ARPHA Preprints

PENSOFT

Project Report

Author-formatted document posted on 10/07/2023

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

DOI: https://doi.org/10.3897/arphapreprints.e109168

D3.1 Inventory of current European network for monitoring. Web-based database

Alejandra Morán-Ordóñez, David Martí Pino, Lluís Brotons

ARPHA Preprints *Author-formatted document posted on 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

EUROPABON

D3.1 Inventory of current European network for monitoring. Web-based database.

30/11/2021

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	**D3.1 Inventory of current European network for monitoring. Web-based database.**
Deliverable number	**D3.1**
Contractual date of delivery	**30.11.2021**
Actual date of delivery	**30.11.2021**
Type of deliverable	**Website, patents fillings, etc.**
Dissemination level	**Public**
Work package number	**WP3**
Institution leading work package	**Centre for Ecological Research and Forestry Applications (CREAF)**
Task number	**T3.1**
Institution leading task	**The Centre for Ecological Research and Forestry Applications (CREAF)**
Author(s)	**Alejandra Morán-Ordóñez, David Martí Pino, Lluís Brotons**
EC project officer	**Laura Palomo-Rios**

Deliverable description	This Deliverable describes the database on biodiversity monitoring initiatives at the European level collected by EuropaBON (WP3 - task 3.1; EuropaBON biodiversity database) and the web-based platform (website) that contains it. The website serves the dual purpose of being the platform for data entry, as well as to allow the visualisation and quick consultation of the collected data. This report also gives a summary of the data collected up to the date of delivery of the report (end November 2021) and explains future developments of the database and how it will be used by other EuropaBON WPs and tasks.
Keywords	**Biodiversity Monitoring, Website, DataBase, Data workflow**

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

Contents

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

EXECUTIVE summary

This **report describes the database on biodiversity monitoring initiatives at the European level collected by EuropaBON** (WP3 - task 3.1; EuropaBON biodiversity database from hereon) and the web-based platform (website) that contains it. The website serves the dual purpose of being the platform for data entry, as well as to allow the visualisation and quick consultation of the collected data. This report also gives a brief summary of the data collected up to the date of delivery of the report (end November 2021).

Previous efforts on collecting information on existing monitoring efforts in Europe have concentrated in describing programs using a generalist approach that has received criticism because of lack of completeness even for well documented groups such as birds. **Including all available information on monitoring in Europe at any spatial scale is difficult** because of the high number of unlinked initiatives, specially at local scales, and the dynamic nature of these projects both in terms of the emergence of new efforts and the disappearance of old ones. To address this challenge, the **EuropaBON project focusses on the monitoring network concept and aims at identifying, with priority, those monitoring efforts that are coordinated** and especially those in which this coordination is consistent at a supranational level and at the European scale. **Coordination in the context of monitoring is related to the integration of data and information across scales which is the underlying concept of a future successful biodiversity monitoring network in Europe**.

The web-based platform https://monitoring.europabon.org/, gathers three types of information for each biodiversity data workflow: **1) Integration nodes** or institutions/projects/initiatives who integrate/process biodiversity data to generate EBVs, EESVs or any other indicators with potential relevance to environmental policy, particularly at the European level; 2) details of the **biodiversity monitoring data** or initiatives/schemes responsible for the collection of biodiversity-related information (what are the species recorded? how is the data collected? who is responsible for the monitoring? etc.); and 3) **Data streams** or how is the data flowing between monitoring programs and institutions that integrate data? what type of products are generated? (Figure 1). **The data included in the database and the novel visual approaches used to illustrate monitoring workflows could be further refined to better communicate how biodiversity monitoring is used in Europe** to produce relevant EBVs, EESVs and indicators, and therefore help articulating the necessary discussion behind the design of a future European Biodiversity network.

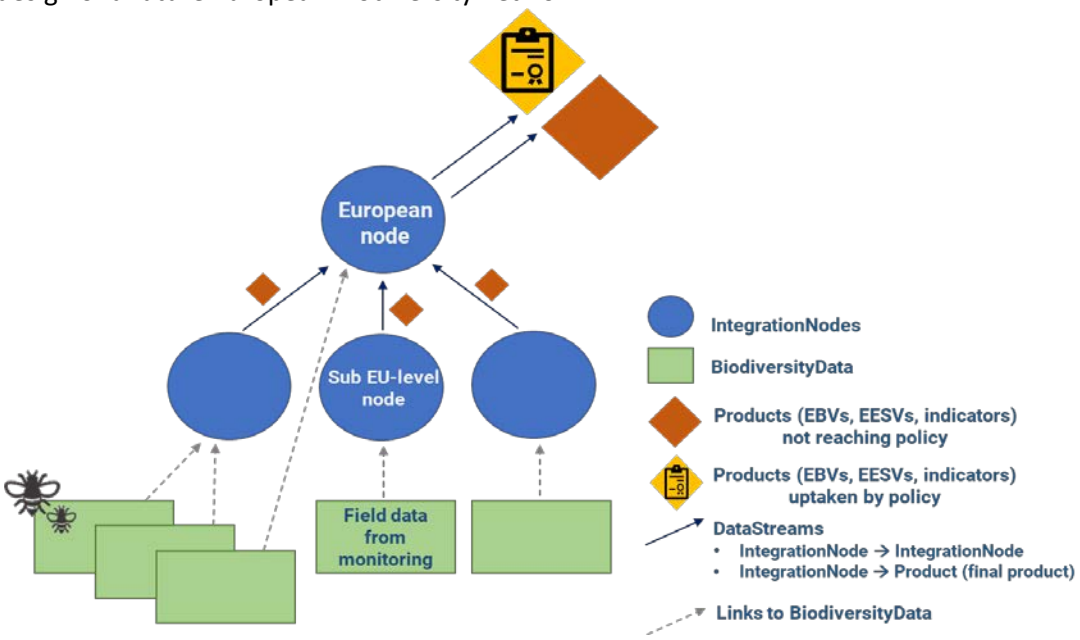


Figure 1. *Schematic diagram of the data flows across a biodiversity data workflow*

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

By end of November 2021, the database gathered data on more than European 74 biodiversity data workflows accounting for a total of 669 biodiversity datasets. An exploratory analysis of the data showed NGOs process and integrate most biodiversity data across Europe (38% of integration nodes), being count data (for abundance and density estimation) the most common type of data collected across all biodiversity and ecosystems data sources gathered (80% of data). Most data flows have their source in data collected through systematic monitoring programs (62%). The purpose of most biodiversity and ecosystems data sources gathered is to advance the ecological knowledge of species. The database is dominated by monitoring programs of terrestrial ecosystems and species (86% of datasets), being plants, birds and invertebrates, the groups better represented. Geographic coverage of data was uneven across Europe for different taxonomic groups, although in general terms there was a small coverage of data across most taxa in southeastern European countries. Around 55% of biodiversity monitoring programs in the database collect time-series with a small fraction of these being open-access. Only around 22 % of data flows generate EBVs and EESVs.

The EuropaBON database does not cover all biodiversity data currently monitored across Europe. For example, it is difficult to trace data integrations made at an academic level based on biodiversity opportunistic observations. Also **monitoring efforts in Europe collecting data at the national or subnational level without integration at the transnational level are not well covered** in the EuropaBON database. Moreover, **further efforts need to be developed in gathering spatial information on sampling locations** for selected, key monitoring networks and exploring ways of representing this information in the database.

Next developments of the database with selected stakeholders will include gathering the information on **data flows linked to the reporting of the Birds and Habitats Directives** (now flows linked to European Directives were only partially covered by the WISE-2 Biology data- Water Framework Directives). The collection and documentation of these data flows will be facilitated by the contact established through the workshops and consultation process made for the *EuropaBON User and Policy Needs Assessment - Deliverable 2.2*). This joint work with stakeholders will also allow to better quantify the comprehensiveness of the current database contents as well as to identify current gaps and bottlenecks of biodiversity monitoring across Europe in terms of thematic content, spatial and temporal coverage (WP3 - task 3.2), data flows and availability (WP3 - task 3.3), as well as cost-effectiveness (WP3 - task 3.4).

Finally, a future European monitoring network will probably be better represented as a **network of networks and the structure of the EuropaBON monitoring database** is designed to represent this idea. In summary, the current information in the database would represent responsibilities focusing on European integration nodes and their key biodiversity data components, whereas the incomplete data at lower levels (integration nodes at the national level not connected to European nodes) would reflect flows at the national and subnational levels and a more dedicated effort here would be required in the future.

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

1. Rationale

This report describes the database on biodiversity monitoring initiatives at the European level collected by EuropaBON (WP3 - task 3.1; EuropaBON biodiversity database from hereon) and the web-based platform (website) that contains it. The website serves the dual purpose of being the platform for data entry, as well as to allow the visualisation and quick consultation of the collected data. This report also gives a brief summary of the data collected up to the date of delivery of the report (end November 2021).

Previous efforts on collecting information on existing monitoring efforts in Europe (i.e. the EUMON database[1]) have concentrated in describing programs using a generalist approach that has received criticism because of lack of completeness even for well documented groups such as birds[2,3]. Including all available information on monitoring in Europe at any spatial scale is difficult because of the high number of unlinked initiatives, specially at local scales (for instance for protected areas), and the dynamic nature of these projects both in terms of the emergence of new efforts and the disappearance of old ones.

To address this challenge, the EuropaBON project is focussing on the monitoring network concept and aims at identifying, with priority, those monitoring efforts that are effectively coordinated and especially those in which this coordination is consistent at a supranational level and at the European scale. Coordination in the context of monitoring is related to the integration of data and information across scale which is the underlying concept of a future successful biodiversity monitoring network in Europe. By identifying current efforts in which this integrated coordination already exists, task 3.1. will provide the starting point for the project to further identify what are the current gaps and bottlenecks that biodiversity monitoring in Europe faces to deliver cost-effective EBVs based products that can effectively inform policy at the European scale.

The goal of EuropaBON task 3.1 was to set up a web-based platform to collect/record the current biodiversity data workflows across Europe and provide an initial assessment of the data entered. Further, more detailed analyses of gaps and bottlenecks in European scale monitoring efforts workflows will be developed in task 3.2 and 3.3. This database will prove a key tool to understand how biodiversity data collected in monitoring schemes across Europe flows through different institutions and programs and gets processed to produce Essential Biodiversity Variables (EBVs)[4,5], Ecosystem Services Variables (EESVs) and /or other European policy-relevant indicators. Some examples of European level biodiversity data workflows are the reporting of the Habitats and Water Framework Directives, the Wild Bird Indicators produced in the Pan-European Common Bird Monitoring Scheme or the European Grassland Butterfly Indicator produced by the European Butterfly Monitoring Schemes. The EuropaBON biodiversity monitoring database builds upon previous efforts of biodiversity data collection on monitoring programs across Europe, including the EuMON[1] (http://eumon.ckff.si/) or the GEOBON (https://portal.geobon.org/datasets). From the databases retrieved in those previous European projects, the EuropaBON database only retained those that are being incorporated into different data flows and have led to the generation of Europe-wide EBVs, EESVs or other indicators. In this report, we use the term "**integration initiatives"** to refer to each one of these biodiversity data workflows. The set of integration initiatives collected in this database will constitute the information flow network of the EuropaBON project and will be used to identify the current gaps and bottlenecks of biodiversity monitoring across Europe in terms of thematic content, spatial and temporal coverage (WP3 - task 3.2), data flows and availability (WP3 - task 3.3), as well as cost effectiveness (WP3 - task 3.4).

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

[1] Wetzel, F.T., et al. (2018). Unlocking biodiversity data: Prioritization and filling the gaps in biodiversity observation data in Europe. *Biological conservation,* 221,78-85.

[2] Voříšek, P., et al. (2018) Wetzel et al. fail to identify the real gaps in European bird monitoring. *Biological Conservation*, 225, 245-246.

[3] Moussy, C., et al. (2021) A quantitative global review of species population monitoring. *Conservation Biology*. DOI: 10.1111/cobi.13721.

[4] Pereira, H.M., et al. (2013) Essential biodiversity variables. *Science*, 339: 277-278.

