



BIODIVERSITY
BUILDING
BLOCKS FOR
POLICY

D6.3 Case study III - Regional indicators in Europe

31/07/2026

Author(s): Jasmijn Hillaert, Lissa Breugelmans, Ward Langerlaert, Tim Adriaens, Soria Delva, Sandra MacFadyen, Damiano Oldoni, Nikol Stoykova, Toon Van Daele & Quentin Groom



Funded by
the European Union

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the EU nor the EC can be held responsible for them.



Prepared under contract from the European Commission

Grant agreement No. 101059592

EU Horizon Europe Research and Innovation Action

Project acronym: **B3**

Project full title: **Biodiversity Building Blocks for policy**

Project duration: 01.03.2023 – 31.08.2026 (42 months)

Project coordinator: Dr. Quentin Groom, Agentschap Plantentuin Meise (MeiseBG)

Call: HORIZON-CL6-2021-GOVERNANCE-01

Deliverable title: Case study III - Regional indicators in Europe

Deliverable n°: D6.3

WP responsible: WP6

Nature of the deliverable: Report and online dashboard

Dissemination level: Public

Licence of use: CC-BY

Lead partner: INBO

Recommended citation: Hillaert J., Breugelmans L., Langerlaert W., Adriaens T., Delva, S., MacFadyen S., Oldoni D., Stoykova N., Vandaele T., & Groom Q. (2026). **Case study III - Regional indicators in Europe**. B3 project deliverable D6.3.

Due date of deliverable: 41

Actual submission date: 42

Deliverable status:

Version	Status	Date	Author(s)
1.0	Final	31 July 2026	Jasmijn Hillaert, INBO Lissa Breugelmans, MeiseBG Ward Langerlaert, INBO Tim Adriaens, INBO Soria Delva, INBO Sandra MacFadyen, SUN Damiano Oldoni, INBO Nikol Stoykova, Pensoft Toon Vandaele, INBO Quentin Groom, Meise BG



Table of contents

Key takeaway messages	4
Executive summary	4
Non-technical summary	5
Graphical summary	6
List of abbreviations	7
1. Introduction	8
1.1. From data to policy	8
1.2. Indicator types	8
1.3. Overview of available GBIF data in Flanders	10
1.4. Goal of the Flanders use case	15
2. Methodology	18
2.1. Creating cubes with rgbif	18
2.2. Performing (dubi)cube quality checks	21
2.3. Overview of indicators tested	24
2.4. Natura 2000 versus non-Natura 2000	30
2.5. How to deal with survey effort bias?	34
3. Results	37
3.1. General indicators	37
3.2. Phylogenetic indicators	50
3.3. Invasive alien species indicators	53
3.4. Natura 2000 versus non-Natura 2000.	70
3.5. WiSDM Flow and cubes	77
4. Discussion and future perspectives	79
4.1. General findings	79
4.2. Modifiable areal unit problem	79
4.3. invasimapr as a regional invasion indicator	80
4.4. Limitations and recommendations	81
4.5. Future prospects	82
Acknowledgements	84
References	85



Key takeaway messages

- Biodiversity indicators can effectively transform GBIF data into policy-relevant information.
- A public, interoperable dashboard makes biodiversity indicators accessible and reusable.
- Accounting for survey effort is essential for reliable biodiversity assessments.
- Data quality and survey effort largely depend on taxa and publishing history of the region of interest.
- The B3 project expands the toolbox for biodiversity monitoring across multiple policy priorities.

Executive summary

Work package 6 of the B3 project focuses on testing biodiversity indicators and modelling approaches developed during the project across several use cases. This deliverable presents a list of indicators used in Case Study III: Regional Indicators in Europe. The objective is to evaluate their performance and suitability for monitoring biodiversity trends in Flanders using real-world data on native and invasive species.

The implemented indicators are subdivided into:

- General biodiversity indicators (based on the **b3gbi** and **dubicube** R packages, for instance cumulative species richness, relative species richness, etc.);
- Phylogenetic indicators (based on the **pdindicator** R package);
- Invasive Alien Species (IAS) indicators (based on the **invasimapr**, **b3alien** and **trias** packages).

Particular attention is given to comparing Natura 2000 and non-Natura 2000 habitats when applying these indicators, quantifying and interpreting indicator uncertainty, and accounting for survey effort.

