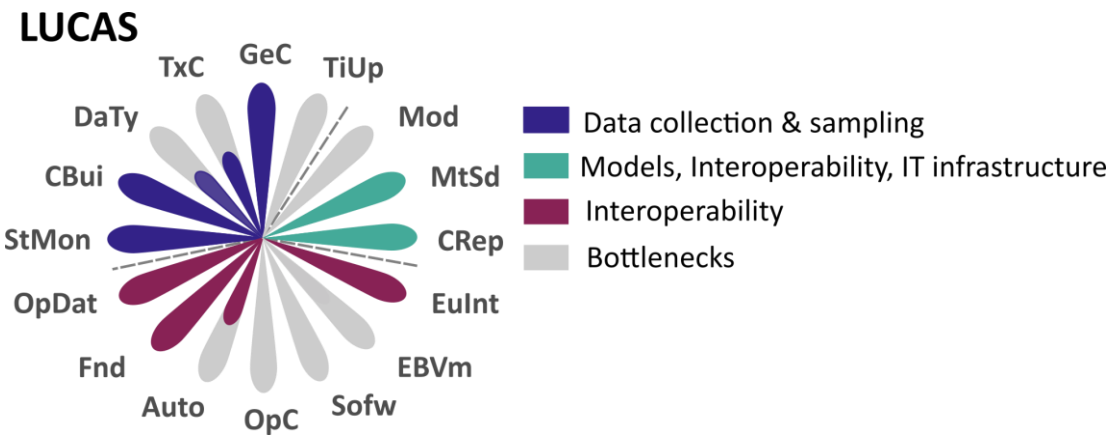
1) Data collection & sampling. Soil data is collected following standardized protocols for a sample of approximately 0.5 kg of topsoil (0-20 cm) in each site. Data quality checks run in parallel to the collection of the data. The LUCAS initiative has capacity building and training, with surveyors receiving training before going into the field (a set of supporting documents, instructions on how to carry out the survey, and a set of quality control procedures). Since the 2018 survey, DNA metabarcoding is used to analyse the soil diversity of the microbial community in a sample of 1,000 surveyed sites. This analysis targets the following attributes: Bacteria and Archaea (16S rDNA), Fungi (ITS), Eukaryotes (18S rDNA), Microfauna (nematodes), Mesofauna (arthropods), Macrofauna (earthworms), Metagenomics. Given the small sample of sites and the large diversity of microbes, it is hard to assume that every microbial taxon is represented in the analyses (**TxC** partial bottleneck); genetic analyses do not inform about abundance of each taxon (**DaTy** bottleneck). TheLUCAS survey is done

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

every 3 years, a smaller frequency than the temporal resolution sought to generate this EBV (**TiUp** bottleneck).

2) Models, interoperability, and IT infrastructure. LUCAS follows metadata standards: surveyors use the same forms and instructions to integrate the data. Additionally, an ad-hoc IT tool, named Data Management Tool (DMT) is used to reinforce the standardization and integration of the LUCAS data into a central repository. The DMT records the data and analyses the quality of the recorded values through an automatic quality control. The data from the land cover and land use is corrected and run through various statistical models, but models are not used to generate functional diversity indices (**Mod** bottleneck).

3) Data integration. The LUCAS integration initiative does not produce the functional diversity indices aimed at in this EBV and it does not even collect data on microbial diversity at the desired temporal resolution (**EBVm** bottleneck). Because there are no models in place for the generation of these indices, there is no user-friendly software or open code to report in this regard (**SoFw** and **OpC** bottlenecks). Besides LUCAS data flows and gets integrated using the Data Management Tool, data flows are not fully automated (e.g., lack of data collection via APIs, visual quality checks; **Auto** partial bottleneck). The data are freely available and can be downloaded after prior registration through the Request Form.



References:
Eurostat (2018). LUCAS - 2018 Quality Report. URL.
Joint Research Centre (JRC) (2018). LUCAS - Soil 2018. URL.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Terrestrial ecosystem structure

EBV: Ecosystem distribution of terrestrial EUNIS Habitats	
ID	66
Realm	Terrestrial
EBV class	Ecosystem structure
EBV name	Ecosystem distribution of terrestrial EUNIS Habitats
Step in identification process	User & Policy Needs Assessment
Definition	The geographical/spatial distribution of terrestrial EUNIS habitats within contiguous spatial units (grid cells) across the EU over time.
Metric	- Binary presence/absence - Probability of occurrence
Spatial resolution unit	10 x 10 km
Temporal resolution unit	1 year
Taxonomic/ ecosystem focus group	EUNIS terrestrial habitats (e.g. level 3 or 4)
Main European initiatives and description of current bottlenecks	