[5] GEO BON (2021). What are EBVs? [Online]. Available from: https://geobon.org/ebvs/what-are-ebvs/

2. The team

Alejandra Morán Ordóñez, David Martí Pino and Lluís Brotons from CREAF led the development of the website and the data collection, with support of other members of the CREAF team (Dani Villero, Sergi Herrando, Robert Manzano, Laura Recoder, Quim Canelles and Magda Pla). The leading team received valuable input on the website and the database structure and contents from other members of WP3 Joana Santana (ICETA), Pedro Beja (ICETA), Francisco Moreira (ICETA), Jannicke Moe (NIVA), Anne Lyche Solheim (NIVA), Liis Tiirmann (UTartu), Simon Potts (UReading) and Tom Breeze (UReading), Ingolf Kühn (UFZ), Roy H.A. van Grunsven (Dutch Butterfly Conservation), Joachim Maes (JRC; coordinating team), Henrique Pereira (MLU; coordinating team) and Jessica Junker (MLU; coordinating team). Additional input was received from Daniel Kissling (UvA) and Maria Dornelas (USTAN) (leaders of WP4).

We gratefully acknowledge the expertise and time of colleagues who have collaborated in entering/reviewing data into the EuropaBON database: Gabriel Gargallo (Institut Català d'Ornitologia - ICO, Spain), Eivind Ekholt Andersen (NIVA, Sweden), Borja Jiménez Alfaro (Universidad de Oviedo, Spain), Pierre Rasmon (Université de Mons, Belgium), Denis Michez (Université de Mons, Belgium), Guillaume Body (Office Français de la Biodiversité, France), Sara Vallecillo (Joint Research Centre, Ispra, Italy), Dag Rossland (Norwegian Environment Agency), Monika Peterlin (European Environment Agency), Luísa Rodrigues (Institute for Nature Conservation and Forests, ICNF, Portugal), Pierre Bonnet (CIRAD UMR Amap, France), Peter Borgen Sørensen (Aarhus Universitet, Denmark), Judy Shamoun-Baranes (University of Amsterdam, The Netherlands), Silke Bauer (Swiss Ornithological Institute, Switzerland), Rowan Edwards (BWARS - Bees Wasps and Ants Recording Society, UK), Michael Edwards (BWARS - Bees Wasps and Ants Recording Society, UK), Claire Carvell (UK Centre for Ecology & Hydrology), Michael Garratt (UReading, UK).

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

3. The WebBase Database: https://monitoring.europabon.org

The main objective of the EuropaBON monitoring database is to be able to explicitly describe current workflows of monitoring efforts delivering biodiversity information in Europe. Monitoring efforts are understood in this deliverable as projects designed to track biodiversity change in time in a spatially explicit context. To do this, and in contrast to previous efforts, the present database has developed a novel network framework allowing to map the flow that data experiences since it is collected on the field until it is used in different aggregation steps across different spatial and temporal scales.

The main objective of the database is therefore, not to include all available monitoring projects obtaining data on biodiversity change, but to be comprehensive in including all representative networks that, at European scale, coordinate and integrate biodiversity data and allow obtaining representative biodiversity estimates of change. By, building on the current approach, the database proposed by the EuropaBON project would be able to be expand in the future and gradually grow by including biodiversity networks at progressively lower spatial scales (i.e., national and subnational) and eventually contribute to a more detailed assessment of the monitoring panorama in Europe.

3.1 Main elements of the database

The main elements of the database are therefore the nodes of the monitoring networks, what we call in this report "'**integration initiatives"**: these are initiatives/projects/programs that integrate biodiversity data from monitoring schemes and process them to produce aggregated biodiversity data or Essential Biodiversity Variables (EBVs), Essential Ecosystem Services Variables (EESVs) and/or other European policy-relevant indicators.

Each integration initiative is composed of three distinctive elements (Figure 1):

1. **IntegrationNodes**: institutions/projects/platforms integrating/processing biodiversity data to generate EBVs, EESVs or any other indicators with potential relevance to environmental policy, particularly at the European level. One single integration initiative can be composed of many nodes operating at different scales (e.g., subnational, national, European-level) in a coordinated manner. One single institution can act as an integration node of different integration initiatives.
2. **BiodiversityData**: biodiversity monitoring initiatives/schemes responsible for the collection of biodiversity-related information, including data collected from the field as well as remote sensing data. The focus of this EuropaBON task is not to map all monitoring initiatives in Europe, but to ensure that those monitoring initiatives providing data for current biodiversity workflows at the European level are well represented in the database.
3. **Datastreams (arrows)**: these represent data flows and connect IntegrationNodes across different scales. These contain two types of information: information related to the data flowing between nodes (**dataset**) and whether this has been integrated/processed ("products": EBVs, EESVs, indicators) or not (raw and aggregated data) (**data process**). If the data has been integrated/processed, it also contains information about the integration method (e.g., statistical modelling, expert opinion). A data stream that does not connect with another integration node represents a final Product (EBVs, EESVs, Indicator) that can eventually be uptaken by policy or not (this is represented in Figure 1 with diamonds of different colours).

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 10/07/2023.* DOI: https://doi.org/10.3897/arphapreprints.e109168

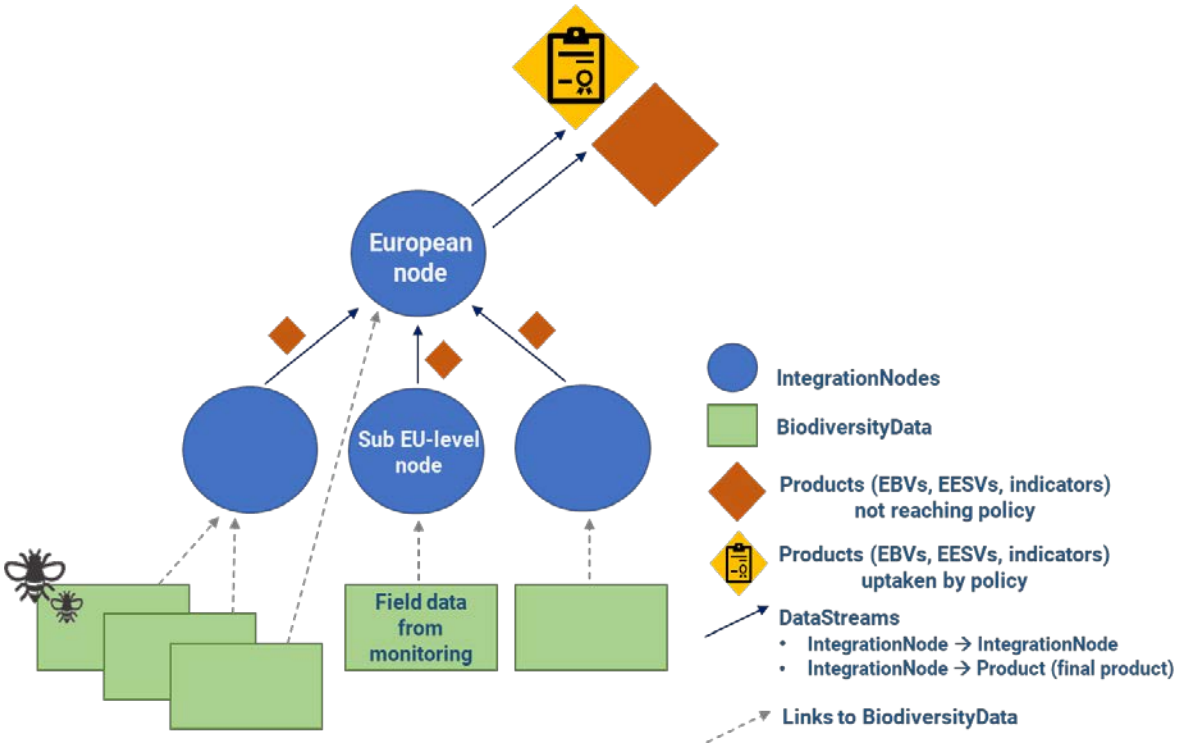


Figure 1. *Schematic diagram of the data flows across an integration initiative*

3.2 The website: accessibility and technical details

The website can be accessed through this link: https://monitoring.europabon.org.

The website has been developed using technologies licensed as free software. The Backend has been programmed using the Python language, taking advantage of the powerful Django web framework.

A PostgreSQL database has been chosen to store all data. We chose PostgreSQL because it's the most powerful relational open-sourced database, it has powerful GIS features, and the developers have advanced knowledge in this kind of database. The Frontend has been developed in HTML, JS, jQuery, and some libraries as Leaflet and d3.js for mapping and data visualizations. The whole website including the database is stored in a VPS located in Frankfurt. The deployed website uses Nginx web server technology.

We created two different user roles: website administrators and external users. The administrators can create projects (one project per integration initiative) and new user profiles, assigning each user to one or various integration initiatives. With this approach, external users can only edit and create data related to the integration initiatives for which they are responsible. Once the user logs in to the system, it can start entering the data forms explained in section 3.3. (Integration Nodes, Biodiversity Data, Data Streams). The users can consult the list of integration initiatives and data already entered in the database, add new data forms, modify their contents, and delete them. Users without login details can only consult the database contents (some fields of the database are not visible to external users -e.g., contact emails of monitoring programs coordinators – for data privacy and sensitivity reasons).

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

The website counts on three data output summary modules: integration node overview and visualization, search, and global stats.

1) The first module allows users to choose one integration node and see the structure of the integration node, products, linked data streams and other integration nodes. It allows the user to easily explore the hierarchy and data flows integrated by each integration node.

2) The second module ("Search and Visualize") allows the users to search integration nodes using several filters (i.e., target taxa, geographical coverage, or project). Once data is filtered, the user can see:

- Products generated by the selected integration nodes.
- Geographical coverage shown on a map.
- Network graph showing data flows between integration nodes, products and linked biodiversity data. The detailed data of each element of the network plot can be consulted (data opens in a pop-up window) (Figure2).

3) The third module ("Stats > Global Stats") allows users to consult basic statistics about the data gathered in the database.

Administrator role users have access to extended features:

- Manage uploaded thesaurus.
- Export all data in CSV files.
- Manage institutions.
- Users' management.

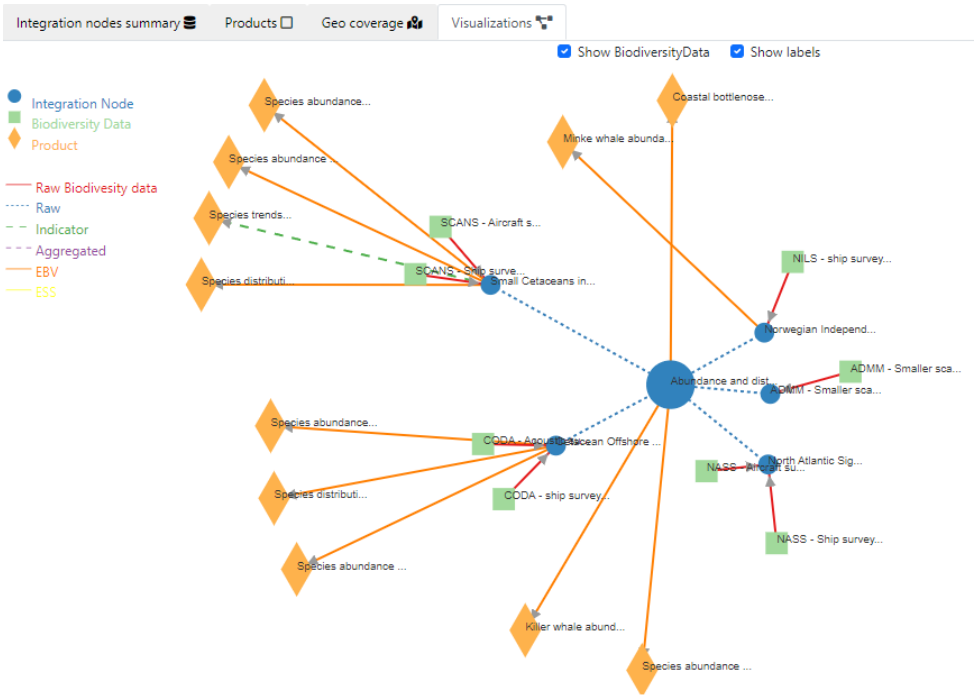


Figure 2. *Visualization of one of the integration initiatives collected in the EuropaBON database (network visualization): blue circles represent integration nodes, green squares represent biodiversity datasets and yellow diamonds are products (EBVs, EESVs or indicators). The arrows indicate data flows between biodiversity data and integration nodes (red), between integration nodes (blue arrows) or final product (yellow arrows that end into a given product).*

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 10/07/2023.* DOI: https://doi.org/10.3897/arphapreprints.e109168

3.3 The website: tables and data forms

The three main elements of the database appear on the top left corner of the website as three independent tabs (**data forms**) (Figure 3):

1) The **Integration Nodes** tab seeks to collect relevant information for each of the nodes within an integration activity (one entry for node – represented with blue circles in Figure 1). Each node is an institution integrating data and/or generating 'products' (EBVs, EESVs and indicators). These data and products can flow into another upper-level integration node (e.g. European-level node) or represent 'final products' that inform - or could potentially inform - policy. Note here that the hierarchy of integration nodes (the number of sub-European levels) will vary depending on the integration initiative. The same institution can act as an "integration node" for different integration initiatives. For example, the European Bird Census Council - EBCC is the European-level integration node of three different integration initiatives: the European Atlas of Breeding Birds - EBBA2, the EuroBirdPortal - EBP and the Pan-European Bird Monitoring Scheme - PECBMS. In this case, the EBCC as a European-level integration node is listed three times in the database. However, each time this node will contain different information since for example, the details of the main coordinator, the funding available and its origin, or how the data is integrated might differ between integration initiatives carried out by the same institution.