The main outcomes are:

- The selected indicators are made available through a public dashboard (<https://flanders.b-cubed.eu/>) that can serve as inspiration for other dashboards (e.g. the Alien Species Portal, the Flemish biodiversity portal, and the Belgium GBIF node).
- Close collaboration with other dashboard maintainers ensures interoperability, for instance all outputs are provided on Zenodo in GeoJSON format.
- Indicator factsheets are available in printed format and digitally through Zenodo.



Non-technical summary

Understanding spatial and temporal patterns in biodiversity is essential for protecting nature and supporting informed management decisions. In this case study, workflows developed within the B3 project are tested in Flanders, a region in Europe with extensive species occurrence data available through GBIF.

Three types of indicators are evaluated using several packages developed within B3. First, general biodiversity indicators are applied to assess broad biodiversity patterns, such as the number of reported records per km² or the trend of total observed species richness over time. Second, phylogenetic diversity indicators are calculated. While closely related to observed species richness, these indicators also account for the evolutionary distance between species. Third, indicators focusing specifically on invasive alien species are assessed. These indicators support the implementation of the European Regulation on invasive alien species (Regulation (EU) No [1143/2014](#)).

Special attention is given to differences in indicator values between Natura 2000 and non-Natura 2000 areas. Natura 2000 is a network of protected areas across the European Union, established to conserve species and habitats of European importance. In addition, several approaches to account for observer bias in the calculation of biodiversity indicators are tested.

The selected indicators are made available through a public dashboard (<https://flanders.b-cubed.eu/>), allowing users to explore biodiversity patterns and trends. The dashboard and underlying code can serve as an inspiration for other dashboards to also adopt the B3 workflow. Finally, clear factsheets were created for each B3 indicator, outlining the indicator type, purpose, policy relevance and limitations.



Graphical summary



Graphical overview of the Flanders use case workflow.



List of abbreviations

B3	Biodiversity Building Blocks for policy
CBD	Convention on Biological Diversity
CR	Critically Endangered
EASIN	Driver-Pressure-State-Impact-Response
EEA	European Environment Agency
EN	Endangered
EU	European Union
LOESS	Locally Estimated Scatterplot Smoothing
IAS	Invasive Alien Species
ISEIA	Invasive Species Environmental Impact Assessment
GAM	Generalized Additive Model
GBF	Global Biodiversity Framework
GBIF	Global Biodiversity Information Facility
GLMM	Generalized Linear Mixed Model
GRIIS	Global Register of Introduced and Invasive Species
MAUP	Modifiable Areal Unit Problem
NMDS	Non-metric Multidimensional Scaling
OIS	'OmgevingsInformatie Samenwerkingsverband'
PD	Phylogenetic Diversity
TrIAS	Tracking Invasive Alien Species
VU	Vulnerable
WPx	Work package "x" of the B3 project



1. Introduction

1.1. From data to policy

Biodiversity monitoring generates large volumes of observations that only become useful for policy once they have been processed into meaningful information. The Data–Information–Knowledge–Wisdom (DIKW) hierarchy provides a conceptual framework describing this transformation (Ackoff, 1989; Rowley, 2007). Within this hierarchy, indicators play a central role by summarising complex datasets into interpretable measures of biodiversity status and change. The B3 project develops workflows and indicators that facilitate this transformation, enabling reproducible analyses of biodiversity data and providing an evidence base for biodiversity assessment and policy (Figure 1).

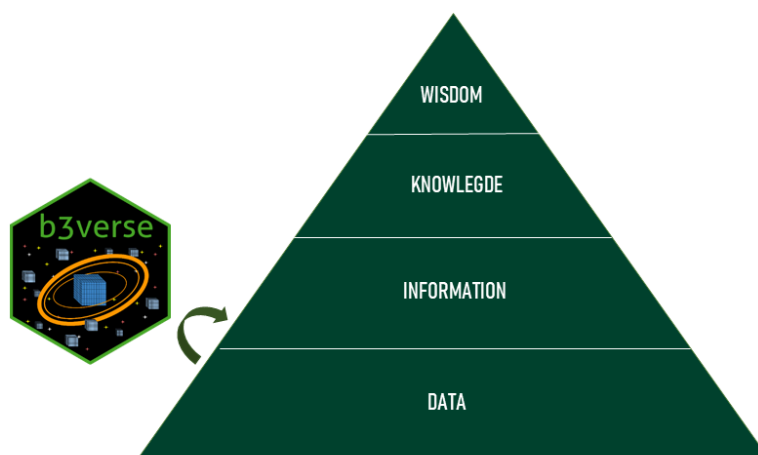


Figure 1: Conceptual representation of the Data–Information–Knowledge–Wisdom (DIKW) hierarchy (adapted from Rowley, 2007), illustrating how the B3 workflows contribute to transforming biodiversity observations into policy-relevant information and knowledge.

