

24/11/2025

Version 1.0



## MS24 Concept to integrate use cases based on the Regulation on Invasive Alien Species available

Author(s): Namsrai Jargal (SGN), Peter Haase (SGN)

Contributor(s): Julian Oeser (UFZ), Taimur Khan (UFZ), Ingolf Kühn (UFZ), Alessandro Oggioni (CNR), Caterina Bergami (CNR), Zhiming Zhao (LifeWatch)



Co-funded by  
the European Union

### Project funded by



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs,  
Education and Research EAER  
**State Secretariat for Education,  
Research and Innovation SERI**

BMD (Biodiversity Meets Data) receives funding from the European Union's Horizon Europe Research and Innovation Programme and the Swiss State Secretariat for Education, Research and Innovation (SERI) (ID No 101181294). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union, the European Research Executive Agency (REA) or SERI. The EU, REA and SERI cannot be held responsible for them.

**Prepared under contract from the European Commission**

Grant agreement No. 101181294

EU Horizon Europe Research and Innovation Action

Project acronym:	<b>BMD</b>
Project full title:	<b>Biodiversity Meets Data</b>
Project duration:	01.03.2025 – 28.02.2029 (48 months)
Project coordinator:	Stichting Naturalis Biodiversity Center (Naturalis)
Call:	HORIZON-CL6-2024-BIODIV-01
Milestone title:	CONCEPT TO INTEGRATE USE CASES BASED ON THE REGULATION ON INVASIVE ALIEN SPECIES AVAILABLE
Milestone n°:	MS24
Means of verification:	Concept
Work package:	WP5
Nature of the milestone:	Concept
Contribution to deliverable n°:	D5.2
Licence of use:	Not applicable
Lead beneficiary:	Senckenberg Society of Nature Research (SGN)
Recommended citation:	Jargal, N., Haase, P., Oeser, J., Khan, T., Kuehn, I., Oggioni, A., Bergami, C., & Zhao, Z. (2025). <i>Concept to integrate use cases based on the Regulation on Invasive Alien Species available</i> . BMD project milestone MS24.
Due date of milestone:	November 30, 2025 <b>(M9)</b>
Actual submission date:	Month n°
Quality review:	Yes

**Milestone status:**

Version	Status	Date	Author(s)	Actions
0.1	Draft	September 11, 2025	Namsrai Jargal (SGN), Peter Haase (SGN), Julian Oeser (UFZ), Taimur Khan (UFZ), Ingolf Kuehn (UFZ), Alessandro Oggioni (CNR), Caterina Bergami (CNR), Zhiming Zhao (LifeWatch)	Sent for review



0.2	Draft	September 29, 2025	Namsrai Jargal (SGN), Peter Haase (SGN) Julian Oeser (UFZ), Taimur Khan (UFZ), Ingolf Kuehn (UFZ), Alessandro Oggioni (CNR), Caterina Bergami (CNR), Zhiming Zhao (LifeWatch)	Reviewed
0.3	Draft	October 31, 2025	Namsrai Jargal (SGN), Peter Haase (SGN) Julian Oeser (UFZ), Taimur Khan (UFZ), Ingolf Kuehn (UFZ), Alessandro Oggioni (CNR), Caterina Bergami (CNR), Zhiming Zhao (LifeWatch)	Finalised, with incorporation of feedback from reviewers
1.0	Final version	November 24, 2025	Namsrai Jargal (SGN), Peter Haase (SGN) Julian Oeser (UFZ), Taimur Khan (UFZ), Ingolf Kuehn (UFZ), Alessandro Oggioni (CNR), Caterina Bergami (CNR), Zhiming Zhao (LifeWatch)	Submitted



## 1. Table of contents

1.	Table of contents .....	4
2.	Summary .....	5
3.	List of abbreviations.....	5
4.	Conceptual framework .....	6
4.1.	Purpose .....	6
4.2.	Scope.....	6
4.3.	Introduction (Why IAS are a problem).....	6
4.4.	Regulatory background.....	7
4.5.	Available resources .....	8
4.6.	Potential integrated use cases across realms .....	9
4.7.	Implementation .....	10
4.8.	Data resources .....	10
5.	Design of a Virtual Research Environment .....	11
5.1.	VRE design workflow.....	11
5.2.	VL list.....	12
6.	Technical implementation .....	19
6.1.	Spatial analyses .....	19
6.2.	Temporal analyses .....	20
6.3.	User interface and outputs .....	21
7.	Conclusion and future steps .....	21
8.	Acknowledgements.....	22
9.	References .....	22



## 2. Summary

This document corresponds to milestone MS24 and presents the conceptual and technical design for integrating use cases of invasive alien species (IAS) of Union concern into a Virtual Research Environment (VRE). Based on the EU Regulation 1143/2014/EC (IAS Regulation), it focuses on key aspects (presence status, distributions, and temporal trends) that can be addressed using spatiotemporal data and modelling. The framework uses the list of species of Union concern, addressing the EU regulation mentioned above, and authoritative species checklists (EASIN and GRIIS) to define the taxa of interest, whose occurrences are then sourced from GBIF and realm-specific systems (OBIS and monitoring datasets). These data are combined with environmental predictors in spatial and temporal analyses to generate outputs such as distribution maps, trends, and potential invasion risk assessments. These outputs support monitoring, reporting, and management at Member State and EU-wide scales, linking legal obligations and stakeholder needs with data and VRE use cases (Virtual Laboratories) in a reproducible, policy-relevant manner.

## 3. List of abbreviations

IAS	Invasive alien species
EASIN	European Alien Species Information Network
EMODnet	European Marine Observation and Data Network
EU	European Union
GBIF	Global Biodiversity Information Facility
GRIIS	Global Register of Introduced and Invasive Species
IUCN	International Union for Conservation of Nature
NGOs	Non-Governmental Organizations
OBIS	Ocean Biodiversity Information System
SDM	Species distribution model
VL	Virtual Laboratory
VRE	Virtual Research Environment



## 4. Conceptual framework

### 4.1. Purpose

The overarching goal of this milestone is to define a conceptual and technical framework for integrating use cases of invasive alien species (IAS) of Union concern into a Virtual Research Environment (VRE), thereby supporting regulatory implementation, monitoring, and management through reproducible data flows and analyses. The VRE is built around specific Virtual Laboratories (VLs), each focused-on species distribution and temporal trends across freshwater, terrestrial, and marine realms. These VLs provide platforms for IAS-specific analyses and are co-designed with stakeholders (NATURA 2000 site managers, EU Member states and associated countries, NGOs) to address specific management needs.

### 4.2. Scope

- Realms: Terrestrial, freshwater, and marine ecosystems
- Spatial scales: Regional (incl. Natura 2000 sites), national, EU-wide assessments
- Biological levels: Species

**Link to EU IAS Regulation:**

<http://data.europa.eu/eli/reg/2014/1143/oj>

**Link to EU List of IAS of Union concern:**

<https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02016R1141-20250807>

**Link to EASIN**

[https://environment.ec.europa.eu/topics/nature-and-biodiversity/invasive-alien-species\\_en](https://environment.ec.europa.eu/topics/nature-and-biodiversity/invasive-alien-species_en)

**Link to relevant national checklists (GRIIS):**

<https://griis.org/>

### 4.3. Introduction (Why IAS are a problem)

IAS are organisms that are introduced intentionally or unintentionally into regions where they do not naturally occur. Once established, they can have severe negative consequences for biodiversity, ecosystems, and human well-being. The IAS impacts extend across multiple dimensions <sup>[1,2]</sup>:

*Environmental impacts*

- IAS are one of the five major drivers of biodiversity loss in Europe and worldwide <sup>[1,3]</sup>.
- They threaten native species, disrupt ecological interactions, and degrade habitats <sup>[2,4]</sup>.
- Their presence can reduce ecosystem resilience and the ability to provide essential services such as clean water, fertile soils, and pollination <sup>[2,5,6]</sup>.