2) The **BiodiversityData** tab seeks to collect information about the biodiversity monitoring initiatives underneath integration nodes (data collection).

3) The **DataSteams ("All Data streams")** tab collects information about how the data is processed (if)? and how does it flow? This tab also accommodates information about "products" (EBVs, EESVs, indicators) generated by integration nodes with potential links to policy (see the example represented with the orange diamond in Figure 1). Therefore, the DataStream tab accommodates information on 1) dataflows between lower-level and upper-level integration nodes (how the products generated in a lower-level integration node flow into an upper-level integration node; 2) dataflows between European-level Integration Nodes and Policy and 3) final products from European-level integration nodes that have the potential to inform policy but currently lack that link.

Details on the data fields collected within the Integration Nodes, BiodiversityData and DataStream tabs can be found in Annex 1. How the different tables are linked in the database (conceptual model of the EuropaBON database) is shown in Annex 2.

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

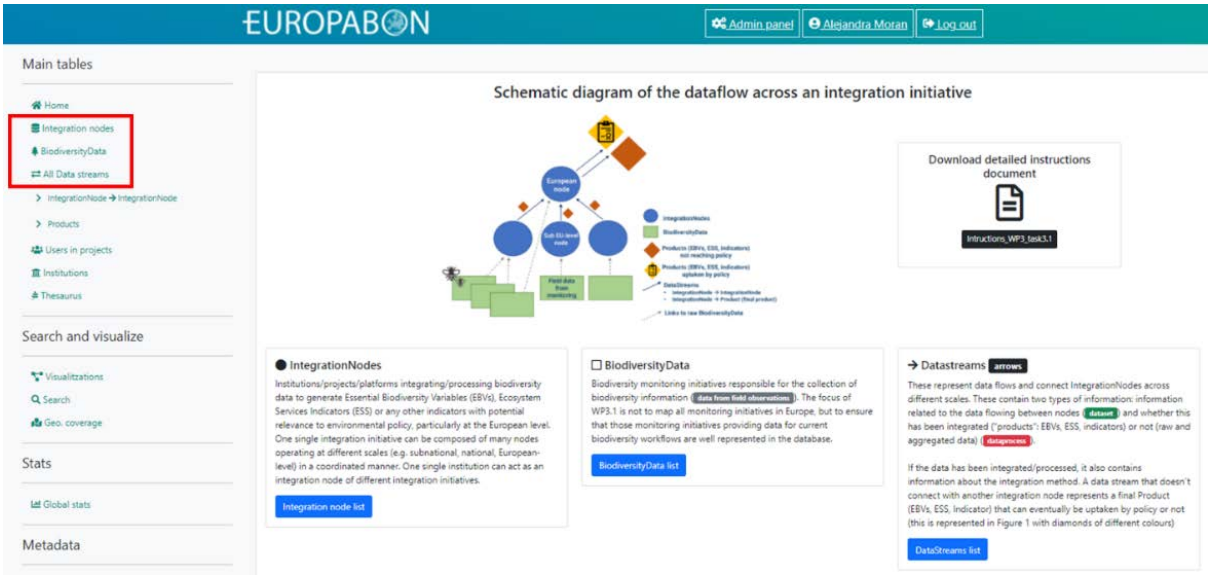


Figure 3. *Home page of the Website database. The three different data forms can be accessed from the top-left corner of the website (highlighted in this figure with a red rectangle)*

4. Data collection

Together with partners included in WP3 activities, other EuropaBON partners and by active contact with stakeholders from the wider EuropaBON community, we searched for information about biodiversity monitoring relevant players in Europe and aimed at describing in more detail the workflows in which they were involved. The search for integration initiatives followed different venues: on the one hand, the partners of the EuropaBON consortium as well as the members registered on the EuropaBON website (stakeholders, data providers, managers, etc.), were asked to participate in the identification of key biodiversity data integration initiatives with relevance at European level; the other hand, the CREAF team made a broad internet search of European-level EBVS, EESVs and biodiversity indicators linked to monitoring programs using different keyword combinations (e.g. biodiversity OR species OR habitats OR indicators AND/OR monitoring OR aggregation AND Europe* OR EU) and consulted the main reference repositories of such data at the European level containing such information like the indicators repository of the European Environmental Agency (EEA). Members of the CREAF team and other WP3 members also screened previous monitoring databases (the EuMON and the GEOBON) and retained biodiversity monitoring programs that belonged to a European-level integration initiative.

Moreover, experts on different taxonomic groups and coordinators of main integration initiatives across Europe (e.g., the PCBMS, the European Butterfly Monitoring Scheme, the European Vegetation Archive) were directly contacted to retrieve information about the initiatives they were responsible for. EuropaBON partners also contributed to populate the database with initiatives they knew, coordinated or were part of. The goal was to be as comprehensive as possible in gathering the relevant biodiversity integration initiatives across Europe, making a special effort in covering as many different taxonomic groups and realms (terrestrial, freshwater, coastal, marine) as possible.

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 10/07/2023.* DOI: https://doi.org/10.3897/arphapreprints.e109168

5. Summary of contents by end November 2021

The EuropaBON biodiversity monitoring data set gathered information from 74 European-level integration initiatives (Annex 3), accounting for a total of 772 integration nodes (8.4% European-level nodes, 81.2% National and approx. 10 % corresponding to subnational, regional or global nodes) and 662 different monitoring programs/biodiversity data sets.

Who is integrating/processing biodiversity data across Europe? 37% of European level integration nodes are run by supra-National public institutions, 25 % by NGOs and 20 % by Research bodies (Universities, research centers), with approx. 17% of these European nodes run by a combination of these. At the national level these percentages are: 7 % of integration nodes run by Governmental institutions, 40% by NGOs and 26% by research bodies, with 26% of integration nodes run by a combination of these (Figure 4). In general terms, and across all spatial scales, **NGOs process and integrate most biodiversity data across Europe (38% of integration nodes)**, followed by research bodies (35%) and governmental or supra-governmental institutions (19.8%).

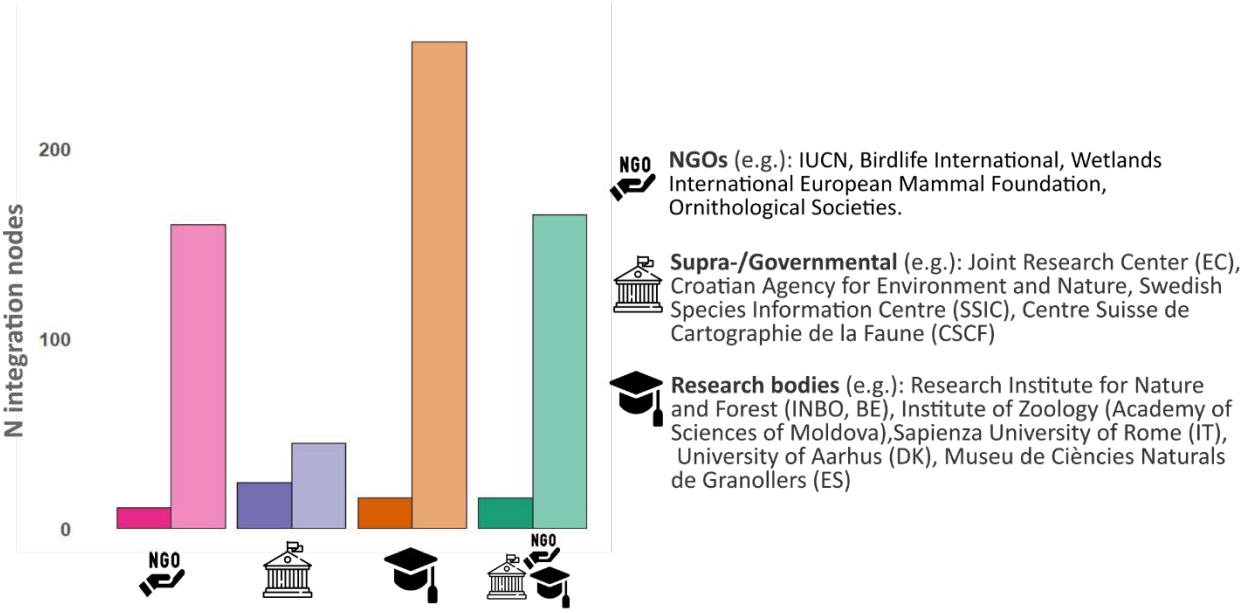


Figure 4. *Number of integration nodes coordinated by NGOs, Governmental or supra-governmental institutions, Research bodies (including Universities) or by a combination of these (EuropaBON database numbers by end November 2021). In each group, the left column corresponds to European nodes and the right column to National nodes.*

Overall, we observed **a bias in the geographic coverage of integration nodes** gathered in the EuropaBON database, with South-Eastern European countries having less integration initiatives at the national level (Figure 5). This bias is much more pronounced when the data collected are evaluated at taxonomic group level (see Annex 4 for some examples across specific taxonomic groups).

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 10/07/2023.* DOI: https://doi.org/10.3897/arphapreprints.e109168

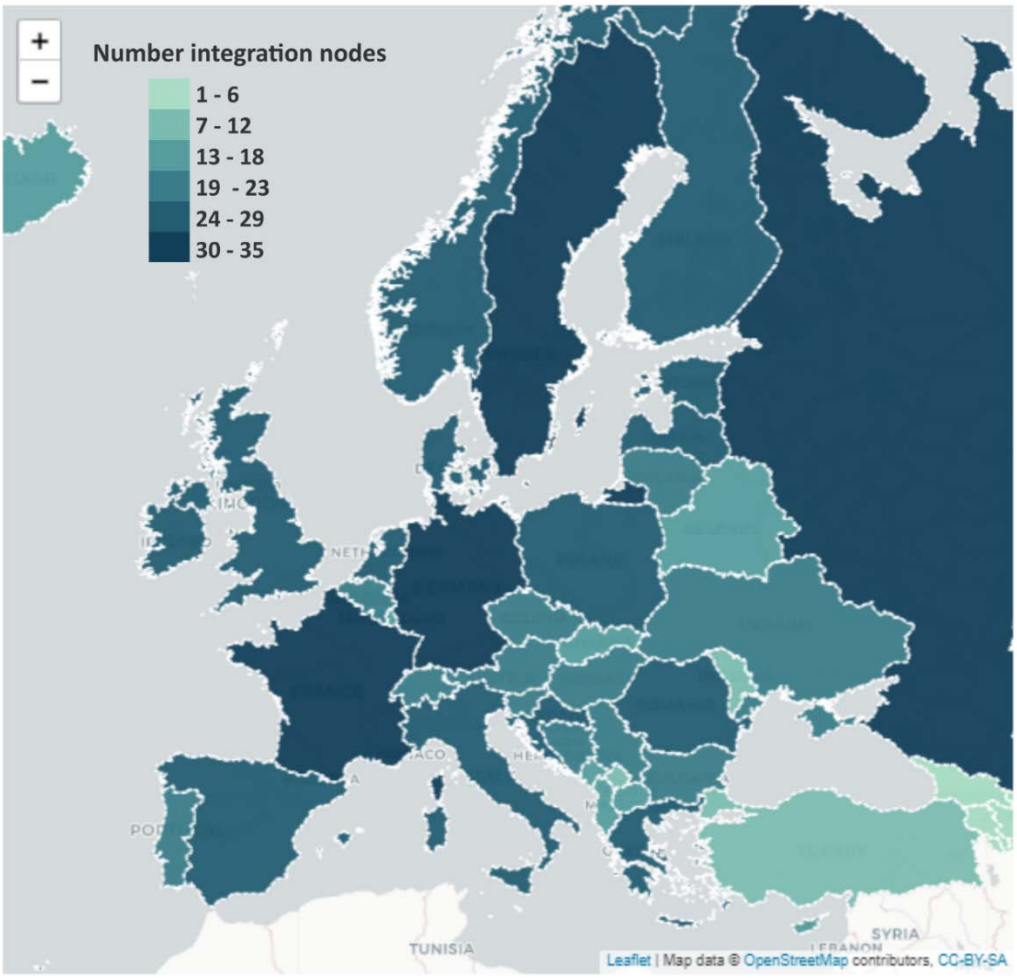


Figure 5. *Number of integration initiatives with national coverage across all taxonomic groups*

Main nature of biodiversity and environmental monitoring programs gathered

Data origin: From the 669 unique biodiversity and environmental data sources collected, **62 % corresponded to systematic monitoring programs**, 3.7 % to databases of opportunistic observations (e.g. data portals such as https://observation.org/ or https://www.ornitho.eus/), 1.6% to remote sensing data and approx. 10 % combined data from different origins (e.g. the Ecosystem monitoring of coastal waters of Norway - ØKOKYST retrieves information combining systematic monitoring programs, automatic recording and remote sensing data). We were unable to trace the information about data origin for 22% of the biodiversity data recorded in the database because detailed information about the workflows promoted by the integration node were not always available (case for example of many European Red Lists).