1.2. Indicator types

Depending on the purpose of an indicator, several indicator types can be distinguished. See Table 1 for an overview of applied indicator types within this case study. This classification is partly in line with the Driver-Pressure-State-Impact-Response (DPSIR) framework (Gari et al., 2015).

**Table 1: Overview of the indicator classification applied in this case study.**

Indicator type	Description
State or Trend	Describes the state of the ecosystem
Pressure	Captures forces or activities that are causing change or stress on a system.
Monitoring management	Tracks the actions taken to manage or influence a system, including policies or interventions.
Effects on ecosystem services	Measures how biodiversity and ecosystem changes affect benefits that ecosystems provide to people, such as clean water, pollination, or recreation.
Effects on economy and society	Assesses how environmental changes influence human wellbeing, livelihoods, health, and economic activity.
Data quality	Describes the accurateness, completeness and validity of the data

Figure 2 provides a graphical overview of the different indicator types and their interrelationships. Pressure indicators and management monitoring indicators both help to understand the external factors that drive ecosystem and biodiversity changes. While pressure indicators focus on factors that exert stress on ecosystems, monitoring management indicators provide information regarding management effectiveness and possible side effects. Ecosystem services indicators, in turn, assess the potential impacts of ecosystems and biodiversity changes on ecosystem functioning. These impacts may subsequently affect economic activities and societal well-being, which are captured through economy and society indicators. Eventually, data quality indicators help to identify potential weaknesses and uncertainties in the data flow. This is particularly important for indicators based on ad hoc or unstructured data, which is largely biased by the observer effect.