#### *Economic impacts*

- IAS impose heavy financial burdens on agriculture, forestry, fisheries, and infrastructure <sup>[7,8]</sup>.
- In the European Union, their overall economic impact is estimated at around EUR 12 billion per year, a figure that continues to rise <sup>[1]</sup>.

#### *Human health impacts*

- Certain IAS affect public health by causing severe allergies, skin burns, or act as vectors or reservoirs of disease <sup>[2,6]</sup>.
- These health effects can increase healthcare costs and reduce quality of life <sup>[2]</sup>.

Given their consequences, effective monitoring, management, and prevention mechanisms targeting IAS are necessary <sup>[2,9,10]</sup>. Expedited detection and timely response can limit ecological and economic damage, while concerted effort by Member States avoids new introductions and diminishes the distribution of established species <sup>[10,11]</sup>. A virtual integration of species occurrence data, environmental predictors, and modeling frameworks can facilitate evidence-based decision-making, prioritize management actions, and inform stakeholders on existing and potential threats, and ultimately support the conservation of biodiversity, ecosystem services, and human health.

## 4.4. Regulatory background

The IAS Regulation (EU Regulation 1143/2014/EC and implementing acts) establishes the Union list and requires monitoring, prevention, and management measures at the Member State level <sup>[1,12]</sup>. Implementing rules and updates to the Union list have been published and remain active policy drivers for coordinated monitoring and reporting <sup>[1,13]</sup>. Effective implementation requires turning lists and legal obligations into operational use cases: (a) where species are (distribution), (b) where they are spreading (trends/range expansion), (c) which sites are at risk (connectivity/propagule pressure), and (d) how management measures perform. EASIN and other EU services already curate and aggregate IAS information that the VRE can reuse.

The IAS Regulation mandates monitoring systems and official controls. Commission Implementing Regulation (EU) 2024/574 <sup>[13]</sup> specifies technical formats for reporting (Article 24(1)), including:

#### **For each IAS of Union or regional concern:**

- Species information: Scientific name and optionally common name.
- Presence status: Current presence, confirmed absence, presence in contained holdings, or uncertain/unknown.
- Distribution: Spatial mapping of species presence across Member States, including marine ecosystems.
- Reproductive status and patterns: Self-sustaining, limited, or occasional reproduction, and reproductive mode (sexual, asexual, both, unclear).



- Introduction and spread patterns: Predominant introduction and spread mechanisms during the reporting period.
- Permits issued: Research, ex-situ conservation, medicinal, or other authorized activities.
- Rapid eradication and management measures: Geographic coverage, methods, effectiveness, and monitoring.
- Impacts (optional): Effects on biodiversity, ecosystem services, protected sites, human health, or the economy.
- Involvement in regional listing: Participation in listing IAS of regional concern.

**Across species and activities:**

- Pathway action plans: Analysis and prioritization of unintentional introduction pathways, status of action plans.
- Monitoring system: Description of coverage and use in monitoring eradication or control effectiveness.
- Official control system: Measures to prevent intentional introductions, including competent authorities.
- Public information and participation: Communication and engagement measures for citizens.
- Cost of actions: Costs or efforts for compliance, potentially broken down by category, including cost recovery systems.

Reporting occurs every six years. Surveillance systems must dynamically detect newly appearing IAS and build upon existing monitoring systems (e.g., Habitats, Water Framework, Marine Strategy Framework Directives). Official controls verify goods and inspect establishments with permits for research or ex-situ conservation.

Among all parameters that require monitoring, the following will be addressed by the VRE based on a spatial analysis and modeling of species occurrence data. Specifically, these include:

- Species information
- Presence status
- Current and future distributions

Since reporting under the IAS Regulation is mainly happening at the national scale, providing summaries of results produced by the VRE at the Member state level might be useful. We will also collaborate with the OneSTOP project partners ([onestop-project.eu](https://onestop-project.eu)) to align modeling efforts and maintain consistency.

## 4.5. Available resources

The IAS Portal (<https://ias.eea.europa.eu/>) is the European Environment Agency platform supporting the implementation of the EU IAS Regulation. Designed for mandatory reporting, it hosts information from the 2015-2018 reporting period and provides key resources such as the Union species list, data viewers





for spatial information, and national summaries from Member States. The most crucial piece of information in this respect is the list of IAS of Union concern

<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02016R1141-20220802&from=EN>

EASIN ([easin.jrc.ec.europa.eu/easin/](https://easin.jrc.ec.europa.eu/easin/)) serves as the European Commission's Joint Research Centre's central platform for IAS of Union concern data and provides authoritative information crucial for implementing the EU IAS Regulation. Its core offerings include the official and updated IAS list of Union concern, detailed species fact sheets with taxonomy, ecology, and introduction pathways, and interactive spatial tools that map species observed distributions over time and across land cover and use. Furthermore, EASIN catalogs historical introduction data, making it an important resource for establishing baselines, tracking spread, and supporting the monitoring and reporting obligations of Member States.

GRIIS (<https://griis.org/>), managed by the IUCN, presents validated and verified national and subnational checklists of alien and invasive alien species.

OBIS (<https://obis.org>) is a global open-access data platform that aggregates and shares information on the diversity, distribution, and abundance of marine life. Established under the Intergovernmental Oceanographic Commission, OBIS provides standardized, quality-controlled data from research, monitoring, and citizen science initiatives. It supports marine biodiversity assessments, ecosystem management, and global initiatives such as the Marine Strategy Framework Directive and the Convention on Biological Diversity.

OneSTOP ([onestop-project.eu](https://onestop-project.eu)) seeks to minimize terrestrial IAS impacts by integrating novel detection (eDNA, AI) with a data-driven prioritization system. It develops an automated pipeline for rapid IAS data dissemination, which includes streamlining data-sharing workflows, automating the publishing of national and European GRIIS checklists, generating timely occurrence alerts, and ensuring data harmonization. Also, it leverages predictive models and future scenarios to map species distributions, integrate real-time data for dynamic forecasts, scan for emerging threats, assess impacts on Red List species, and strategically guide resource allocation.

GuardIAS ([guardias.eu](https://guardias.eu)) is a sister project of OneSTOP, which focuses on aquatic systems (marine and freshwater) and applies a similar integrated approach to aquatic IAS.

## 4.6. Potential integrated use cases across realms

- **Species information**

Generate counts and trends of IAS per Member State and Natura2000 site. Provide standardized identifiers (scientific names, with optional common names) to ensure compatibility with regulatory reporting formats. These outputs can be reused directly for Article 24 reporting<sup>[13]</sup> and allow stakeholders to track changes in IAS over space and time.

- **Presence status**

Assess the current state of each species: confirmed presence in the environment, presence only in contained holdings (e.g., aquaculture, botanical gardens), or uncertain/unknown status. VRE



workflows can combine authoritative lists (EASIN/Union list and GRIIS) with GBIF, OBIS, and monitoring datasets to produce up-to-date presence grids and state-level summaries.