Origin of data collection: From the biodiversity data programs from which we were able to collect information about their data origin (N=389), in 31 % of cases data were collected by citizen scientists, 13.6% were collected by government authorities, while 32.14% was professionally executed (e.g., NGOs, consultancies, researchers) (Figure 4). **Count data for abundance estimation were the most common type of data collected in the biodiversity and environmental data sources gathered (80% of data)**, followed by presence-only data (occurrences; 12.9%), forest inventory data type (parameters such as diameter at breast height - DBH, forest canopy cover, etc.; ~4.5% of data) and presence-absence data and remote sensing data (< 3% of data) (Figure 6).

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 10/07/2023.* DOI: https://doi.org/10.3897/arphapreprints.e109168

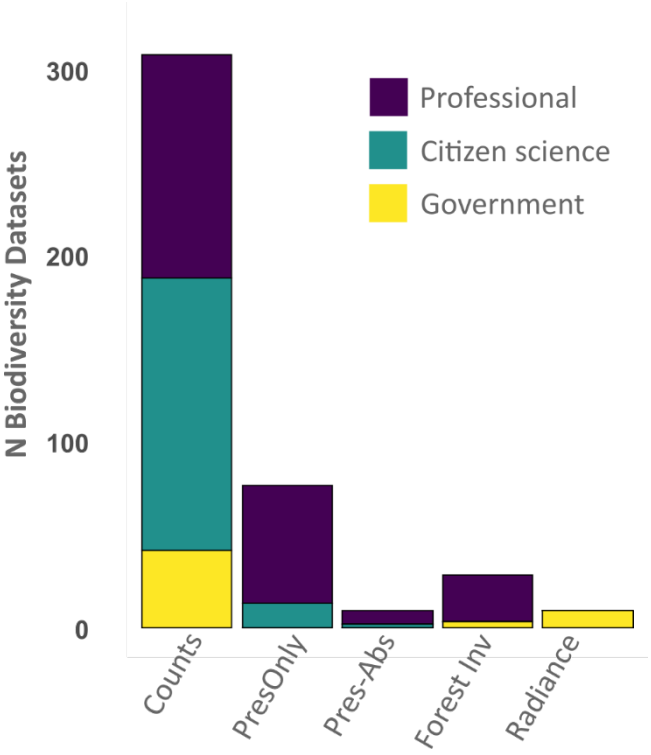


Figure 6. *Number of Biodiversity Datasets categorized by their data origin (citizen science, governmental authorities, professional monitoring) and data type (counts, presence-only, Presence-Absence, Forest Inventory data and spectral radiance - remote sensing)*

The purpose of most biodiversity and environmental data sources gathered is to advance ecological knowledge of species, habitats, communities and ecosystems (76% of data; e.g. common monitoring bird surveys, small mammals monitoring programs), while around 15% had multiple purposes including monitoring reference sites, monitoring linked to restoration programs or environmental impact assessments (e.g. monitoring of waterbodies link to the reporting of the water framework directive - see ***Box 2***). **Only a small fraction of all biodiversity and environmental data sources collected are open access** (10.6%), while 38% might be available on demand.

Monitoring programs on **plants and plant communities account for most datasets in the EuropaBON database (25% of biodiversity and ecosystems data collected)**, followed by those targeting birds (22.4%), invertebrates (15 %), mammals (11%) and amphibians and reptiles (~5% each group) with other groups (e.g., fish, fungi, Phytoplankton, macroalgae or land cover data) having a representation of less than 5% in the database (Figure 7). **The database is dominated by monitoring programs of terrestrial ecosystems and species (86% of datasets)**, being marine, freshwater, and coastal species and ecosystems less represented (15%, 14% and 9% of datasets, respectively).

The temporal resolution of data collection varied greatly across biodiversity and environmental monitoring programs, from daily (e.g. radar data from meteorological stations or data collected through citizen science platforms such as https://www.fauna.hr/ and eBird) to > 10 years (e.g. monitoring target efforts to the development of the European Atlas of Breeding Birds or the Atlas of Mammals). **Around 55% of programs in the database collect time-series** data and 45% one-off data. Around 78% of programs collecting time series (N=287) correspond are currently ongoing. The average duration of these time-series monitoring programs is about 30.7 years, with the longest-lived programs (around a century) being the National Forest Inventories of Norway and Finland and the Woodland surveys in the United Kingdom and the marine environmental monitoring programs in Sweden and France.

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

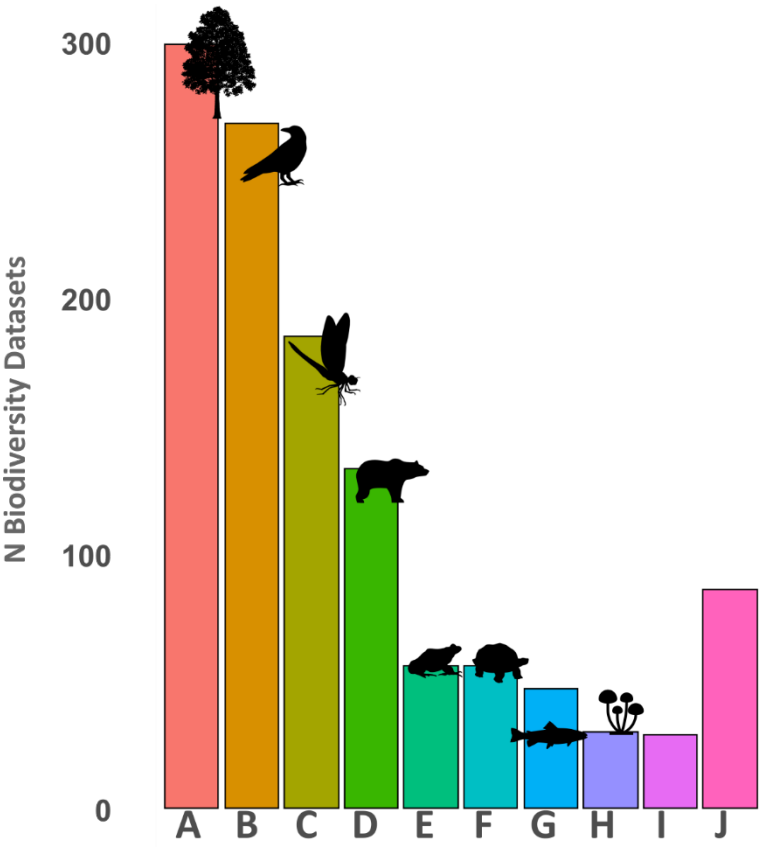


***Figure 7**. Distribution of biodiversity and ecosystems data collected in the EuropaBON database: A) Plant and plant communities; B) Birds, C) Invertebrates; D) Mammals; E) Amphibians; F) Reptiles; G) Fish; H) Funghi; I) Phytoplankton and J) Others (e.g., Phytobenthos, Macroalgae, land cover data)*

How does the data flow between integration nodes? Raw and aggregated data, EBVs, EESVs and indicators

The data is processed and flows among integration nodes in different ways: raw data (35.3 %), aggregated (27%), EBVs and EESVs (around 22.4%) and indicator values (15%). The majority of generated and flowing data are maps (55%), species trends (11%), statistical parameters (e.g. population size; 13%) and other data types (e.g. species counts, casual occurrence records; 21%). The main integration methods are statistical modelling (17.2%), extrapolation from limited data (6.4%), expert elicitation (19.8%) and others (~52%; e.g., aggregation of casual records at the grid level). The majority of dataflows occur with a frequency > 10 years (~54.2% of flows recorded): these mostly corresponded to integration initiatives that have generated products only once such as for example the Atlas of Amphibians and Reptiles or the IUCN Red list of birds; 30.2% of data flows are annual (e.g. PECMS, WISE-2 Biology data - see ***Boxes 1 - 4***) and around 9 % of flows occur with a frequency between 5- 10 years (e.g. harmonization of National Forest Inventory plots) or on a daily basis (4.5%). The latter are mainly flows of data with origin in data portals such as EBird, ornitho or Trektellen, all constituting the data basis of the EuroBirdPortal.

Only around 22 % of data flows have been documented to generate EBVs and EESVs: 55% of EBVs corresponded to measures of species distributions, 17% to measures of species abundance and 7% to measures of taxonomic /phylogenetic diversity, being EESVs less represented in the database (< 1 % of flows, all related to two European-level integration initiatives, the Mapping and Assessment of Ecosystems and their Services - MAES and the EFISCEN (European Forest Information SCENario Model). Regarding the indicators, these mostly corresponded to measures of species population trends (64% of indicators) and to the Biodiversity-related WFD indicators (Ecological Quality Rates for phytobenthos, phytoplankton, macrophytes, macroinvertebrates and fish)(16.2%) and measurements of conservation status. Figures 8 and 9 show how biodiversity data (by taxonomic

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

group) flows into the generation of European-level and National-level products (EBVs, EESVs, indicators), respectively (Table 1).

Regarding the spatial scale of the data flowing between nodes, as well as of final products, it is worth highlighting that ~8 % of the data processed flowing has a European-wide coverage and 30.8% a National Coverage. Almost 29 % of biodiversity data processed has a grid format with minimum spatial resolution of 1 km (e.g. Aggregated counts of butterflies belonging to the European Butterfly Monitoring Scheme) and a maximum resolution of 50 km (species distribution maps of the Second Atlas of European Breeding Birds - EBBA2). The lowest spatial resolution on ecosystems data is 100 m (Corine Land Cover data)

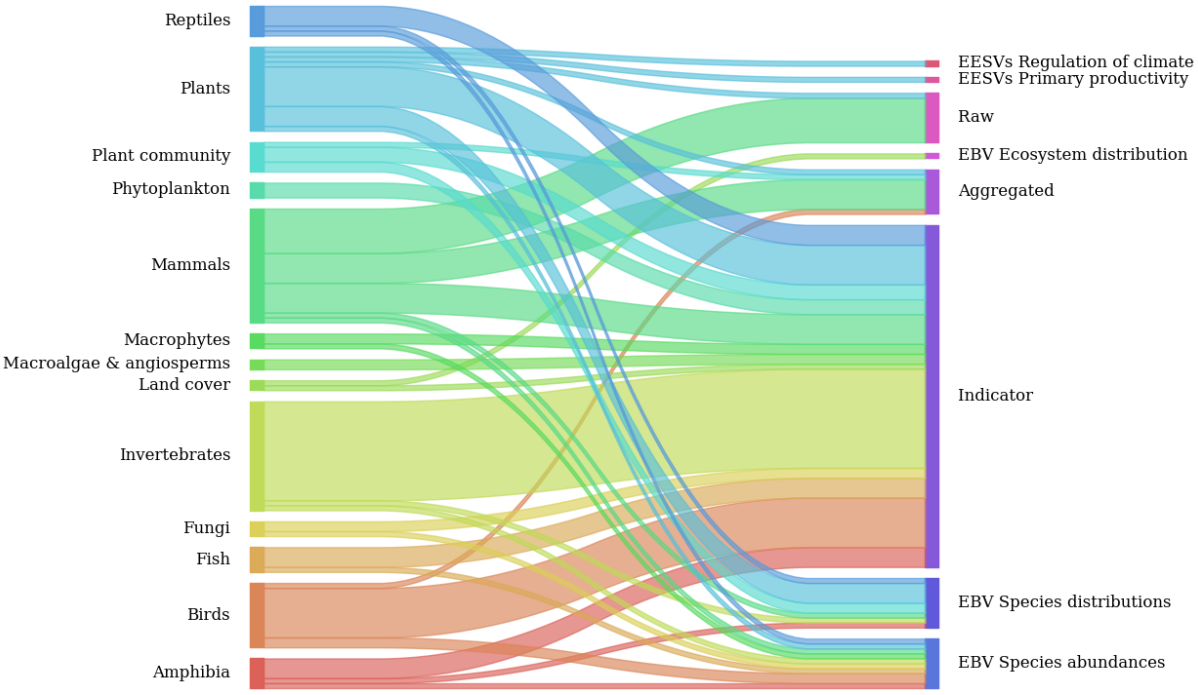


***Figure 8**. Data flows of raw, aggregated or processed data (EBVs, EEBs, or indicators) at the European level by taxonomic group*