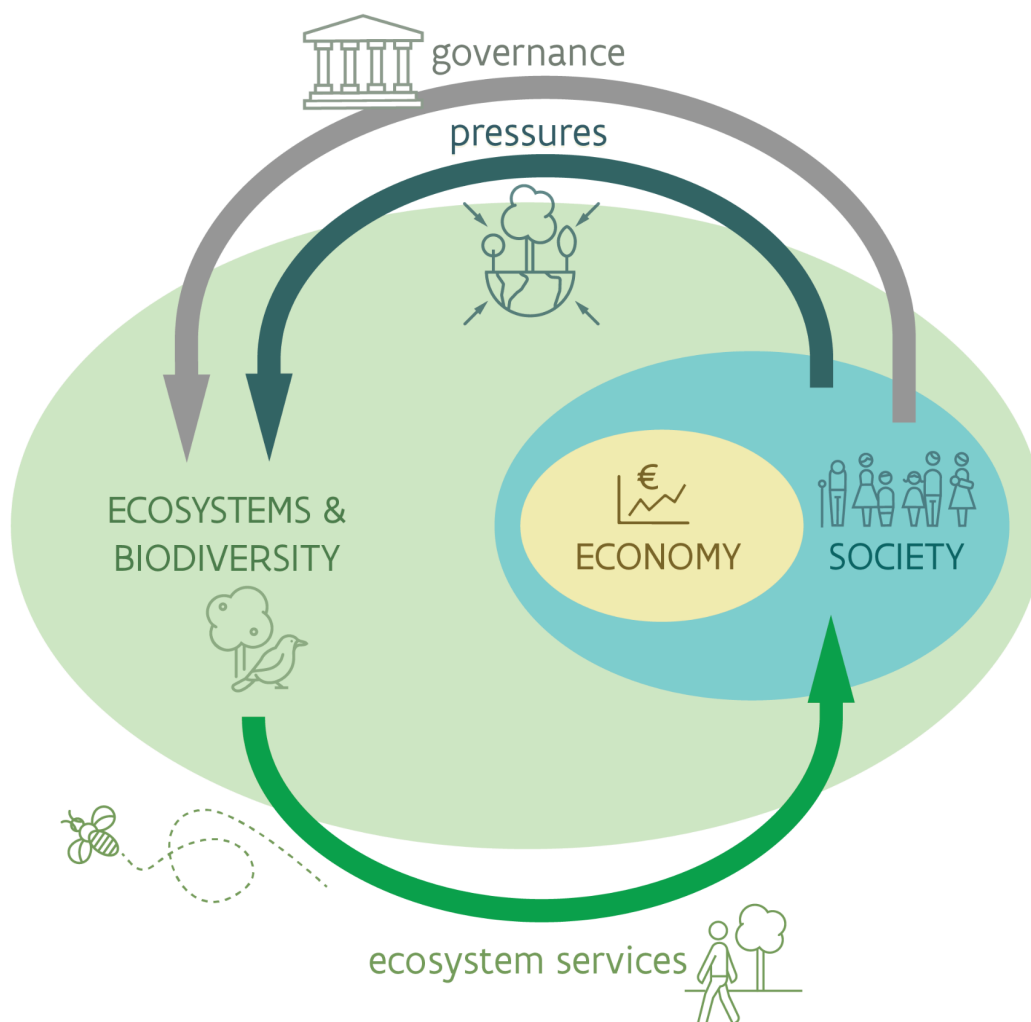


Figure 2: Graphical representation of indicator types applied in this case study and how these are interlinked. © INBO

1.3. Overview of available GBIF data in Flanders

The use case focuses on Flanders, the northern region of Belgium in northwestern Europe (Figure 3). Flanders consists of five provinces and has a population of approximately 6.9 million



inhabitants (Statistiek Vlaanderen, 2026).



Figure 3. Geographic location of Flanders (hatched area) within Belgium (dark grey) in northwestern Europe. Flanders, located in the northern part of Belgium, comprises five provinces. Wallonia occupies the southern part of Belgium, while the Brussels-Capital Region is centrally located and geographically enclosed by Flanders.

In total for Flanders, more than 36 million records are available on GBIF since 1975. This number largely fluctuates over time with a strong increasing trend since 2008 when [Waarnemingen.be](https://www.waarnemingen.be) started to collect citizen science observations in Flanders. The number of records suddenly drops in 2018 (Figure 4). Since that year, Natuurpunt stopped publishing data of native species from the [Waarnemingen.be](https://www.waarnemingen.be) platform and only alien species are currently published on GBIF. This fact biases the observed species trends and complicates the process of distinguishing ecological reality from observer bias. Because of a publication time lag, data of the last three years (2023, 2024, and 2025) are not yet complete.



The spatial distribution of total records published since 1975 in Flanders is visualised in Figure 5. Although the map provides an overview of recording intensity, the spatial pattern should be interpreted with caution. Some datasets are published as centroids at relatively coarse spatial resolutions based on specific grid reference systems, resulting in an artificial rectangular pattern that does not necessarily reflect the true spatial distribution of observations (Figure 5).