- **Distributions**

Map both the realized and potential distributions of IAS. Current distributions can be inferred from occurrence data, while potential distributions can be modeled using species distribution models (SDMs) under present and future climate conditions. These outputs help identify at-risk habitats, assess invasion stages (emerging vs. widespread), and guide prioritization of management efforts.

- **Temporal trends**

Assess the temporal trends of IAS at the levels of different spatial units (Natura 2000 sites, Member states). Trends can be calculated based on “raw” occurrence data (accounting for spatiotemporal trends in sampling effort), or, if data availability and quality permit, based on, for example, occupancy models.

## 4.7. Implementation

The potential VRE and its VLs will be implemented using the R programming language, which provides extensive libraries for biodiversity informatics, spatial analysis, and species distribution modeling. User-facing components will be delivered through interactive dashboards, offering an intuitive graphical user interface (GUI).

The design of the VRE will strictly follow the FAIR principles (Findable, Accessible, Interoperable, Reusable) to ensure long-term scientific and policy relevance:

- Rich and standardized metadata: All datasets and workflows will include detailed provenance documentation, specifying data source, date, version, and spatial/temporal resolution.
- Data governance and quality assessments: Built-in checks will ensure data consistency and reliability by adhering to data standards of authoritative systems (e.g., EASIN, GRIIS, OBIS, and GBIF), where such standards are available and applicable.
- Reproducible workflows: Version-controlled scripts and containerized environments (e.g., via Docker) will enable fully repeatable and auditable analyses, supporting trustworthiness for regulatory reporting. A workflow engine will centralize workflow runs and act as a compute backend to the VRE.

Additionally, Virtual Laboratories (VLs) in the VRE, based on different use cases of IAS of Union concern, will be scalable, from individual Natura 2000 sites to EU-wide analyses, and interoperable with existing EU infrastructures such as EASIN, GBIF, and Copernicus. This ensures that outputs can be directly used for both regulatory compliance and stakeholder decision support.

## 4.8. Data resources

### Authoritative checklists

- EASIN / EU list: Authoritative records for Union concern species



- GRIIS: National checklists for invasive species

The EU-level and national-level checklists for invasive species need to be ingested into the BMD data catalogue to consistently identify relevant species when running analyses at the levels of Natura200 sites, within Member states, or across Member states. For spatial and temporal analysis, the Union list will be filtered to include only species with a confirmed presence in Europe for which sufficient occurrence data is available.

#### Species occurrence data

- GBIF and OBIS: Occurrence records for species defined by the authoritative checklists.
- Subsets of monitoring datasets: Relevant species records in monitoring datasets where available.

#### Environmental predictor data

- CHELSA Bioclimatic variables
- CORINE land cover/use data
- Other realms/taxon-specific environmental factors

## 5. Design of a Virtual Research Environment

### 5.1. VRE design workflow

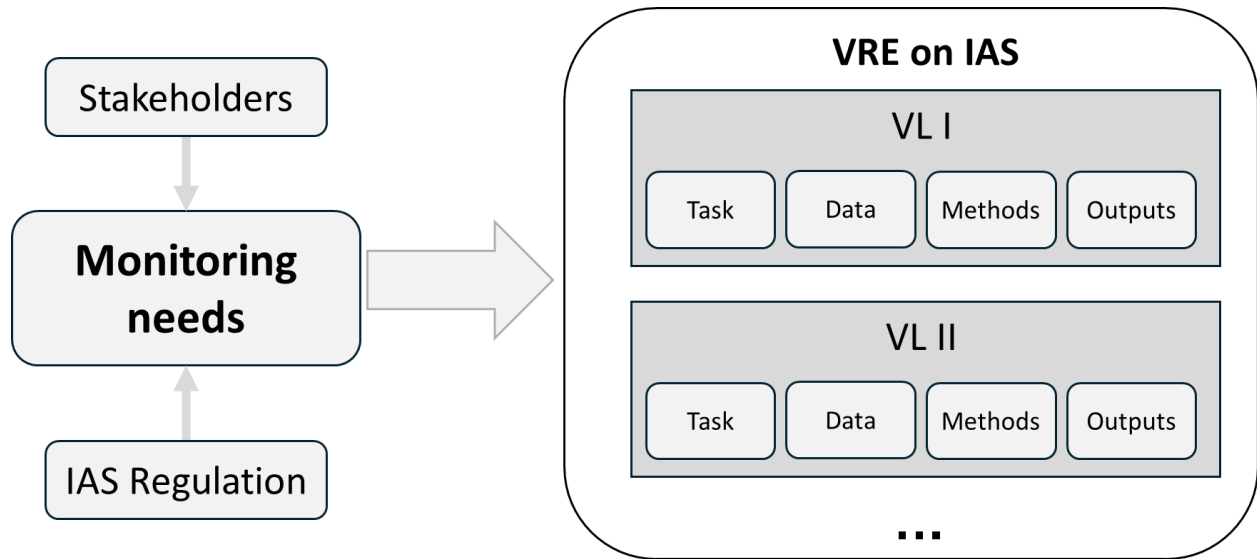
Fig. 1 illustrates the conceptual framework used for defining the VRE and its VLs, focused on IAS of Union concern <sup>[14]</sup>. The workflow begins by defining monitoring requirements based on:

- Input from relevant stakeholders. Stakeholder input will be gathered through co-design workshops, for example, in the form of user stories, and will inform workflow design, user interfaces, and data presentation to ensure the VLs are practical, user-friendly, and aligned with end-user needs. They will be incorporated iteratively throughout the design and prototyping phases to refine outputs and confirm usability.
- Requirements defined by the IAS Regulation

Based on the identified monitoring needs, we can design multiple VLs addressing IAS of Union concern. This requires the input of multiple work packages of BMD and thus should be a collaborative process. To design specific VLs in the VRE, we need to define four conceptual key components:

- Task: Short description of the specific monitoring task/question the VRE addresses. These tasks are derived from monitoring needs defined based on stakeholder requirements (e.g., based on user stories) and directives. [WP1] [WP5]
- Data: Description of relevant required and available datasets. [WP2] [WP5]
- Methods: Description of data analysis and modeling methods. [WP5]
- Outputs: Description of outputs (required/provided) and their formatting. [WP1] [WP5]





**Figure 1: Conceptual representation of the VRE design workflow. The VRE is designed to integrate multiple VLs for IAS use cases, based on monitoring needs.**

The complete workflow as well as model parametrization will be documented in a FAIR way, following the protocol suggested by Zurell et al. (2020) <sup>[15]</sup>. While this is specifically for species distribution models, we will adapt the respective parts also for temporal analyses.

## 5.2. VL list

VLs in the VRE correspond to specific IAS use cases across terrestrial, freshwater, and marine realms <sup>[14]</sup>. The tables summarize the relevant task, data, methods, outputs, and key issues for each realm.