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

Figure 9. *Data flows of raw, aggregated, or processed data (EBVs, EEBs, or indicators) at the National level by taxonomic group*

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

Table 1. Summary of European-level and grid-based products gathered in the EuropaBON database, detailed by taxa and product type (EBVs, EESVs, indicators)

Target taxa	Data type	European-level products (N; examples)	Grid-level products (N; examples)
Amphibia	EBV	N = 1; New Atlas of European Amphibians and Reptiles (NA2RE)	N = 1; Species richness
Amphibia	Indicator	N = 4; Cumulative numbers of non-indigenous species in European Seas; Temporal variability in numbers of new marine non-indigenous species per functional group, recorded in European Seas; Species trends; Species conservation status	N = 1; EU Ecosystem Assessment - Invasive Alien Species
Birds	Aggregated	N = 1; Migration pattern map	N = 2; New Migration Mapping Tool (MMT); EBP viewer
Birds	EBV	N = 1; Waterbird species-specific population size	N = 1; Species richness
Birds	Indicator	N = 9; DISTRIBUTION SHIFTS OF PLANT AND ANIMAL SPECIES - EEA - Copy; Species trends; All common birds trend indicator; DISTRIBUTION SHIFTS OF PLANT AND ANIMAL SPECIES - EEA; Species population trends; Farmland birds trend indicator; Forest birds trend indicator; Change Index - Europe; Species conservation status	N = 3; EU Ecosystem Assessment - Invasive Alien Species; Species breeding evidence - Europe (50km); Species distribution change - Europe (50km)
Fish	Indicator	N = 4; Species conservation status; Cumulative numbers of non-indigenous species in European Seas; Temporal variability in numbers of new marine non-indigenous species per functional group, recorded in European Seas; Species trends	
Fish	EBV		N = 1; Species richness
Fungi	Indicator	N = 2; Temporal variability in numbers of new marine non-indigenous species per functional group, recorded in European Seas; Cumulative numbers of non-indigenous species in European Seas	
Invertebrates	EBV	N = 1; European bee Red List	N = 5; Species richness; Species richness; Species richness; Species richness; Species richness
Invertebrates	Indicator	N = 20; DISTRIBUTION SHIFTS OF PLANT AND ANIMAL SPECIES - EEA - Copy; European Bees RedList; Species trends; Species conservation status; Species conservation status; Species trends; Species conservation status; Species conservation status; Species trends; DISTRIBUTION SHIFTS OF PLANT AND ANIMAL SPECIES - EEA; Cumulative numbers of non-indigenous species in European Seas; Woodland Butterfly Indicator; Widespread Butterfly Indicator; Urban Butterfly Indicator; Natura 2000 Butterfly Indicator; Species conservation status; Species trends; A Butterfly Indicator on climate change impacts; Temporal variability in numbers of new marine non-indigenous species per functional group, recorded in European Seas; Species trends	N = 1; Grassland butterfly indicator
Land cover	Indicator	N = 1; Ecosystem coverage - EEA Indicator	N = 1; Vegetation Productivity - Indicator assessment
Land cover	EBV		N = 2; CLC EU Changes; CLC EU Map

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

Target taxa	Data type	European-level products (N; examples)	Grid-level products (N; examples)
Land cover	EESVs		N = 1; Net Ecosystem Productivity
Macroalgae & angiosperms	Indicator	N = 2; Cumulative numbers of non-indigenous species in European Seas; Temporal variability in numbers of new marine non-indigenous species per functional group, recorded in European Seas	
Macrophytes	Indicator	N = 2; Temporal variability in numbers of new marine non-indigenous species per functional group, recorded in European Seas; Cumulative numbers of non-indigenous species in European Seas	
Mammals	Aggregated	N = 6; Hibernating Bats - 2. Regional trends; Hibernating Bats - 3. European trends; Hibernating Bats - 1. National trends; List of internationally important underground sites; Internationally important underground sites for bats; Underground sites with most bat species	
Mammals	EBV	N = 1; ENETWILD	N = 1; Mammal richness
Mammals	Indicator	N = 6; Species conservation status; Temporal variability in numbers of new marine non-indigenous species per functional group, recorded in European Seas; Species trends; Bat species dependence on overground roost types; Geographic pattern of dependence of bat species on overground roost types; Cumulative numbers of non-indigenous species in European Seas	N = 1; EU Ecosystem Assessment - Invasive Alien Species
Mammals	Raw	N = 1; Species distribution	
Phytoplankton	Indicator	N = 3; WISE-2 Dashboard; Temporal variability in numbers of new marine non-indigenous species per functional group, recorded in European Seas; Cumulative numbers of non-indigenous species in European Seas	
Plant community	Indicator	N = 2; Cumulative numbers of non-indigenous species in European Seas; Temporal variability in numbers of new marine non-indigenous species per functional group, recorded in European Seas	
Plant community	EBV		N = 1; Habitat suitability of acidophilous beech forests and base-rich fens
Plants	EBV	N = 2; European atlas of forest tree species; European Forest Information Scenario model (EFISCEN)	N = 2; Species richness; Species richness
Plants	EESVs	N = 1; European Forest Information Scenario model (EFISCEN)	
Plants	Indicator	N = 6; Temporal variability in numbers of new marine non-indigenous species per functional group, recorded in European Seas; Species trends; Cumulative numbers of non-indigenous species in European Seas; Species conservation status; Species conservation status; Species trends	N = 2; EU Ecosystem Assessment - Invasive Alien Species; Fire Weather Index
Plants	Aggregated		N = 1; Plant species distribution
Reptiles	EBV	N = 1; New Atlas of European Amphibians and Reptiles (NA2RE)	N = 1; Species richness

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

Target taxa	Data type	European-level products (N; examples)	Grid-level products (N; examples)
Reptiles	Indicator	N = 4; Species conservation status; Species trends; Temporal variability in numbers of new marine non-indigenous species per functional group, recorded in European Seas; Cumulative numbers of non-indigenous species in European Seas	N = 1; EU Ecosystem Assessment - Invasive Alien Species

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

Box 1. The Pan-European Common Bird Monitoring Scheme (PECBMS)

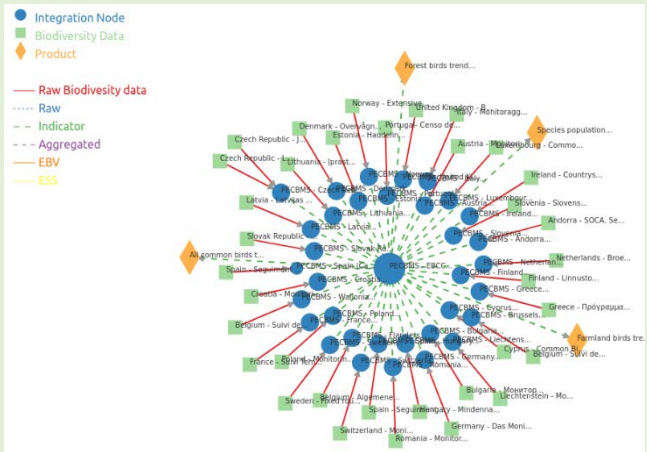
The Pan-European Common Bird Monitoring Scheme (PECBMS) is a European-level integration initiative whose main goal is to collect and harmonize data from large-scale and long-term bird monitoring schemes in European countries to generate indicators of the general state of nature. The common bird monitoring schemes that feed data to this initiative are mainly based on fieldwork of volunteers and follow a standardized methodology and formal design. Most of the monitoring schemes participating in this initiative started during the period 1980-1990.

The initiative is driven by NGOs (the European Bird Census Council (EBCC) and BirdLife International). It began motivated by the interest and enthusiasm of European ornithologists and their organisations across Europe to share knowledge across the continent and build together a robust platform to produce pan-European bird population trends and multi-species indicators on the state of nature with the final aim to improve bird conservation in Europe. The initiative was initially funded by ornithological NGOs but promptly received interest from the European Commission who has funded PECBMS since 2002 (3-year tender).

PECBMS integrates data from 34 national and sub-national integration nodes. The coordination unit of this initiative (European Integration node) is the Czech Society for Ornithology (CSO). This coordination unit integrates data on species population trends of common birds statistical modelling using TRIM) reported by all the national and subnational nodes on an annual basis, and calculates European-wide long-term trends of species populations (170 species), as well as a set of wild bird indicators (all common birds - 168 species; common farmland birds - 39 species and common forest birds - 34 species).

The multispecies population indices (indicators) produced by PECBMS have been the main outputs used for various policy purposes. Our European common bird indicators are included among the Streamlining European Biodiversity Indicators (SEBI), EU common bird indicators are part of the Indicators of Sustainable Development of the EU, and the European farmland bird indicator has been accepted as biodiversity indicator for EU´s Structural Indicator. National versions of the Farmland Bird Indicators (FBI) have also been approved as the Regulation indicators in the EU´s Rural Development Plans (Council Regulation (EC) No 1698/2005). The indicators produced by PECBMS have been used by other international institutions too, Organisation for Economic Co-operation and Development (OECD), United Nations Environment Programme (UNEP), European Environment Agency (EEA), Secretariat of the Convention on Biological Diversity, or European Court of Auditors among others. The indicators have also been included in the Living Planet Index (LPI). PECBMS also contributed the latest EU indicators to the last update of the European Red List of Birds, published in October 2021.

PECBMS is also a network of cooperation among organisations. As such, it provides a robust platform to periodically share and update knowledge on bird monitoring issues (from fieldwork design to statistical modelling) among its partners, which becomes of great value at national level to enhance the standardisation of multiple reporting tasks, such as that of the Birds Directive (Art. 12 reporting).



Network visualization of PECBMS data flows collected in the EuropaBON database

PECBMS data are regularly used by research institutes to address a series of topics of interest for policy makers, such as the development of indicators that track the impact of climate change on bird populations, or the effect protection in natura 2000 sites in decreasing the impact of farmland practices. As happened with the FBI, these developments can eventually provide official indicators of impact on biodiversity at national or international levels.

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

Box 2 - WISE-SoE Biology flow: a biodiversity data flow driven by European water policy

Biological data from aquatic ecosystems are collected and integrated by the European Environment Agency through the Water Information System for Europe. The data are collected on a yearly basis through the European Environment Information and Observation Network (EIONET), and used by EEA in State of the Environment (SoE) reports which are published every 5 years. The reporting to WISE-SoE is voluntary, while the reporting of ecological status to WFD is mandatory. Therefore, the WISE-SoE data are often sub-samples of national data assembled for the purpose of providing comparable indicators of pressures, state and impact of waters on a pan-European scale. The WISE-2 dataflow was established to obtain a harmonized flow of biology data reported as Ecological Quality Ratios (EQRs) from all surface water categories, rivers, lakes, transitional and coastal waters. In the context of biodiversity information, benefits of data reported to WISE-2 compared to WFD are (1) more frequent reporting: every year vs. every 6 years; (2) higher resolution: values on continuous scale (0-1) vs. categorical (5 status classes); (3) the determinants can be related to physical or chemical impacts (e.g. eutrophication vs. general degradation).