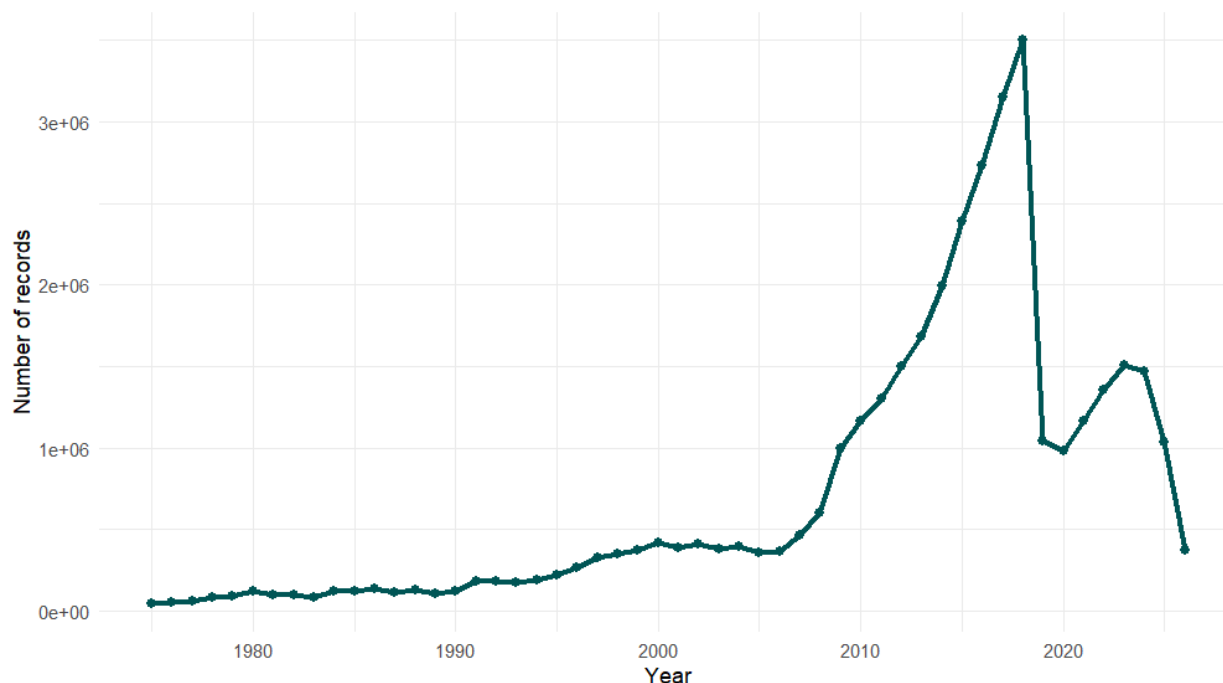


Figure 4: An overview of the total number of records published on GBIF since 1975 for Flanders. Data source: [GBIF.org](https://gbif.org) (2026a).

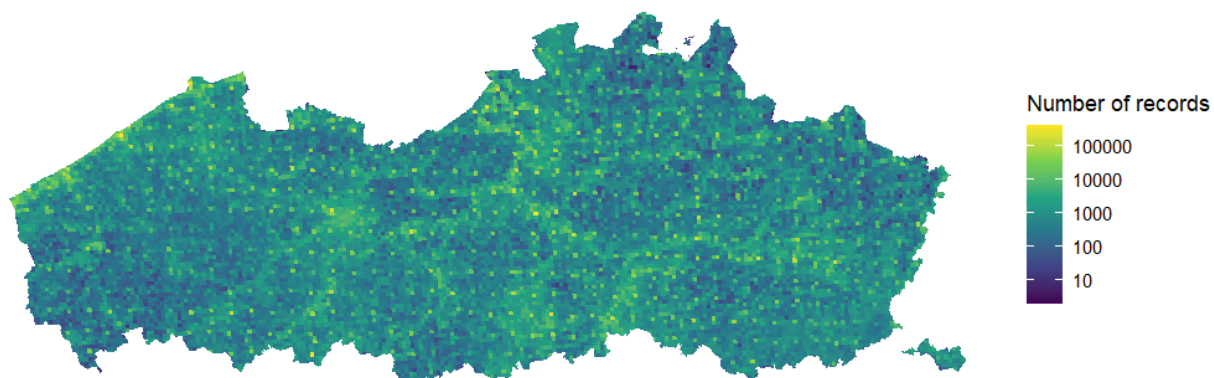


Figure 5: The spatial distribution of the total number of records published per EEA- 1km gridcell on GBIF since 1975 for Flanders. Spatial randomisation was not applied to the occurrence records and no boundaries were set for spatial uncertainty. Data source: [GBIF.org](https://gbif.org) (2026a).



When looking at the most important publishers of records to GBIF, Natuurpunt, who hosts the *Waarnemingen.be* website, largely exceeds the others with almost 20 million records published since 2000, followed by the Institute of Nature and Forest with over 10 million records (Figure 6). A detailed overview of the number of records annually published is available in Figure 7.