**Table 1: VL for assessing the spatial distribution of terrestrial IAS**

Task	Assess the distribution of terrestrial IAS: Current distribution (realized and potential) and future potential distribution of terrestrial IAS
------	---



Data	<p>Occurrence data:</p> <ul style="list-style-type: none"> <li>• GBIF records on all terrestrial species on Union List and National checklists</li> <li>• Global occurrence data is important to characterize the full environmental niche of the species</li> <li>• Regional occurrence data (e.g., within Europe) can then be used to build regional SDMs that use outputs of global models as an input (nested distribution modeling approach)</li> </ul> <p>Predictor data:</p> <ul style="list-style-type: none"> <li>• CHELSA Bioclim: 19 bioclimatic variables as input to global SDMs</li> <li>• Chelsa Bioclim + other environmental predictors as inputs to regional SDMs. Potentially useful variables include datasets on human landscape modification, transportation infrastructure (roads &amp; railways), soil conditions, land cover, topographic variables <ul style="list-style-type: none"> <li>◦ The selection of specific datasets depends on their integration in the BMD data catalogue</li> </ul> </li> </ul>
Methods	<ol style="list-style-type: none"> <li>1. Current presence status inside Member states and/or Natura2000 sites: Occurrence based on GBIF data, summaries per Natura2000 site (e.g., identity and number of invasive species)</li> <li>2. Current potential/realized distributions: Nested/hierarchical SDMs, contrast potential habitat suitability predictions from SDMs with realized distributions from GBIF data to infer range filling and invasibility of habitats, or contrast global SDMs with regional SDMs to infer stage of invasion</li> <li>3. Future potential distributions (derived from climate-based SDMs)</li> </ol>
Outputs	Presence status for distribution maps and area trends, identity and number of potential invaders under current and future conditions
Issues	<ul style="list-style-type: none"> <li>• Invasive species are not at equilibrium with their environment, meaning that tailored modeling approaches should be used</li> </ul>

**Table 2: VL for assessing temporal trends for terrestrial IAS**

Task	Assess temporal trends in the occurrence of terrestrial IAS
Data	<p>Occurrence data:</p> <ul style="list-style-type: none"> <li>• GBIF records on all terrestrial species on Union List and National checklists within Member states</li> </ul> <p>Predictor data:</p> <ul style="list-style-type: none"> <li>• Data on spatiotemporal trends in sampling effort, for example number of occurrences for target-group taxa</li> <li>• For occupancy models: environmental predictor data similar to SDMs (climate, soil, topography, etc.)</li> </ul>
Methods	Trends could be calculated using:



	<ol style="list-style-type: none"> <li>1. Occupancy models, which separately model species occurrence and the observation process.</li> <li>2. The Frescalo method, which relies on local frequency scaling to correct for spatiotemporal sampling biases</li> <li>3. Simple summaries of IAS occurrence number trends relative to spatiotemporal indicators of sampling effort.</li> </ol>
Outputs	Graphs and statistics summarizing trends in IAS abundance/ richness per spatial unit (Natura2000 site, Member state) over time
Issues	<ul style="list-style-type: none"> <li>• Thorough calculation of population trends usually requires systematically collected data (i.e., survey-type data), including information on absences and sampling effort. The lack of this information will make inferences of trends less robust.</li> </ul>

**Table 3: VL for assessing freshwater IAS distributions**

Task	Assess the spatial distribution of freshwater IAS: Current and future potential distributions of freshwater invasive species
Data	<p>Occurrence data:</p> <ul style="list-style-type: none"> <li>• GBIF records of all freshwater IAS on the Union list and National checklists within Member states.</li> <li>• IAS records in monitoring data TREAM <sup>[16]</sup>, RivFishTIME <sup>[17]</sup>, &amp; BioTIME <sup>[18]</sup></li> </ul> <p>Predictor data:</p> <ul style="list-style-type: none"> <li>• CHELSA Bioclim (key predictors for this VL)</li> <li>• Land cover/use</li> <li>• Elevation</li> <li>• Hydrographic data <sup>[19]</sup> <ul style="list-style-type: none"> <li>• River basin and sub-catchment</li> <li>• Flow accumulation</li> <li>• Stream slope</li> <li>• Stream order</li> <li>• Stream power and compound topographic index</li> </ul> </li> <li>• Other factors: Human footprint index, inland water navigation, and distance to nearest port</li> </ul> <p>The selection of specific datasets depends on their integration in the BMD data catalogue.</p>
Methods	<ol style="list-style-type: none"> <li>1. Current presence status: Summarize GBIF occurrences per Member State and/or per river catchment within Member States (identity and number of invasive species).</li> <li>2. Current potential distributions: Develop SDMs, preferably using the Random Forest algorithm with down-sampling; compare predicted suitable habitats with observed occurrences to infer range filling and habitat invasibility.</li> </ol>



	3. Project the climate-based SDMs onto future climate scenarios to forecast changes in suitable habitats.
Outputs	Presence status of freshwater IAS per spatial unit (Member state and/or river catchment within Member States), distribution maps and area trends, identity and number of potential invaders under current and future conditions
Issues	<ul style="list-style-type: none"> <li>Invasive species are not at equilibrium with their environment, meaning that tailored modeling approaches should ideally be used.</li> <li>Whether and how to reflect dispersal capacity and dispersal pathways in projections of future distributions. This would require information on dispersal capacities (e.g., species-specific dispersal kernels).</li> </ul>

**Table 4: VL for evaluating temporal trends for freshwater IAS**

Task	Assess temporal trends in the occurrence of freshwater IAS
Data	<p>Occurrence data:</p> <ul style="list-style-type: none"> <li>GBIF records of all freshwater IAS on the Union list and National checklists within Member states.</li> <li>IAS records in monitoring data TREAM <sup>[16]</sup>, RivFishTIME <sup>[17]</sup>, &amp; BioTIME <sup>[18]</sup></li> </ul> <p>Predictor data:</p> <ul style="list-style-type: none"> <li>Data on spatiotemporal trends in sampling effort, for example number of occurrences for target-group taxa</li> <li>CHELSA Bioclim, land cover/use, hydrographic and other relevant factors' data</li> </ul>
Methods	<p>The following approaches may be employed to adjust for sampling and detection biases in species occurrence data to obtain spatiotemporally adjusted estimates:</p> <ul style="list-style-type: none"> <li>Occupancy models</li> <li>The Frescalo method</li> <li>Sampling effort-corrected summaries</li> </ul> <p>Temporal trends can then be analysed using linear models or generalized additive models, with potential drivers of change included to explain observed patterns.</p>
Outputs	Graphs and summary statistics describing temporal changes in IAS occurrence/ richness per spatial unit (Member State and/or river catchment within Member States) over time
Issues	Uneven spatial and temporal coverage of occurrence records; sampling and reporting bias may affect trend interpretation; absence data are often missing, limiting precise trend estimates

**Table 5: VL for assessing spatial distributions of marine IAS**

Task	Assess the presence status and spatial distribution of marine IAS: Current and future potential distributions of marine invasive species
Data	Occurrence data:



	<ul style="list-style-type: none"> <li>GBIF records of all marine IAS on the Union list and National checklists within Member states;</li> <li>OBIS (Ocean Biodiversity Information System; <a href="https://obis.org">https://obis.org</a>) for global marine species observations;</li> <li>EMODnet Biology (<a href="https://emodnet.ec.europa.eu/en/biology">https://emodnet.ec.europa.eu/en/biology</a>) for harmonized biological datasets across European marine regions;</li> <li>AquaNIS (<a href="https://aquanisresearch.com/">https://aquanisresearch.com/</a>) for verified records of aquatic non-indigenous and cryptogenic species;</li> <li>BioTIME <sup>[18]</sup></li> </ul> <p>Predictor data:</p> <ul style="list-style-type: none"> <li>Bio-ORACLE v2.2 (<a href="https://www.bio-oracle.org/downloads-to-email-v2.php?version=2_2">https://www.bio-oracle.org/downloads-to-email-v2.php?version=2_2</a>) and MARSPEC (<a href="https://marspec.weebly.com/">https://marspec.weebly.com/</a>) datasets for marine environmental variables (e.g., sea surface temperature, salinity, chlorophyll-a, current velocity, dissolved oxygen, pH);</li> <li>EMODnet Seabed Habitats (<a href="#">EMODnet Seabed Habitats</a>) for benthic substrate type, depth zones, and habitat classification;</li> <li>EMODnet Chemistry marine chemical data (<a href="https://emodnet.ec.europa.eu/en/chemistry">https://emodnet.ec.europa.eu/en/chemistry</a>);</li> <li>EMODnet Physics in situ ocean physics time-series data and vertical profiles (<a href="https://emodnet.ec.europa.eu/en/physics">https://emodnet.ec.europa.eu/en/physics</a>);</li> <li>Copernicus Marine Service (CMEMS - <a href="https://marine.copernicus.eu/">https://marine.copernicus.eu/</a>) for physical and biogeochemical parameters at high spatial and temporal resolution.</li> <li>Main shipping routes (<a href="https://github.com/newzealandpaul/Shipping-Lanes">https://github.com/newzealandpaul/Shipping-Lanes</a>)</li> </ul>
Methods	<ul style="list-style-type: none"> <li>Current presence status: Summarize marine IAS occurrences per Member State, MSFD subregion, and regional sea (identity and number of species), based on harmonized data from occurrence data repositories (e.g., OBIS, EMODnet Biology, AquaNIS, GBIF).</li> <li>Current potential distributions: Fit SDMs (e.g., MaxEnt, Random Forest) using marine environmental predictors (Bio-ORACLE, MARSPEC, EMODnet, CMEMS); compare predicted suitable habitats with observed occurrences to assess range filling and habitat invasibility.</li> <li>Future potential distributions: Use climate-based SDMs to project suitable habitats for marine IAS under future oceanographic and climate scenarios (e.g., CMIP6, SSP2-4.5, SSP5-8.5) to identify areas at potential risk of establishment or expansion.</li> </ul>
Outputs	Presence status of marine IAS per spatial unit (Member State and/or MSFD subregion, and regional sea), distribution maps and area trends, identity and number of potential invaders under current and future conditions
Issues	<ul style="list-style-type: none"> <li>Invasive species are not at equilibrium with their environment, meaning that tailored modeling approaches should ideally be used.</li> </ul>





	<ul style="list-style-type: none"> <li>Whether and how to reflect dispersal capacity and dispersal pathways in projections of future distributions. This would require information on dispersal capacities (e.g., species-specific dispersal kernels).</li> </ul>
--	---

**Table 6: VL for evaluating temporal trends for marine IAS**

Task	Assess temporal trends in the occurrence of marine IAS
Data	<p>Occurrence data:</p> <ul style="list-style-type: none"> <li>GBIF records of all marine IAS on the Union list and National checklists within Member states;</li> <li>OBIS (Ocean Biodiversity Information System; <a href="https://obis.org">https://obis.org</a>) for global marine species observations;</li> <li>EMODnet Biology (<a href="https://emodnet.ec.europa.eu/en/biology">https://emodnet.ec.europa.eu/en/biology</a>) for harmonized biological datasets across European marine regions;</li> <li>AquaNIS (<a href="https://aquanisresearch.com/">https://aquanisresearch.com/</a>) for verified records of aquatic non-indigenous and cryptogenic species;</li> <li>BioTIME <sup>[18]</sup>.</li> </ul> <p>Predictor data:</p> <ul style="list-style-type: none"> <li>Bio-ORACLE v2.2 (<a href="https://www.bio-oracle.org/downloads-to-email-v2.php?version=2_2">https://www.bio-oracle.org/downloads-to-email-v2.php?version=2_2</a>) and MARSPEC (<a href="https://marspec.weebly.com/">https://marspec.weebly.com/</a>) datasets for marine environmental variables (e.g. sea surface temperature, salinity, chlorophyll-a, current velocity, dissolved oxygen, pH);</li> <li>EMODnet Seabed Habitats (<a href="#">EMODnet Seabed Habitats</a>) for benthic substrate type, depth zones, and habitat classification;</li> <li>EMODnet marine chemical data (<a href="https://emodnet.ec.europa.eu/en/chemistry">https://emodnet.ec.europa.eu/en/chemistry</a>);</li> <li>EMODnet in situ ocean physics time-series data and vertical profiles (<a href="https://emodnet.ec.europa.eu/en/physics">https://emodnet.ec.europa.eu/en/physics</a>);</li> <li>Copernicus Marine Service (CMEMS - <a href="https://marine.copernicus.eu/">https://marine.copernicus.eu/</a>) for physical and biogeochemical parameters at high spatial and temporal resolution.</li> <li>Main shipping routes (<a href="https://github.com/newzealandpaul/Shipping-Lanes">https://github.com/newzealandpaul/Shipping-Lanes</a>)</li> </ul>
Methods	<p>The following approaches may be employed to adjust for sampling and detection biases in marine species occurrence data to obtain spatiotemporally adjusted estimates:</p> <ul style="list-style-type: none"> <li>Occupancy models</li> <li>The Frescalo method</li> <li>Bias-correction methods (e.g., target-group background, spatial thinning)</li> <li>Sampling effort-corrected summaries</li> </ul> <p>Temporal trends can then be analysed using linear models or generalized additive models, with environmental and anthropogenic drivers included to explain observed patterns.</p>



Outputs	Graphs and summary statistics describing temporal changes in IAS occurrence/ richness per spatial unit (per Member State, MSFD subregion, and regional sea) over time
Issues	Uneven spatial and temporal coverage of occurrence records; sampling and reporting bias may affect trend interpretation; absence data are often missing, limiting precise trend estimates.

**Table 7: VL for DNA-based analysis of invasive taxa for the marine environment**

Task	Environmental DNA based monitoring of the arrival and spread of new non-indigenous taxa in marine environments (also known as Non-Indigenous Species (NIS)) and for assessing trends in their introduction. And provide insight into community structure and ecological processes.
Data	<ul style="list-style-type: none"> <li>• Marker genes (16S/18S rRNA, ITS and COI)</li> <li>• High throughput sequence (PEMA Sequence Retriever)</li> <li>• WoRMS: An authoritative classification and catalogue of marine names (<a href="https://www.marinespecies.org/">https://www.marinespecies.org/</a>)</li> <li>• WriMS: World Register of Introduced Marine Species (<a href="https://www.marinespecies.org/introduced/">https://www.marinespecies.org/introduced/</a>)</li> <li>• FactorsFile is the metadata contains the SampleID and AreaType for metaMDS</li> </ul>
Methods	<ul style="list-style-type: none"> <li>• Three existing eDNA based workflows for identification of invasive taxa in marine environments can be improved. All three workflows share a common core analysis structure: The process begins by employing the PEMA pipeline to analyze the raw eDNA sequencing data. The resulting Operational Taxonomic Units are then linked to species records using the WoRMS and WriMS data. This step determines the status of each identified species (i.e. indigenous, non-indigenous or newly introduced) at the sample site. Besides the linking of PEMA output data to WoRMS and WriMS data the workflows each feature a unique addition: <ul style="list-style-type: none"> <li>○ Workflow 1 - DNA-based trend for the arrival of new non-indigenous taxa: does Nonmetric Multidimensional Scaling of the PEMA output data using metaMDS.</li> <li>○ Workflow 2 - DNA-based trend in the distributional range of invasive taxa: analyses the geographical expansion of the invasive species.</li> <li>○ Workflow 3 - DNA-based taxonomic relatedness: measures the phylogenetic and taxonomic relatedness among the identified invasive species using taxondrive.</li> </ul> </li> <li>• The three workflows use the following algorithms / pipelines: <ul style="list-style-type: none"> <li>○ Pipeline for Environmental DNA Metabarcoding Analysis (PEMA) <sup>[20]</sup> is used to process raw sequencing data from environmental samples and provide information on the community composition of organisms present. It supports the downstream analysis of four marker genes (16S/18S rRNA, ITS and COI), by allowing the user to train the classifiers with custom reference databases (<a href="https://github.com/hariszaf/pema">https://github.com/hariszaf/pema</a>).</li> </ul> </li> </ul>