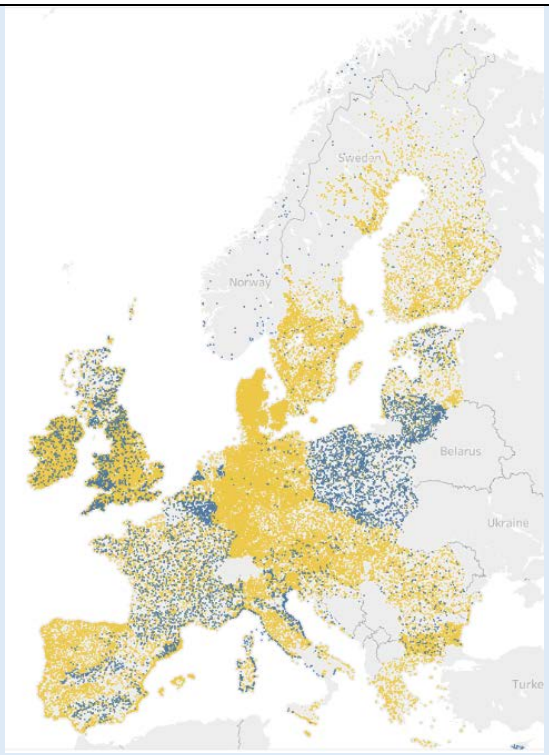
Biology data are requested for the following taxon groups (termed biological quality elements - BQEs) and water categories:

- phytoplankton in lakes, transitional and coastal waters
- phytobenthos in rivers
- macrophytes in lakes
- macroalgae in transitional and coastal waters
- angiosperms in transitional and coastal waters
- benthic invertebrates in rivers, transitional and coastal waters
- fish in rivers, lakes and transitional waters

The EQR value represents a measure of the deviation from reference conditions for each biological quality element. National EQR values are calculated for water bodies based on the value of a biological metric in relation to the reference value of the given metric (i.e. the expected metric value of water bodies with minimal anthropogenic impact), and reported to EEA according to the WISE-SoE Biology data dictionary. The national EQR values are subsequently converted to normalized EQR values which are comparable across countries, using information on the national classification systems. The original metric values are not reported to WISE, but information on the metrics used is reported together with the classification systems. The biological metrics typically represent a response to increasing pressure seen as a decrease of the sensitive taxa that usually dominate under reference conditions and an increase of tolerant taxa.

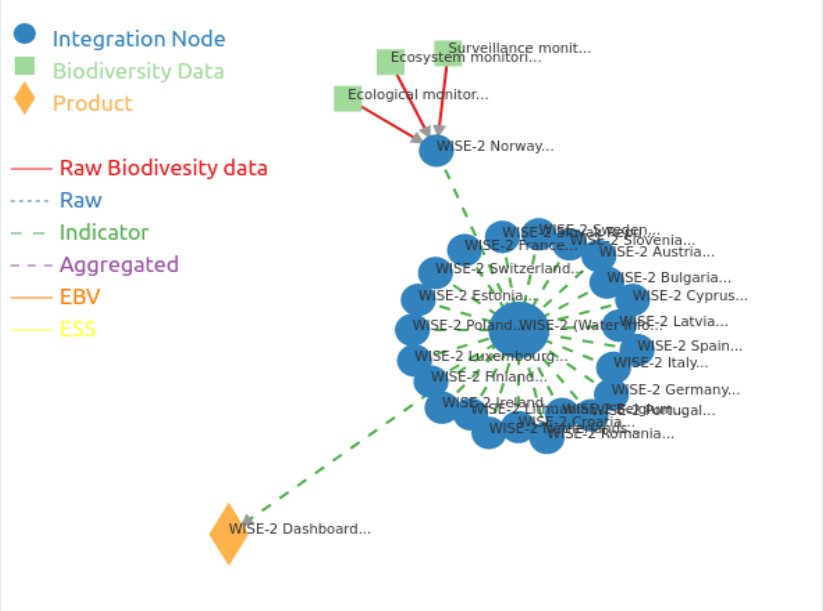
The quality-checked and final processed data are available from EEA's database Waterbase - Biology. The database currently holds freshwater biology data from countries from the reporting years 2011-2013 and 2015-2019, and marine biology data reported in 2019. The data are reported from more than 8000 waterbodies in 25 countries. The map shows the geographic distribution of water bodies reported to WISE-2 (blue) in comparison to those reported to WFD. (NB: the map is a draft version intended for internal use by EEA).

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



Spatial coverage of monitoring sites by the reporting WFD 2016 (yellow points) vs WISE-2 (all years; blue points)

Information from WISE-2 has been partially entered into the EuropaBON WP3.1 database, as illustrated in the figure below. The European-level integration node (WISE-2; by EEA) and the national-level integration nodes (national WISE-2 deliveries; typically, by national environmental agencies) have been completed with the information that is publicly available for all countries. The national integration node for Norway has been elaborated with more details on the underlying Biodiversity Data from three ecological monitoring programmes. We plan to contact the WISE-2 reporters from all countries to ask for more detailed information on the underlying biodiversity data, but currently this request is pending at EEA. Reporters have also been encouraged to report more of their existing monitoring data to WISE-2 to fill the gaps both in space, time and taxonomy (biological quality elements).



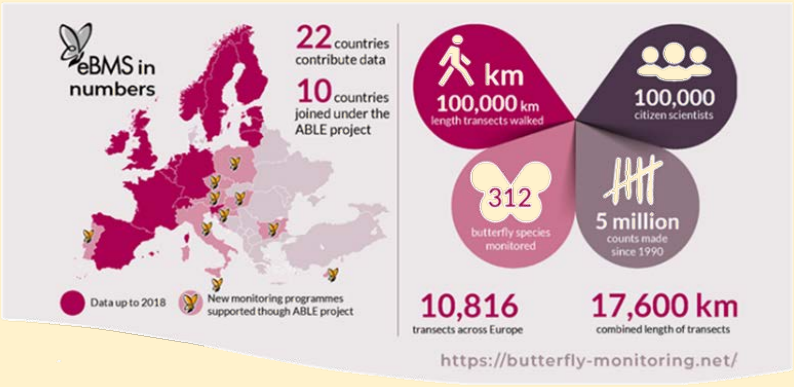
Network representation of WISE-2 data flows collected in the EuropaBON database

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

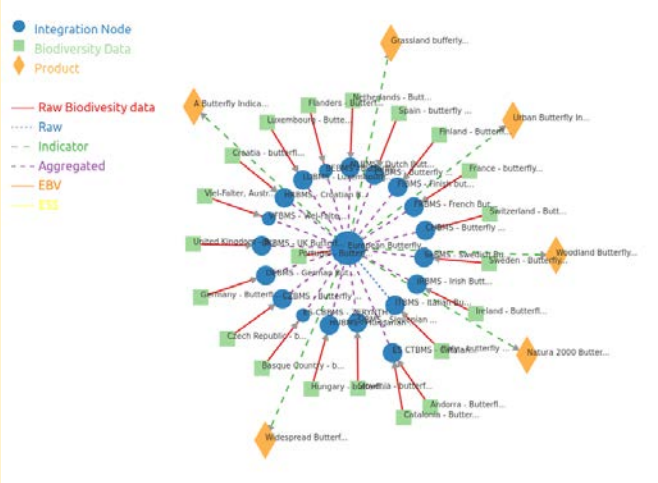
Box 3 - European Butterfly Monitoring Schemes (EBMs)

Monitoring butterflies is an important activity because butterflies are good indicators of biodiversity and their sensitivity makes them quick to react to environmental change. The butterflies of Europe breed in a broad range of habitats including grassland, woodland, wetland and montane areas. They are frequently used by ecologists as model organisms to study the impact of habitat loss and fragmentation, plus climate change. Areas rich in butterflies typically have a wide range of other invertebrates, so they are indicators of a healthy environment and well-functioning ecosystems. Butterflies and associated species form a key part of the food chain, they are an important part of the prey for among others birds and bats. Butterflies also pollinate a large range of flowering plants including vegetables, giving them economic importance.



Butterflies are one of the best monitored insect groups in Europe. Long-term monitoring has been running for decades in some countries. The first Butterfly Monitoring Scheme began in the United Kingdom in 1976 (UKBMS), since then, this methodology has been adopted in many other European countries to monitor butterflies. All use the same standardised survey method, fixed routes called transects, designed by Ernie Pollard, of the Monks Wood Experimental Station in the UK. Since 1990, many other countries have joined in and they now submit data to the European Butterfly. There are currently 22 countries contributing data. This is coordinated by Butterfly Conservation Europe and UK Centre for Ecology & Hydrology.

In several European projects, including the ABLE project, a lot of time and effort is invested in training local partners and developing tools for both monitoring and identification to facilitate local volunteers to monitor butterflies, especially in regions where there is little history of monitoring by citizen science. By using standardized methods with well-developed analyses the data can be combined and used for many different goals. Besides the calculation of pan-European and national species trends the data can be used for several indicators including the Grassland Butterfly Indicator, the Woodland Butterfly Indicator, the Urban Butterfly Indicator, the Natura 2000 Butterfly Indicator (comparing trends inside and outside of N2000 areas) and the Climate change indicator.

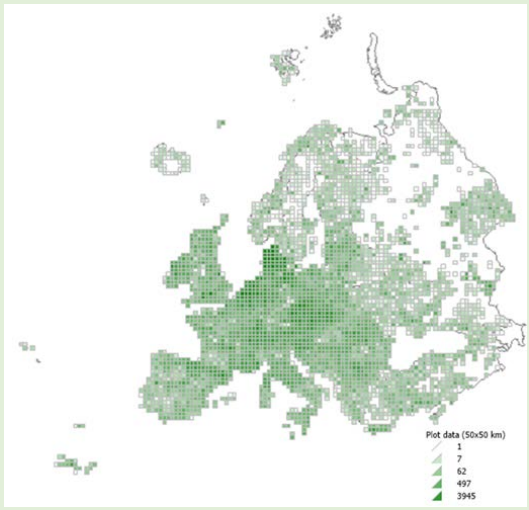


Network representation of EBMs data flows collected in the EuropaBON database

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

Box 4 - *Mapping habitat distributions at the European level: the European Vegetation Archive*

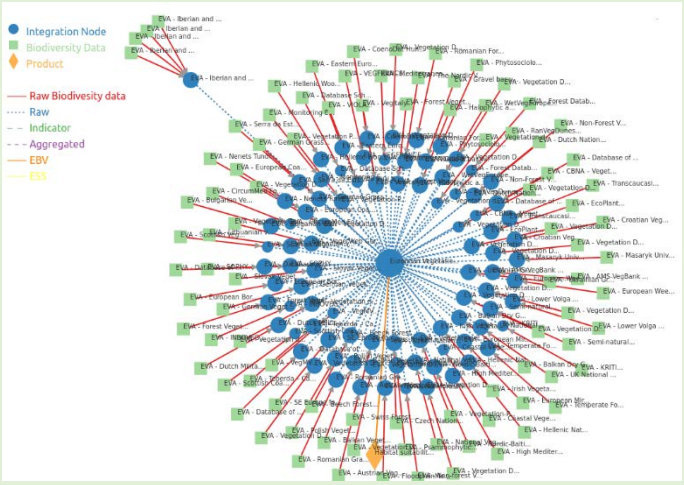
The European Vegetation Archive (EVA)[1] is an integrative database of vegetation plots across Europe. The purpose of EVA is to establish and maintain a single data repository of vegetation-plot observations (i.e. records of plant taxon co-occurrence at particular sites, also called phytosociological 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 comprises 99 national and supranational vegetation plots databases and contains 1,804,985 vegetation plots from 53 countries.



Number of vegetation plots included in EVA per 50 x 50 km (last update 05/05/2021)

The EVA data have been used for the generation of European-wide habitat maps (EBV as a measure of ecosystem´s distributions) in a series of publications of the European Environmental Agency [2,3,4]. Methodological issues have been tested through the development of community-level species distribution models and has allowed 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 at 1-km resolution for the entire Europe[5]. Further developments of the method suggest the possibility of predicting and mapping the "local area of occupancy" of habitat typologies of Conservation Community Interest (e.g., shrublands, pastures and different forest types) at smaller spatial resolutions (30 m) and extents thanks to the combination of spatial modelling and remote sensing data[6].

EVA data is mostly based on single surveys conducted over the last decades in Europe. However, the consortium is currently developing a parallel project to identify and gather vegetation-plot data with repeated measurements over time (ReSurveyEurope). The goal of this initiative is to allow a robust assessment of biodiversity trends in European vegetation. The database of ReSurveyEurope will have the same properties of EVA in terms of data contribution and data access. First analyses are expected to start in 2022.



Network representation of EVA data flows collected in the EuropaBON database

[1] Chytrý M. et al. (2016) *Applied Vegetation Science*, 19, 173–180; [2] Schaminée J.H.J., et al. (2014) Vegetation analysis and distribution maps for EUNIS habitats. Report EEA/NSV/14/006. EEA, Copenhagen.; [3] Schaminée J.H.J, et al. (2016) Development of distribution maps of grassland habitats of EUNIS habitat classification. Report EEA/NSV/16/005. EEA, Copenhagen.; [4] Schaminée J.H.J., et al. (2016) 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.; [5] Jiménez-Alfaro, B. et al. (2018) *Diversity and Distributions*, 24(7), 978-990; [6] Álvarez-Martínez, J.M., et al. (2018) . *Methods in Ecology and Evolution*, 9(3), 580-593.