The European Vegetation Archive (EVA) is an integrative database of vegetation plots across Europe. The purpose of EVA is to establish and maintain a single data repository of vegetation-plot observations (i.e., records of plant taxon co-occurrence at particular sites, also called phytosociological relevés) from Europe and adjacent areas and to facilitate the use of these data for non-commercial purposes, mainly academic research and applications in nature conservation and ecological restoration. The EVA is an initiative of the Working Group European Vegetation Survey (EVS) of the International Association for Vegetation Science (IAVS), and it is coordinated by a board of members distributed across different European institutions that gets renewed every 4 years. By April 2021, EVA comprised 99 national and supranational vegetation plots databases and contains 1,804,985 vegetation plots from 53 countries. At the end of 2021 the EVA launched ReSurveyEurope, an initiative that seeks to mobilize vegetation-plot resurvey data with repeated measurements over time and establish a collaborative initiative as a basis for nuanced and robust assessment of biodiversity trends on small spatial grains over longer periods in Europe. Data from EVA has already been used to generate spatially-explicit predictions of distributions of a few EUNIS grassland, shrub and forest habitats across Europe (Schaminée et al. 2014, 2016a,b).

Possible bottlenecks in the generation of this EBV from existing data collected by the EVA data relate to:

1) Data collection and sampling: The data gathered in ResurveyEurope comes from monitoring that follows systematic sampling protocols (data from the general EVA database also follows systematic sampling protocols but plots are sampled just once and therefore, data could somehow be considered opportunistic observations). The geographic coverage of the EVA data is very broad (53 countries) and vegetation plots are sampled across a broad gradient of ecosystems in Europe so potentially the EVA/ReSurveyEurope databases contain enough information as to map the distribution of most if not all EUNIS terrestrial habitats. However, the database is mostly based on single surveys conducted over the last decades in Europe and

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

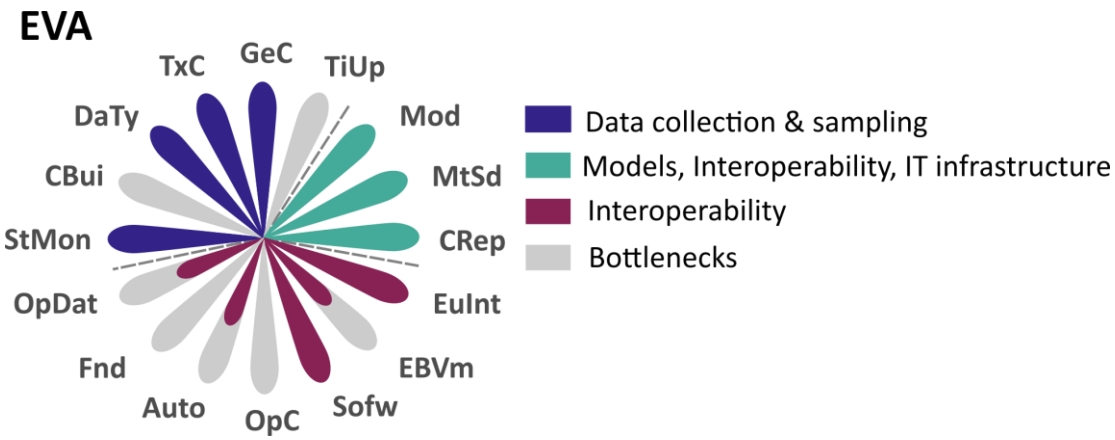
therefore represents a snapshot of abundance of vascular plants in vegetation plots so it cannot be used to generate this EBV at the desired temporal resolution (**TiUp** bottleneck). The ReSurveryEurope initiative seeks to overcome this bottleneck, by compiling temporal series data that will allow to estimate changes in distribution and abundance of vascular plants over time (it includes plots, transects or relevès that have at least two repeated measures using the same of comparable sampling methods); however, the spatial and taxonomic coverage of ReSurveyEurope is smaller than that of the generalist EVA database. The ReSurveyEurope initiative will seek to overcome the current bottleneck in capacity building that EVA has because of its opportunistic character (**Cbui** bottleneck): it will promote the generation of time series data from EVA sampling plots in areas where they have been sampled only once.