	<ul style="list-style-type: none"> <li>○ Data analysis by calling the taxondive <sup>[21]</sup> (solution provided by vegan), which finds indices of taxonomic diversity and distinctness, which are averaged taxonomic distances among species or individuals in the community (<a href="https://github.com/vegandevs/vegan">https://github.com/vegandevs/vegan</a>).</li> <li>○ metaMDS <sup>[22]</sup>: Nonmetric Multidimensional Scaling with stable solution from random starts, axis scaling and species scores (<a href="https://github.com/vegandevs/vegan">https://github.com/vegandevs/vegan</a>).</li> </ul>
Outputs	<p>Clustering into Operational Taxonomic Units (OTUs) or inferring Amplicon Sequence Variants (ASVs).</p> <p>Figures of indices of taxonomic diversity and distinctness, which are averaged taxonomic distances among species or individuals in the community.</p> <p>By integrating with the <a href="#">Biodiversity meets Cubes</a> framework, the species occurrence data from our workflows will be transformed into standardized data cubes, making them readily usable across other virtual research environments.</p>
Issues	<p>Reference database incompleteness: The single biggest hurdle. The taxonomic assignment step is only as good as the reference databases (e.g., SILVA, UNITE)</p> <p>PEMA uses tools like Swarm and VSEARCH for clustering and error mitigation, but these errors can never be fully eliminated</p>

## 6. Technical implementation

### 6.1. Spatial analyses

Species distribution models will be built using state-of-the-art approaches for modeling invasive species. Specifically, considering the following aspects will be important:

1. Accounting for spatial sampling bias: Given the use of GBIF data mainly containing opportunistically sampled presence-only data, the models should account for spatial sampling bias. This can be achieved, for example, via specific background-sampling approaches (e.g., target-group background sampling <sup>[20]</sup>).
2. Minimizing impacts of niche truncation: When SDMs are fitted within spatial extents that do not encompass the whole species range, the estimated environmental niche can be truncated, making the model less reliable when projecting across space and time. To avoid this issue, global occurrence data (encompassing the entire range) should be used. Spatially nested SDMs offer an effective approach to address this issue <sup>[21]</sup>. The combination of global and regional models has also been shown to be effective for improving SDMs of IAS and understanding invasion dynamics <sup>[22]</sup>.
3. Use of well-performing SDM algorithms: Flexible machine learning algorithms such as Maxent, random forest, and boosted regression trees yield the highest predictive performance for presence-only data in most cases <sup>[23]</sup>. However, hyperparameter optimization is important (see below).
4. Optimization (tuning) of algorithm hyperparameters: Optimizing algorithm hyperparameters is important to help regularize models and avoid overfitting. This can be achieved based on cross-



validations of the model. To simulate model transfer, spatial/environmental block cross-validation can be used <sup>[24, 25]</sup>

5. Identification of model extrapolation: The training data available for building SDMs often do not capture the full domain of environmental conditions for which predictions should be made. This causes model extrapolation. Regions where extrapolation occurs should be flagged as having higher uncertainty. Extrapolation can be detected, for example, using a so-called MESS analysis (univariate extrapolation <sup>[26]</sup>), or based on the Mahalanobis distance in order to also capture multivariate extrapolation (new combinations of variable values <sup>[27]</sup>).
6. Considering multiple climate scenarios and models: For SDM projections to future climate conditions, options should be provided to compare different SSP (emission) scenarios. Bioclimatic variables based on different climate models should either be combined into ensemble predictions (i.e., averaging across all climate models), or the range/uncertainty across predictions based on individual climate models should be provided.

## 6.2. Temporal analyses

Without systematically collected data on species abundance, which most likely will not be available for this use case, approaches that robustly estimate population trends are limited. A robust approach for opportunistically sampled presence-only data (i.e., corresponding to most data available via GBIF) are occupancy models. While occupancy models require so-called detection histories containing presence-absence data collected through repeated visits to a set of study sites (e.g., collected through surveys), opportunistic occurrence data can be aggregated into detection histories for their use in occupancy models <sup>[28]</sup>. Occupancy models employ a hierarchical modeling approach to separately model species occupancy (presence/absence) and the observation process (detection/non-detection). In the context of opportunistic data, non-detections can be inferred from visits to a site (e.g., pixel in a raster grid) without the target species being detected, while the length of species lists compiled by observers (“list length”) can be used as a proxy for sampling effort. For details on the implementation of occupancy models for opportunistic data, see, for example: Kéry et al 2010 <sup>[29]</sup>, van Strien et al 2013 <sup>[30]</sup>, Dennis et al 2017 <sup>[31]</sup>. In R, occupancy models can be fit, for example, using the unmarked package (<https://cran.r-project.org/web/packages/unmarked/index.html>).

A second robust approach for analyzing temporal trends based on opportunistic occurrence data is the Frescalo method <sup>[32]</sup>. Instead of the site-level, Frescalo operates at a broader, neighborhood scale (typically around 10-100 km size) to correct for uneven recording effort. Its core principle is local frequency scaling, which adjusts the observed frequency of a species based on how well-recorded a given area and time period were, using other species as a reference. For details on the implementation of the Frescalo method for unstructured data, see Goury et al 2025 <sup>[33]</sup>. In R, the Frescalo method can be implemented using the sparta package (<https://github.com/BiologicalRecordsCentre/sparta>).

Finally, trend indices could be generated by setting occurrence record numbers in relation to indicators of sampling effort (for example, number of records for target-group taxon, or number of observers), could be calculated <sup>[43]</sup>. In addition, a resampling-based abundance index could be calculated <sup>[35]</sup>.



The three described approaches differ in their levels of data availability requirements (occupancy models > Frescalo > trend indices) and which of them can be used in VREs thus depends on data availability for different regions/taxa. For comparability and consistency, using the same approaches across target IAS and sites (Natura2000 sites and/or Member states) would be desirable. Which of these approaches will be feasible will depend on the spatiotemporal availability of occurrence records for IAS, as well as for reference species used to infer sampling effort.

### 6.3. User interface and outputs

User interfaces for the VREs / VLs should be easily accessible to non-technical users. The selection of target species and/or target spatial units (e.g., Natura2000 sites, Member states) could thus be implemented via a user-friendly dashboard, for example using (searchable) dropdown menus, or similar. The selection of target species should be informed by the EU- and national-level species lists, depending on the area of interest specified by the user. Given the non-technical user audience, VRE / VLs should run largely automatized based on minimal user input (i.e., not requiring manual parametrization of models or analyses).