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

6. Limitations, opportunities, and future directions

The EuropaBON monitoring database is the first publicly available database at the European scale providing detailed information on existing biodiversity and ecosystems data monitoring workflows across Europe. Describing how these workflows are currently organized in Europe, what data is collected and how it is integrated at the transnational level opens up the opportunity to build a more coherent biodiversity observation network across Europe. The analyses of this data set should allow EuropaBON to identify where current monitoring networks fail in mobilizing data (future task 3.2) and which bottlenecks currently prevent biodiversity information gathered in the field to be integrated and processed to relevant EBVs (future task 3.3).

The EuropaBON database does not cover all biodiversity data currently monitored across Europe. On the one hand, this relates to the fact that this task focused only on monitoring programs linked to integration initiatives leading to products at the European level (e.g., EBVs, ESS, Indicators). This **excluded a large volume of monitoring programs that currently do not go beyond data collection** or that get only integrated at a scale lower than European. Moreover, the task only partially explored the potential of data repositories (e.g., GBIF) that do not include classic monitoring programs to generate EBVs, EESv and indicators that could be used by policy. These r**epositories gather a large amount of biodiversity data collected via opportunistic observations across Europe that often get integrated at an academic level** (i.e., generation of indicators, EBVs and EESVs within the framework of scientific publications) making it difficult to trace. For example, the Pl@ntNet citizen science platform allows to share on the data repository GBIF a large volume of plant occurrences at the European scale (> 8.3 M records by November 2021). This is done with two complementary datasets, one of automatically identified occurrences (through photo identification by machine learning) and second on opportunistic observations. To date data from Pl@nNet has been used in more than 100 scientific publications, many of them generating species distributions for European plant species. In general, applications grounded in the academic environment that generate EBVs, EEVs or indicators - and regardless their data origin is in monitoring programs or opportunistic observations - are difficult to identify and map out and need to be better documented. We need frameworks to better categorize this type of information and eventually devise approaches to include it in the EuropaBON monitoring database.

The precise geographic coverage of current monitoring networks (i.e., the exact location of sampling sites) has proven to be very difficult to obtain for various reasons (e.g., data sensitivity and ownership). Moreover, when information about the location may be available, the heterogeneity in the spatial description of the sampling sites was very high. **Further efforts need to be developed in gathering spatial information on sampling locations for selected, key monitoring networks and exploring ways of representing this information in the database**. Within the EuropaBON project, examples in WP5 will be explored to cover this gap.

Monitoring efforts in Europe collecting data at the national or subnational level without further coordination at the transnational level are not well covered in the present database. A future European monitoring network will probably be better represented as a network of networks and the structure of the EuropaBON monitoring database is designed to represent this idea. In summary, the current information in the database would represent responsibilities focusing on European integration nodes and their key biodiversity data components, whereas the incomplete data at lower levels (integration nodes at the national level not connected to European nodes) would reflect flows at the national and subnational levels and a more dedicated effort here would be required in the future.

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

Opportunities and future directions

The database only partially covers the data flows of the Birds and Habitats Directives due to their complexity and the need to better liaise with National Agencies and stakeholders involved in data collection and integration (task now facilitated by the contact established through the workshops and consultation process made for the *EuropaBON User and Policy Needs Assessment - Deliverable 2.2*). Now that the contacts with the National Agencies and key stakeholders are well established and that the EuropaBON database structure is well set and has a fully functional interface (website), it will be easier to gather the information about the **data flows linked to the reporting of these Directives**. In particular, the WP3 will focus on mapping the data flows associated with these Directives in three or four countries, looking for commonalities, strengths, weaknesses, etc.

Besides the data flows linked to the EU Directives, EuropaBON will continue to work with its stakeholder network to **identify emerging European monitoring initiatives** to be gathered in the database. This joint work will allow to better quantify the comprehensiveness of the current database contents as well as to identify current gaps and bottlenecks of biodiversity monitoring across Europe in terms of thematic content, spatial and temporal coverage (WP3 - task 3.2), data flows and availability (WP3 - task 3.3), as well as cost effectiveness (WP3 - task 3.4).

Finally, the data included in the database and the novel visual approaches used to illustrate monitoring workflows could be further refined to **better communicate how biodiversity monitoring is used in Europe to produce relevant EBVs, EESVs and indicators**, and therefore help in articulating the necessary discussion behind the design of a future European Biodiversity network. In this regard, the monitoring EuropaBON database framework can be used to **describe new proposed workflows;** discussions are already in place to use this framework in the context of WP4 (task 4.3 Co-design of a European BON) and case studies within WP5 (i.e., BON in support of different European Directives) to illustrate and describe the new advances that EuropaBON will develop in the near future.

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

Annex 1. Data tables and fields in the EuropaBON monitoring database

Table A1. Data fields in the Integration Fields Form

Field name	Description
Name	Name of the integration activity/initiative
Institution	Who coordinates the node? (institution/s). Multiple institutions are allowed.
Geographic coverage	At what scale is the node integrating data? (e.g., EU, Biogeographic Region)
NUTS country and covered NUTS	If the geographic coverage is sub-National, this tab allows the selection of the corresponding NUTs levels 1-2-3 (multiple choice allowed)
Geographic coverage detail	Further details of the geographic scope. For example, if integration occurs at the Biogeographic region level, this field collect information about the specific regions covered (e.g., Mediterranean)
Website	Link to the website with info about the integration activity/node (if available).
Description	Brief description of the Integration Node (which institution is running the integration activity and what does it do exactly?). This field allows the collection of additional information including URL links to the institution, the integration activity or relevant documents or bibliographic references.
Coordinator	Name of the main contact person (for the specific integration initiative)
Coordinator email	Email address of the main contact person
Funding type	What type of funds support the integration activity? Options (multi-choice): public; private; unknown
Funding source	Who funds? main funding source. Options (multi-choice): EU; National; Regional; Scientific Grant; Private; Other; Unknown
Funding_program	Which program? main funding program. Options: 3 yrs tender; DG Environment; HORIZON 2020; Interrreg; LIFE; Other; None; Unknown
Funding_stability	Are funds stable in time? Options: Yes; No; Unknown
Funding_stability_detail	Approx. funds available for this activity in the last 5 years (€/year) (if available)
Node_purpose	What is the nature of the data integration node? Options: company bodies; Governmental; NGOs; Research bodies
Biodiversity_data	If the integration node collects and process raw data from monitoring programs, select here the BiodiversityData programs feeding this integration node. If this form contains information of an upper-level integration node (e.g., European-level node), ignore this field (the links between the European-level node and the BiodiversityData are generally indirect, through sub-European integration nodes). Remember to create the "BiodiversityData" forms to be able to fill this field.

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

Table A2. Data fields in the BiodiversityData Form

Field name	Description
Name	Name of the biodiversity monitoring initiative
Geographic coverage	At what scale is the node integrating data? (e.g., EU, Biogeographic Region)
NUTS country and covered NUTS	If the geographic coverage is sub-National, this tab allows the selection of the corresponding NUTs levels 1-2-3 (multiple choice allowed)
Geographic coverage detail	Further details of the geographic scope. For example, if integration occurs at the Biogeographic region level, this field collect information about the specific regions covered (e.g., Mediterranean)
Data origin	Select from the list: systematic monitoring, opportunistic observations, automatic recording, remote sensing, others
Data origin (details)	Brief description of the monitoring initiative: how the data is collected (e.g., sampling transects), what is the frequency of data collection, etc. This field contains details relevant for the dataflow that builds on this biodiversity data, including relevant URLs and bibliographic references
Purpose	What is driving the monitoring initiative? Options: Ecological Knowledge; Environmental Impact Assessment/Restoration Monitoring; Monitoring reference sites
Data type	Options: Abundance; Forest Inventory; Genetic; Presence-absence; Presence-only; Spectral Radiance
Data collection origin	Options: Citizen Science; Government; Professional; Unknown. Citizen science data includes data collected by volunteers of different backgrounds for example, most of the data uploaded to EBP comes from non-professional birdwatchers and the EBBA2 (European Atlas of Breeding Birds) trained volunteers over different countries to collect the data following the Atlas methodology. Both are examples of citizen science data. Government data includes monitoring programs run by employees of state/governmental agencies and public institutes such as the Nature Conservation Institute. Professional monitoring will include data collected by consultancies or NGOs for example.
Data accesibility	Is the data openly available? Options: Yes; No; Unknown
Data URL	Please, indicate the URL to access or request the data (if available)
Target real	Target real. Options: Terrestrial; Freshwater; Coastal; Marine
Target taxa	Target taxa (broad groups)
Target species	This field allows to upload a list of taxa monitored
Spatially explicit	Does the monitoring initiative collect/generate spatially explicit data? Options: Yes/No
Spatial resolution	What is the spatial scale of the data collected? Options: exact location; grid; other
Resolution	Spatial scale resolution value (numeric; for Grids only)
Resolution units	Spatial scale resolution units (km or m; for Grids only)
Number of sites	Indicate the approximate number of monitoring sampling sites covered
Temporal scale	One-off (occasional records) or temporal series
Temporal resolution	If time series is selected in the 'temporal scale' field, this field indicates the temporal resolution of the data
Start year	Since when is this monitoring initiative running? (oldest data records available for this monitoring program, regardless of how these data gets used in the integration initiative under consideration).
Ongoing	Is the monitoring initiative currently ongoing? Options: Yes/No
Last Year	Last date the database was updated

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

Table A3. Data fields in the DataStreams Form

Field name	Description
Data Origin (from where the data is coming from)	
Data Origin	Indicates the integration nodes from where the data flows
Data process – how is the data processed (if)? and how does it flow?	
Data type	Options: 1) Raw: data as collected in the field; 2) Aggregated: raw data is aggregated (not processed!); e.g., occurrence data collected in linear sampling transects of 1 km is aggregated (summed) at 10 km resolution, e.g., data from seasonal sampling is aggregated at annual level; 3) Essential Biodiversity Variables (EBVs); 4) Ecosystem Services values (EESVs); 5) Indicators: other processed indicators such as for example species trends
EBVs and ESS types	If the data flowing or generated by a given integration node is an EBV or an EESV, this field gives the option to select an option from a list of EBVs (Pereira et al. 2013) and EESVs (IPBES 2017) categories
Format	Format of the data flowing between biodiversity monitoring initiatives and integration nodes or between integration nodes. Options: conservation status; Map; Statistical parameter; Trend; Others
Method	If there is any type of data integration (i.e., if data type is other than Raw or Aggregated), this field details the method used to integrate data. Options: Best estimate; Expert; Extrapolation from limited amount of data; Statistically Robust Estimate; Statistical modelling; Others
Frequency	Frequency of the data flow (e.g., annual, seasonal)
Data set – which dataset/integrated product is processes and how	
Name	Name of the dataset
Description	Brief description of the data flowing (e.g., if the dataset represents aggregated data, this field contains information about how the aggregation is made); if the data set is a "product" – EBV, EESV, indicator – this field contains information about how its values are estimated and provide relevant bibliographic references or URL links if available
Target taxa	Target taxa
Target species	This field allows to upload a list of taxa for a data stream. If the DataStream generates a product integrating data for specific taxa, this field seeks to collect the list of taxa included in the product
Spatial scale	Spatial scale coverage of the dataset. Options: EU; Europe; Exact location; Grid; National; Subnational; Unknown; Other.
Resolution	Spatial scale resolution value (numeric)
Resolution Units	Spatial scale resolution units (km or m)
Data Destination - where does the data /product flow to?	
Does the data Flow into an upper-level integration node?	This field allows linking the data origin and destination nodes
The data is a generated product	Yes; No
Used by Policy	Is the indicator generated used by policy? Options: Yes; No