2) Models, interoperability, and IT infrastructure. There are not bottlenecks to report in this regard: The EVA initiative has already published a series of reports commissioned by the European Environmental Agency where vegetation plots data is used to generate spatially-explicit predictions of EUNIS habitats distributions across Europe using the Maxent model (Schaminée et al. 2014, 2016 a,b). To set up the EVA database a software platform was developed (TurboVeg3) to facilitate data harmonization across the multiple databases on vegetation plot data collected across Europe (the software defines metadata standards for harmonization). The EVA database is centralized and curated by the EVA coordination board.

3) Data integration. There is a partial mismatch between the products generated by EVA and the description of this EBV, especially in terms of temporal resolution (**EBVm** partial bottleneck). The model used by EVA to model EUNIS habitats' distributions was a machine-learning presence-background model Maxent (Phillips et al. 2006, 2008). Maxent has a user-friendly interface facilitating its use by non-modellers. Code of the models has not been made publicly available (**OpC** bottleneck). Data flows are only partially automated: contributing national and subnational data bases must upload their data using the TurboVeg3 that requires filling different tables and forms; there is no reference to the use of Apps or other software to automatically transfer data from the field to the EVA (**Auto** partial bottleneck). EVA data management has been partly funded by the Czech Science Foundation, and partially by European-funded projects involving the staff of Masaryk University and the program developer S. Hennekens (ALTERRA, Wageningen, UR). No other information seems to be available related to current funding, but the project is ongoing and new initiatives are developed within the framework of EVA (e.g., the ReSurveyEurope was launched in 2022); however, funding for contributing parties is not warranted (**Fnd** bottleneck). The data in the EVA database is not fully open: at the time of data submission or update, the custodians assign one of the following data availability regimes to the data contributed by them, either for the whole database or its individual subsets: 1) Restricted-access data are available for data contributors only; 2) Semi-restricted-access data are available for data contributors only and 3) Free-access data are available to a wider community of users. These data can be released based on the proposal to the EVA Coordinating Board, with no need for special approval. The EVA initiative encourages a gradual transfer of data contributed to EVA from regime 1 to regime 3, but the decision on the regime is entirely upon the custodian (**OpDat** partial bottleneck).

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

ARPHA Preprints *Author-formatted document posted on 22/03/2023*. DOI: https://doi.org/10.3897/arphapreprints.e103765



References:

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological modelling*, *190*(3-4), 231-259.

Phillips, S. J., Dudík, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31(2), 161-175.

Schaminée J.H.J., et al. (2014) Vegetation analysis and distribution maps for EUNIS habitats. Report EEA/NSV/14/006. EEA, Copenhagen. URL.

Schaminée J.H.J., et al. (2016 a) Review of grassland habitats and development of distribution maps of heathland, scrub and tundra habitats of EUNIS habitats classification. Report EEA/NSV/15/005. European Environment Agency, Copenhagen. URL.

Schaminée J.H.J, et al. (2016 b) Development of distribution maps of grassland habitats of EUNIS habitat classification. Report EEA/NSV/16/005. EEA, Copenhagen.URL.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