Outputs can be grouped into visual and non-visual outputs (e.g., graphs/maps vs. numbers), as well as displayed outputs and downloadable outputs. Displayed outputs could be shared with users via the visualization engine developed in Work Package 4. Occurrence data points and SDM prediction maps can be displayed in interactive web maps. Similarly, graphs for temporal trends could be displayed in interactive plots.

Downloadable outputs could be provided to users via the Single Access Point (SAP) developed within BMD (Work Package 6). These may include:

- Continuous and binary (thresholded) distribution maps, for example in a GeoTIFF format
- Species occurrence points, for example in a CSV or Shapefile format
- Tables containing summaries of occurrence numbers and distribution area trends, for example in a CSV format
- Underlying R scripts that run the implemented analyses (only if this is requested by stakeholders/users)

VREs / VLs should be accompanied by non-technical explanations of the implemented analyses. Uncertainties associated with the outputs should be clearly stated, also regarding the general limitations of SDMs. For example, presence-only SDMs only predict the potential distribution of a species (environmentally suitable habitat), while the realized distribution of a species is furthermore influenced by other factors (biotic interactions with other organisms and dispersal limitations). When calculating range sizes based on model predictions (i.e., binarizing continuous predictions of habitat suitability), the influence of threshold selection should be indicated through the comparison of multiple thresholds.

## 7. Conclusion and future steps



This MS24 presents the conceptual framework for integrating invasive alien species (IAS) use cases of Union concern in the Virtual Research Environment (VRE). By aligning regulatory requirements under the IAS Regulation (EU Regulation 1143/2014/EC) with standardized data, workflows, and outputs, the proposed VRE, which cover species information, presence status, distributions, and trends, provide a reproducible and scalable solution for monitoring, reporting, and management.

Stakeholder feedback will play a central role in refining the VRE. Input from regulatory authorities, site managers, and NGOs will inform user interfaces, workflow design, and data presentation, ensuring the VREs are practical, user-friendly, and directly supportive of decision-making.

The next steps include workflow prototyping (Q4 2025), stakeholder co-design workshops, and Single Access Point integration, in line with deliverable D5.2 and the upcoming milestones MS5.3.1 (IAS prototype) and MS5.4.1 (evaluation). By involving stakeholders throughout these phases, we can ensure the VRE not only meets regulatory requirements but also genuinely addresses the needs of the end-users.

## 8. Acknowledgements

The authors gratefully thank Quentin Groom for his helpful feedback and suggestions. We also acknowledge the European Commission for its financial support through the EU Horizon Europe Research and Innovation Action (Grant Agreement No. 101181294, BMD project).

## 9. References

1. European Commission. *Invasive Alien Species*, Environment – Nature and biodiversity. Accessed September 5, 2025. [https://environment.ec.europa.eu/topics/nature-and-biodiversity/invasive-alien-species\\_en](https://environment.ec.europa.eu/topics/nature-and-biodiversity/invasive-alien-species_en)
2. IPBES (2023) Summary for Policymakers of the Thematic Assessment Report on Invasive Alien Species and their Control of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. In: Roy H.E, Pauchard A, Stoett P, Renard Truong T, Bacher S, Galil BS, Hulme PE, Ikeda T, Sankaran KV, McGeoch MA, Meyerson LA, Nuñez MA, Ordonez A, Rahlao SJ, Schwindt E, Seebens H, Sheppard AW, Vandvik V (eds), IPBES secretariat, Bonn, Germany, <https://doi.org/10.5281/zenodo.7430692>
3. Jaureguiberry, P., Titeux, N., Wiemers, M., Bowler, D. E., Coscieme, L., Golden, A. S., et al. (2022). The direct drivers of recent global anthropogenic biodiversity loss. *Science advances*, 8(45), eabm9982. <https://doi.org/10.1126/sciadv.abm9982>
4. Bradley, B. A., Laginhas, B. B., Whitlock, R., Allen, J. M., Bates, A. E., Bernatchez, G., et al. (2019). Disentangling the abundance–impact relationship for invasive species. *Proceedings of the National Academy of Sciences*, 116(20), 9919–9924. <https://doi.org/10.1073/pnas.1818081116>
5. Gallardo, B., Bacher, S., Barbosa, A. M., Gallien, L., González-Moreno, P., Martínez-Bolea, V., et al. (2024). Risks posed by invasive species to the provision of ecosystem services in Europe. *Nature Communications*, 15(1), 2631. <https://doi.org/10.1038/s41467-024-46818-3>
6. Vilà, M., Basnou, C., Pyšek, P., Josefsson, M., Genovesi, P., Gollasch, S., Nentwig, W., Olenin, S., Roques, A., Roy, D., Hulme, P.E., Andriopoulos, P., Arianoutsou, M., Augustin, S., Bacher, S., Bazos, I., Bretagnolle, F., Chiron, F., Clergeau, P., Cochar, P.O., Cocquempot, C., Coeur d’Acier,