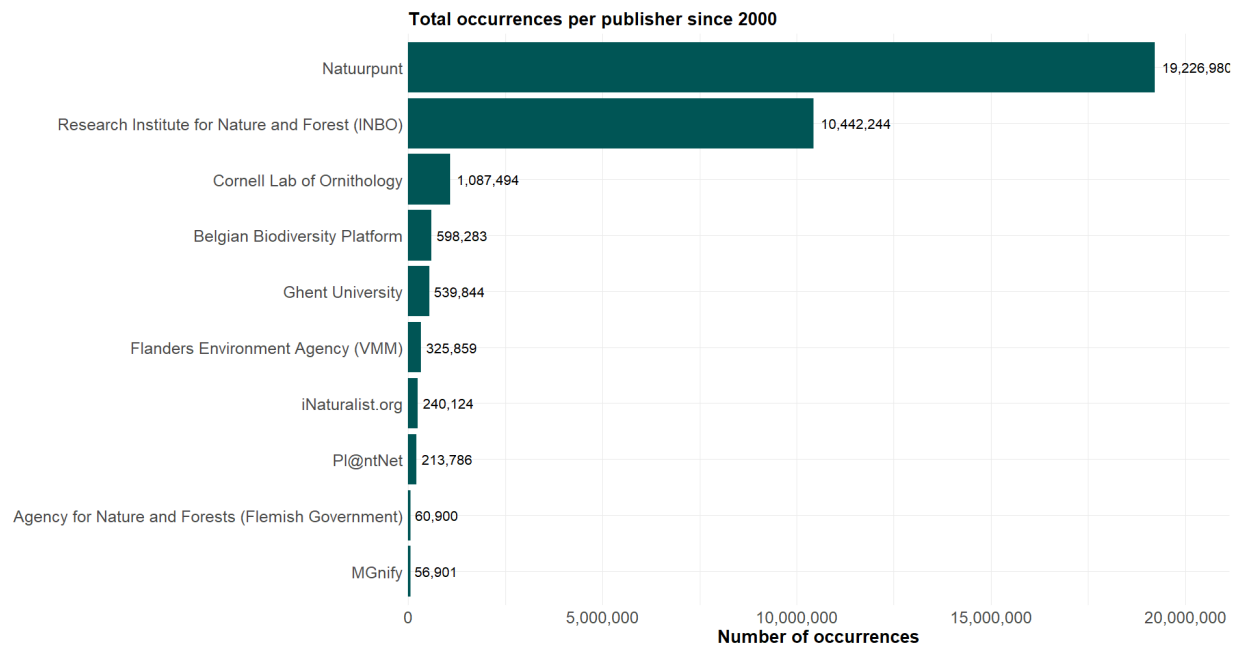


Figure 6: Overview of the top 10 publishers contributing records to GBIF in Flanders since 2000. Data source: [GBIF.org](https://www.gbif.org) (2026a).