Annex IV: EBVs not included in the assessment

ID	Realm	EBV class	EBV name	Definition
43b	Freshwater	Genetic composition	Genetic diversity of selected freshwater taxa	Genetic richness (number of alleles in a population) and genetic evenness (expected proportion of heterozygotes in a population at equilibrium) of taxa.
58c	Freshwater	Species traits	Phenology of migration of wetland birds	The annual timing of arrival and departure of European wetland migratory bird species at breeding, staging and wintering sites over time.
14	Freshwater	Community composition	Ecological Quality Ratio (EQR) of freshwater zooplankton	The ecological status of zooplankton in European lakes, measured as Ecological Quality Ratio (EQR).
15	Freshwater	Ecosystem structure	River Connectivity/Free river flow	The length of free-flowing rivers (without barriers) and the natural longitudinal and lateral connectivity of rivers and lakes.
16	Freshwater	Ecosystem structure	Ecosystem distribution of freshwater EUNIS Habitats	The presence/absence or probability of occurrence of freshwater EUNIS habitats in contiguous spatial units (grid cells) over time.
17	Freshwater	Ecosystem structure	Structural complexity of riparian habitats	The vegetation structure, width and length, or topographic heterogeneity of riparian habitats over time, representing the density, cover, variability and three-dimensional arrangement of vegetation and other structural features.
20	Freshwater	Ecosystem function	Freshwater primary productivity	The amount of carbon that is removed by lake habitats (and large rivers) from the atmosphere over time and stored in biomass, roots and sediments.
42c	Marine	Genetic composition	Genetic diversity of selected marine taxa	Genetic richness (number of alleles in a population) and genetic evenness (expected proportion of heterozygotes in a population at equilibrium) of taxa.
28	Marine	Species populations	Species distributions of benthic marine invertebrates	The presence/absence or probability of occurrence of benthic invertebrate species form the Habitats Directive in EU's benthic habitats within contiguous spatial units (grid cells) over time
30	Marine	Species traits	Phenology of migration of marine birds and mammals	The annual timing of arrival and departure of European marine migratory bird and mammal species at breeding, staging and wintering sites over time.
31	Marine	Community composition	Functional composition of marine phyto/zooplankton (based on traits)	The functional composition and diversity (e.g., based on morphological, physiological or behavioral traits) of marine phyto/zooplankton in EU's marine waters within contiguous spatial units (grid cells) over time
35	Marine	Ecosystem structure	Ecosystem distribution of oyster reef habitats	Presence/absence or probability of occurrence of oyster reef habitats in EU's marine waters within contiguous spatial units (grid cells) over time.
36	Marine	Ecosystem function	Degree of seabed disturbance	The estimated amount of permanent or temporal disturbance of seabed substrate or morphology caused by human activities such as construction, dredging, sand and gravel extraction, deposition of dredged material, shipping and bottom trawling.
37	Marine	Ecosystem function	Harmful marine algal blooms	Distribution, intensity, frequency and position of harmful algal blooms in European coastal waters which occur when cyanobacteria accumulate in water, with the potential to harm the health of humans, plants, and animals
38	Marine	Ecosystem function	Phenology of marine spring phytoplankton bloom	The annual timing and intensity of spring phytoplankton blooms in EU's marine waters over time.
39	Marine	Ecosystem function	Marine primary productivity	Productivity of organic compounds from atmospheric or dissolved carbon dioxide by cyanobacteria, algae and marine plants in EU's marine waters within contiguous spatial units (grid cells) over time.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.

EUROPABON

ID	Realm	EBV class	EBV name	Definition
42a	Terrestrial	Genetic composition	Genetic diversity of selected terrestrial taxa	Genetic richness (number of alleles in a population) and genetic evenness (expected proportion of heterozygotes in a population at equilibrium) of taxa.
53	Terrestrial	Species populations	Species distributions of lichens (as indicators of pollution)	The presence/absence or probability of occurrence of ecological quality indicator lichen species within contiguous spatial units (grid cells) across the EU over time.
55a	Terrestrial	Species populations	Species abundances of selected terrestrial disease vectors	The estimated count of individuals of animal vectors within contiguous spatial units (grid cells) across the EU over time.
55b	Terrestrial	Species populations	Species abundances of selected terrestrial crop pests	The estimated count of individuals of crop pests within contiguous spatial units (grid cells) across the EU over time.
56	Terrestrial	Species traits	Phenology of fructification of mushrooms and wild fruits	The annual timing of the fructification of wild mushroom species and wild fruits within contiguous spatial units (grid cells) across the EU over time.
57	Terrestrial	Species traits	Phenology of flowering and leaf senescence	The annual timing of flowering and leaf senescence of European flowering plants and deciduous trees within contiguous spatial units (grid cells) across the EU over time.
59/60	Terrestrial	Community composition	Community biomass of selected functional groups of terrestrial arthropods (e.g., predator, decomposer)	Estimated community biomass of arthropod functional groups.
65	Terrestrial	Ecosystem structure	Vertical structure of terrestrial vegetation	The vertical structure of terrestrial vegetation over time, representing vegetation height, cover, density, structural variability and three-dimensional arrangement of vegetation biomass.
67	Terrestrial	Ecosystem structure	Connectivity of terrestrial ecosystem habitat types	The degree of connection of EUNIS habitats within a landscape, in terms of their components, spatial distribution and ecological functions.
69	Terrestrial	Ecosystem function	Terrestrial primary productivity	The amount of CO2 fixed by terrestrial plants through the photosynthetic reduction of CO2 into organic compounds minus the CO2 emitted by autotrophic respiration within contiguous spatial units (grid cells) across the EU over time.
71	Terrestrial	Ecosystem function	Ecosystem disturbance as measured by HANPP	Human Appropriation of Net Primary Production (HANPP) is the proportion of terrestrial NPP consumed directly and indirectly through human land.
73	Terrestrial	Ecosystem function	Standing and lying deadwood	Amount of non-living standing and on the ground woody biomass within contiguous spatial units (grid cells) across the EU forest and other wooded lands over time.

Further details for each EBV (specifications in terms of spatio-temporal resolution, metrics, and taxonomic focus) can be found in the Deliverable 4.1.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003553.