- A., David, M., Delipetrou, P., Desprez-Loustau, M.L., Didžiulis, V., Dorkeld, F., Essl, F., Galil, B.S., Gasquez, J., Georgioudis, K., Gudžinskas, Z., Hatzofe, O., Hejda, M., Jarošík, V., Kark, S., Kokkoris, I., Kühn, I., Lambdon, P.W., Lopez-Vaamonde, C., Marcer, A., Migeon, A., McLoughlin, M., Minchin, D., Navajas, M., Panov, V.E., Pascal, M., Pergl, J., Perglová, I., Pino, J., Poblons, K., Rabitsch, W., Rasplus, J.Y., Sauvard, D., Scalera, R., Sedláček, O., Shirley, S., Winter, M., Yannitsaros, A., Yart, A., Zagatti, P., Zikos, A., 2010. How well do we understand the impacts of alien species on ecosystem services? A pan-European, cross-taxa assessment. *Frontiers in Ecology and the Environment*, 8, 135–144. <https://doi.org/10.1890/080083>
7. Diagne, C., Leroy, B., Vaissiere, A.C., Gozlan, R.E., Roiz, D., Jaric, I., Salles, J.M., Bradshaw, C.J.A., Courchamp, F., 2021. High and rising economic costs of biological invasions worldwide. *Nature* 592, 571–576. <https://doi.org/10.1038/s41586-021-03405-6>
  8. Zenni, R.D., Essl, F., García-Berthou, E., McDermott, S.M., 2021. The economic costs of biological invasions around the world. *NEOBIOTA*, 67, 1–9. <https://doi.org/10.3897/neobiota.67.69971>
  9. Bellard, C., Thuiller, W., Leroy, B., Genovesi, P., Bakkenes, M., & Courchamp, F. (2013). Will climate change promote future invasions? *Global Change Biology*, 19(12), 3740–3748. <https://doi.org/10.1111/gcb.12344>
  10. Carboneras, C., Genovesi, P., Vilà, M., Blackburn, T. M., Carrete, M., Clavero, M., et al. (2018). A prioritised list of invasive alien species to assist the effective implementation of EU legislation. *Journal of Applied Ecology*, 55(2), 539–547. <https://doi.org/10.1111/1365-2664.12997>
  11. European Environment Agency. (2010). *Towards an early warning and information system for invasive alien species (IAS) threatening biodiversity in Europe* (EEA Technical Report No. 5/2010). EEA. <https://doi.org/10.2800/4167>
  12. EU Regulation 1143/2014/EC on Invasive Alien Species. <http://data.europa.eu/eli/reg/2014/1143/oj>
  13. Commission Implementing Regulation (EU) 2025/1422. [http://data.europa.eu/eli/reg\\_impl/2025/1422/oj](http://data.europa.eu/eli/reg_impl/2025/1422/oj)
  14. Oeser, J. (2025). *BMD WP5 Handbook* (2nd draft, 30 June 2025). Unpublished internal document.
  15. Zurell, D., Franklin, J., König, C., Bouchet, P.J., Dormann, C.F., Elith, J., Fandos, G., Feng, X., Guillerá-Arroita, G., Guisan, A., Lahoz-Monfort, J.J., Leitão, P.J., Park, D.S., Peterson, A.T., Rapacciuolo, G., Schmatz, D.R., Schröder, B., Serra-Díaz, J.M., Thuiller, W., Yates, K.L., Zimmermann, N.E., Merow, C., 2020. A standard protocol for reporting species distribution models. *Ecography*, 43, 1261–1277. <https://doi.org/10.1111/ecog.04960>
  16. Welti, E. A., Bowler, D. E., Sinclair, J. S., Altermatt, F., Álvarez-Cabria, M., Amatulli, G., ... & Haase, P. (2024). Time series of freshwater macroinvertebrate abundances and site characteristics of European streams and rivers. *Scientific Data*, 11(1), 601. <https://doi.org/10.1038/s41597-024-03445-3>
  17. Comte, L., Carvajal-Quintero, J., Tedesco, P. A., Giam, X., Brose, U., Erős, T., ... & Olden, J. D. (2021). RivFishTIME: A global database of fish time-series to study global change ecology in riverine systems. *Global Ecology and Biogeography*, 30(1), 38–50. <https://doi.org/10.1111/geb.13210>
  18. Dornelas, M., Antão, L. H., Bates, A. E., Brambilla, V., Chase, J. M., Chow, C. F., ... & Fryxell, J. (2025). BioTIME 2.0: Expanding and improving a database of biodiversity time series. *Global Ecology and Biogeography*, 34(5), e70003. <https://doi.org/10.1111/geb.70003>
  19. Amatulli, G., Garcia Marquez, J., Sethi, T., Kiesel, J., Grigoropoulou, A., Üblacker, M.M., Shen, L.Q. and Domisch, S., 2022. Hydrography90m: A new high-resolution global hydrographic



- dataset. *Earth System Science Data*, 14(10), 4525-4550. <https://doi.org/10.5194/essd-14-4525-2022>
20. Zafeiropoulos, H., Viet, H. Q., Vasileiadou, K., Potirakis, A., Arvanitidis, C., Topalis, P., ... & Pafilis, E. (2020). PEMA: a flexible pipeline for environmental DNA metabarcoding analysis of the 16S/18S ribosomal RNA, ITS, and COI marker genes. *GigaScience*, 9(3), giaa022. <https://doi.org/10.1093/gigascience/giaa022>
  21. Clarke, K. R., & Warwick, R. M. (2001). A further biodiversity index applicable to species lists: variation in taxonomic distinctness. *Marine Ecology Progress series*, 216, 265-278. <https://doi.org/10.3354/meps>
  22. Kruskal, J. B. (1964). Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 29(2), 115-129. <https://doi.org/10.1007/BF02289694>
  23. Barber, R. A., Ball, S. G., Morris, R. K., & Gilbert, F. (2022). Target-group backgrounds prove effective at correcting sampling bias in Maxent models. *Diversity and Distributions*, 28(1), 128-141. <https://doi.org/10.1111/ddi.13442>
  24. Guisan, A., Chevalier, M., Adde, A., Zarzo-Arias, A., Goicolea, T., Broennimann, O., ... & Mateo, R. G. (2025). Spatially nested species distribution models (N-SDM): An effective tool to overcome niche truncation for more robust inference and projections. *Journal of Ecology*, 113(7), 1588-1605. <https://doi.org/10.1111/1365-2745.70063>
  25. Gallien, L., Douzet, R., Pratte, S., Zimmermann, N. E., & Thuiller, W. (2012). Invasive species distribution models—how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography*, 21(11), 1126-1136. <https://doi.org/10.1111/j.1466-8238.2012.00768.x>
  26. Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J. J., & Elith, J. (2022). Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. *Ecological Monographs*, 92(1), e01486. <https://doi.org/10.1002/ecm.1486>
  27. Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., ... & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913-929. <https://doi.org/10.1111/ecog.02881>
  28. Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2018). blockCV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Biorxiv*, 357798. <https://doi.org/10.1101/2041-210X.13107>
  29. Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1(4), 330-342. <http://www.respond2articles.com/MEE/>
  30. Mesgaran, M. B., Cousens, R. D., & Webber, B. L. (2014). Here be dragons: a tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. *Diversity and Distributions*, 20(10), 1147-1159. <https://doi.org/10.1111/ddi.12209>
  31. Isaac, N. J., van Strien, A. J., August, T. A., de Zeeuw, M. P., & Roy, D. B. (2014). Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution*, 5(10), 1052-1060. <https://doi.org/10.1111/2041-210X.12254>
  32. Kery, M., Royle, J. A., Schmid, H., Schaub, M., Volet, B., Häfliger, G., & Zbinden, N. (2010). Site-occupancy distribution modeling to correct population-trend estimates derived from opportunistic observations. *Conservation Biology*, 24(5), 1388-1397. <https://doi.org/10.1111/j.1523-1739.2010.01479.x>
  33. Van Strien, A. J., Van Swaay, C. A., & Termaat, T. (2013). Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *Journal of Applied Ecology*, 50(6), 1450-1458. <https://doi.org/10.1111/1365-2664.12158>





34. Dennis, E. B., Morgan, B. J., Freeman, S. N., Ridout, M. S., Brereton, T. M., Fox, R., ... & Roy, D. B. (2017). Efficient occupancy model-fitting for extensive citizen-science data. *PloS one*, 12(3), e0174433. <https://doi.org/10.1371/journal.pone.0174433>
35. Hill, M. O. (2012). Local frequency as a key to interpreting species occurrence data when recording effort is not known. *Methods in Ecology and Evolution*, 3(1), 195-205. <https://doi.org/10.1111/j.2041-210X.2011.00146.x>
36. Goury, R., Bowler, D. E., Harrower, C., Münkemüller, T., Vallet, J., Yearsley, J., ... & Pescott, O. L. (2025). A practical guide to species trend detection with unstructured data using local frequency scaling (Frescalo). Preprint. <https://ecoevorxiv.org/repository/view/9467/>
37. Knape, J., Coulson, S. J., van der Wal, R., & Arlt, D. (2022). Temporal trends in opportunistic citizen science reports across multiple taxa. *Ambio*, 51(1), 183-198. <https://doi.org/10.1007/s13280-021-01550-w>
38. Zbinden, N., Kéry, M., Häfliger, G., Schmid, H., & Keller, V. (2014). A resampling-based method for effort correction in abundance trend analyses from opportunistic biological records. *Bird Study*, 61(4), 506-517. <https://doi.org/10.1080/00063657.2014.969679>