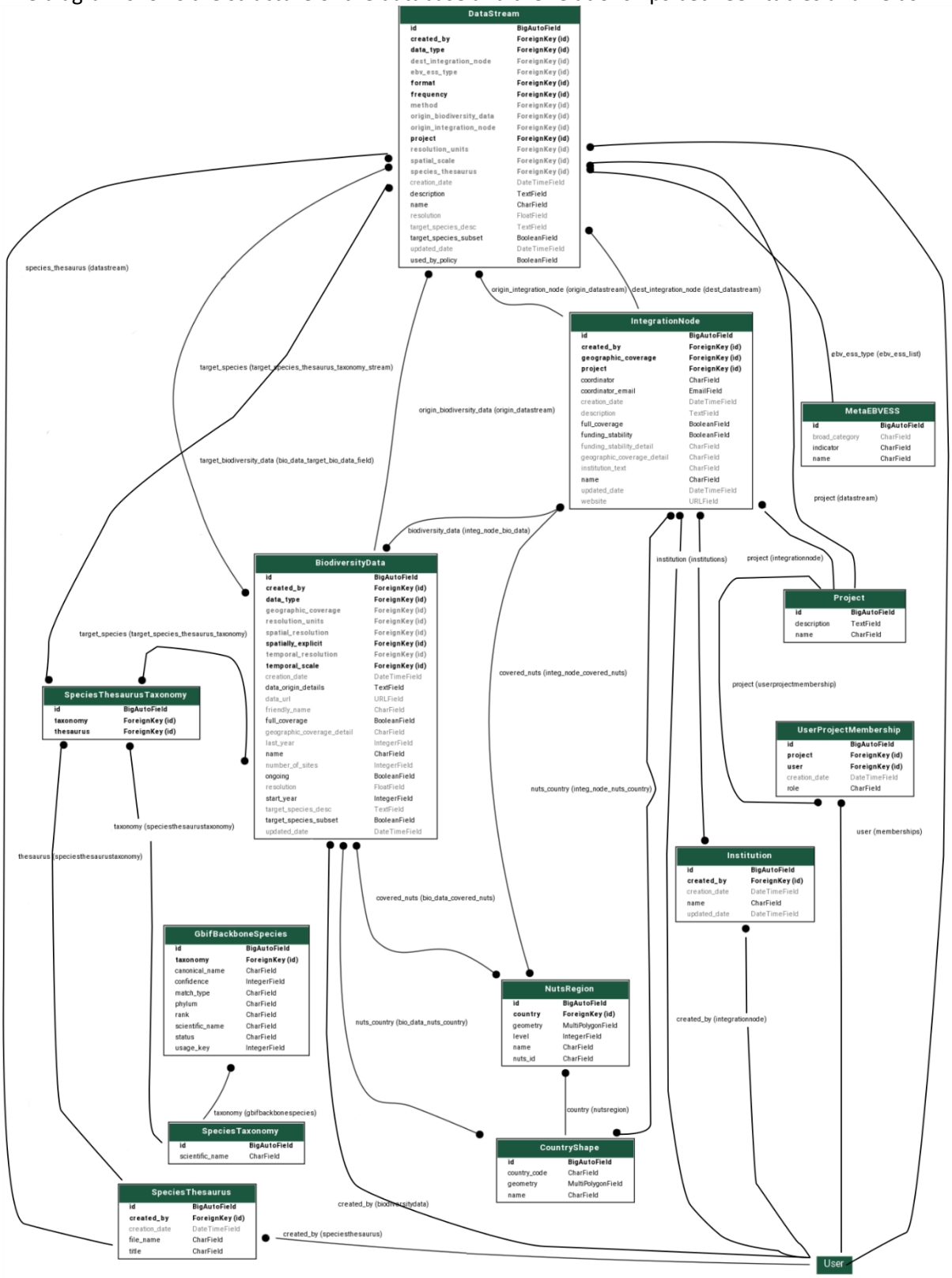
- IPBES. (2017). Update on the classification of nature's contributions to people by the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES/5/INF/24).
- Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H.M., Cardoso, A.C. and Coops, N.C., 2013. Essential biodiversity variables. Science, 339(6117), pp.277-278.

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

Annex 2. EuropaBON database conceptual model

The diagram shows the structure of the database and the relationships between tables and fields.



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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

Annex 3. Integration Initiatives gathered in the EuropaBON database by end November 2021

#	Integration initiative	Integration nodes	Data streams
1	Abundance and Distribution of marine mammals (ADMM) - OSPAR	2	8
2	Assessment of coastal habitats in relation to nutrient and/or organic enrichment - OSPAR	1	0
3	Atlas Florae Europaeae (AFE) - FMNH	1	1
4	Atlas of Marine Life - EMODnet Biology	1	0
5	Atlas of breeding birds of Catalonia - ICO	1	0
6	Atlas of breeding birds of Spain - SEO/Birdlife	1	0
7	BioTIME - University of St. Andrews	1	0
8	Bird Migration Mapping Tool - EURING	1	0
9	Birds directive	2	9
10	CORINE Land Cover (CLC) - Copernicus programe	40	40
11	Catalan LPI	1	0
12	Cetacean Offshore Distribution and Abundance in the European Atlantic (CODA) - University of St. Andrews	1	3
13	Change in Average Trophic Level of Marine Predators in the Bay of Biscay - OSPAR	1	0
14	Changes in Plankton Diversity (CPD) - OSPAR	4	6
15	Changes in phytoplankton and zooplankton communities (CPZC) - OSPAR	1	0
16	Changes in phytoplankton biomass and zooplankton abundance (CPBZA) - OSPAR	5	6
17	Climatic Risk and Distribution Atlas of European Bumblebees (CRDAEB)	22	23
18	Distribution shifts of plant and animal species - EEA	1	0
19	EBBA2	52	108
20	EBP	34	43
21	EFISCEN (European Forest Information SCENario Model) - EFI	10	29
22	ENETWild - IREC	6	7
23	EU-PoMS	1	0
24	Ecosystem mapping and assesment (EMA) - MAES	2	7
25	Eurasian African Bird Migration Atlas - EURING	1	0
26	European Alien Species Information Network (EASIN) - JRC	1	0
27	European Bat Population Trends	16	22
28	European Forest Fire Information System (EFFIS) - JRC	1	0
29	European Red List of amphibians (ERLA) - IUCN	3	5
30	European Red List of bees (ERLBE) - IUCN	1	3
31	European Red List of birds (ERLB) - Birdlife	23	69
32	European Red List of butterflies (ERLBU) - IUCN	47	49
33	European Red List of dragonflies (ERLD) - IUCN	1	3
34	European Red List of freshwater fishes (ERLFF) - IUCN	4	6
35	European Red List of medicinal plants (ERLMP) - IUCN	1	4
36	European Red List of non-marine molluscs (ERLNMM) - IUCN	1	3
37	European Red List of reptiles (ERLR) - IUCN	2	5
38	European Red List of saproxylic beetles (ERLSB) - IUCN	47	49
39	European Red List of vascular plants (ERLVP) - IUCN	1	3
40	European Vegetation Archive (EVA) - IAVS	91	91
41	European atlas of forest tree species - EC	45	46
42	Forest fires in Europe (FFE) - EEA	2	4
43	GBIF - Global Biodiversity Information Facility	1	0
44	GloBAM	2	1
45	Grey seal pup production (GSPP) - OSPAR	7	7

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

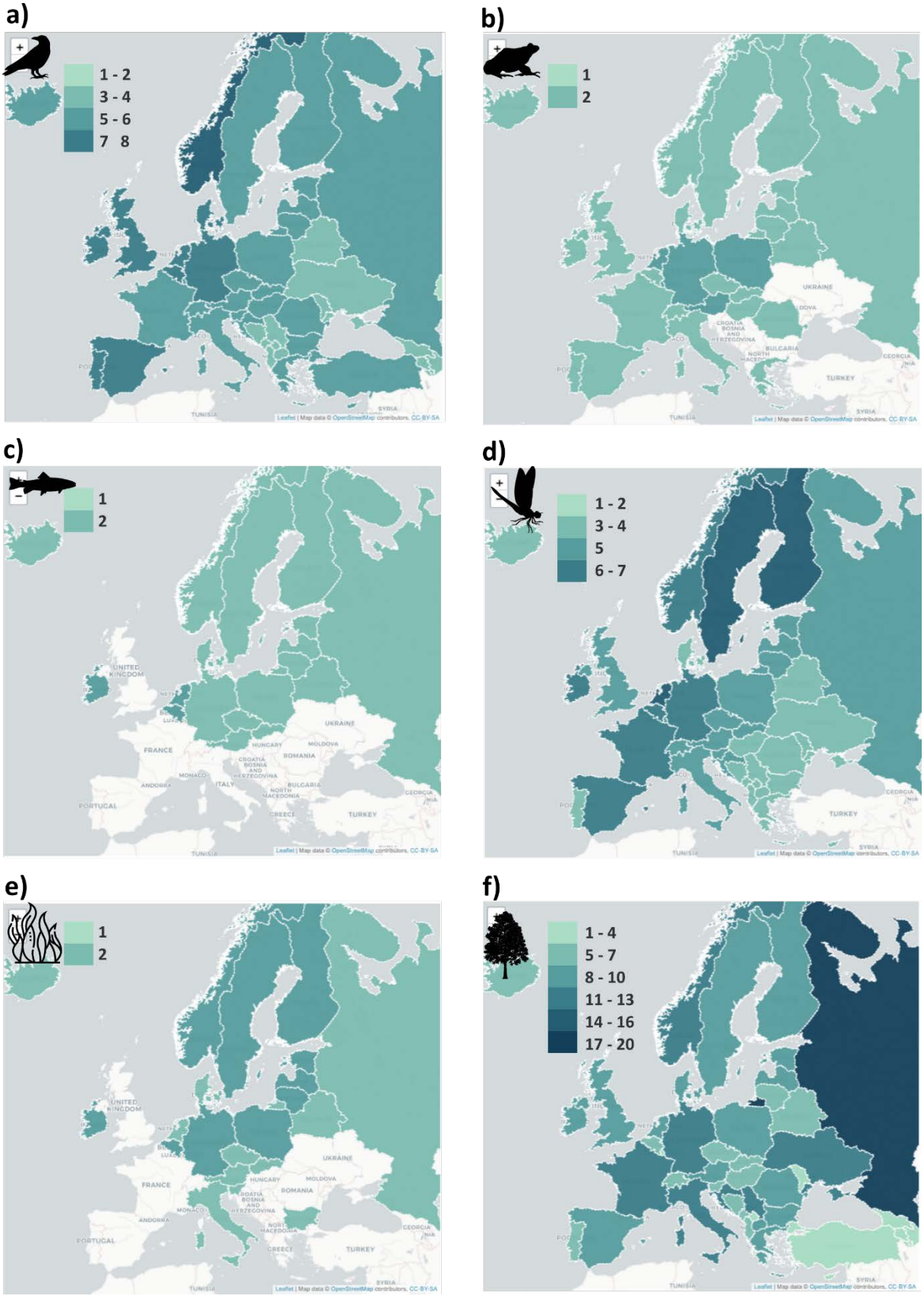
#	Integration initiative	Integration nodes	Data streams
46	Harbour porpoise abundance (HHPA) - HELCOM	11	19
47	Health status (HHS) - HELCOM	2	2
48	International Waterbird Census	44	47
49	Marine bird abundance (MBA) - OSPAR	13	13
50	Marine bird breeding success/failure (MBBSF) - OSPAR	13	13
51	Marine breeding birds abundance and distribution (HMBBAD) - HELCOM	11	22
52	New Atlas of Amphibians and Reptiles of Europe (NA2RE) - SEH	21	22
53	North Atlantic Sighting Surveys (NASS) - NAMMCO	1	0
54	Norwegian Independent Line Transect Surveys (NILS) - Institute of Marine Research Bergen	1	1
55	PECBMS	35	38
56	Phenology of plant and animal species - EEA	1	0
57	Proportion of large fish (Large Fish Index) - OSPAR	1	0
58	Recovery in the population abundance of sensitive fish species - OSPAR	1	0
59	STEP Project	7	10
60	Seal abundance (HSA) - HELCOM	7	13
61	Seal abundance and distribution - OSPAR	1	3
62	Second European Mammal Atlas (EMMA2) - EMF	34	33
63	Size composition in fish communities - OSPAR	1	0
64	Small Cetaceans in European Atlantic waters and the North Sea (SCANS) - University of St. Andrews	1	4
65	Status and distribution of European mammals (EMA) - IUCN	7	9
66	Status of marine fish and shellfish stocks in European seas - EEA	1	0
67	Subtidal habitats of the Southern North Sea - OSPAR	1	0
68	Trektellen - BTO/ICO/SOVON/Birdlife Flanders/Birdlife Netherlands/Aves-Natagora	1	3
69	Trends in marine non-indigenous species (TMNIS) - EEA	9	11
70	Trends in new records of non-indigenous species introduced by human activities - OSPAR	1	0
71	University of Amsterdam Bird Tracking System (UvA-BiTS)	1	0
72	Vegetation productivity (VP) - EEA	1	1
73	WISE-2	36	25
74	eBMS	20	25

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 10/07/2023*. DOI: https://doi.org/10.3897/arphapreprints.e109168

Annex 4. Geographic coverage of biodiversity data by taxonomic group (examples)

Number of national integration nodes in the EuropaBON database by end November 2021; a) Birds; b) Amphibians; c) Fish; d) Invertebrates; e) Macrophytes and f) plants and plant communities.



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