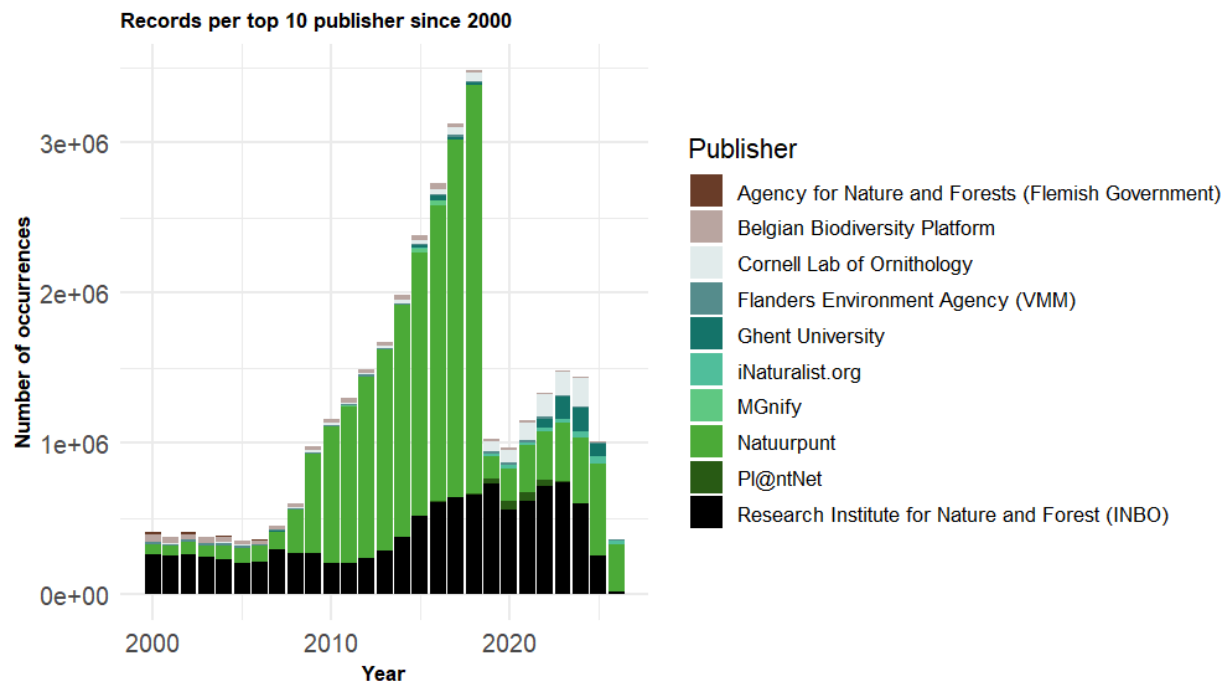


Figure 7: Detailed time series of the number of records published on GBIF by each of the 10 most important publishers since 2000 in Flanders. Data source: [GBIF.org](https://gbif.org) (2026a).

When examining the number of records per kingdom per year in Flanders (Figure 8), it is evident that animals and plants account for the majority of recorded observations. The considerably lower number of records for other kingdoms suggests differences in recording effort and detection probability among taxonomic groups. This pattern may partly be explained by the higher visibility and public interest in plants and animals, which are more frequently targeted by both professional surveys and citizen science initiatives.

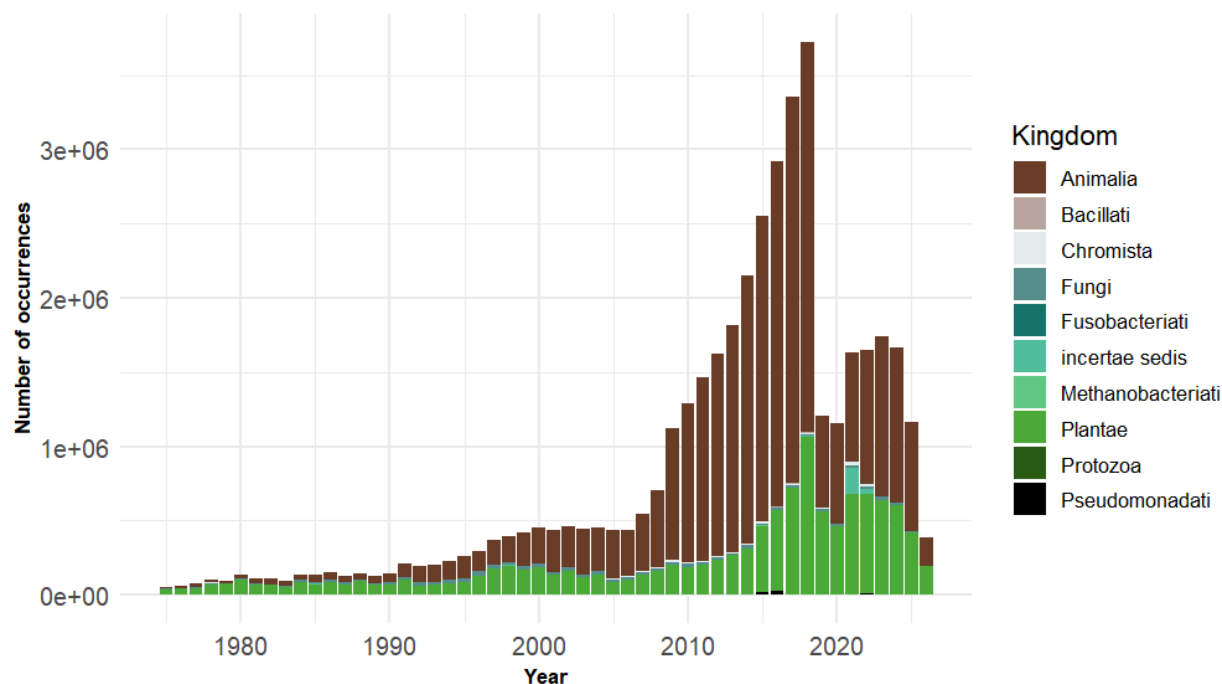


Figure 8: Comparison of the number of records per kingdom per year in Flanders. Only the 10 most frequent kingdoms in Flanders are displayed. Data source: [GBIF.org](https://gbif.org) (2026a).

1.4. Goal of the Flanders use case

The main objective of the Flanders use case is to transform the large volume of biodiversity data available for Flanders on GBIF into informative indicators that can support evidence-based policymaking.

To facilitate access to these outputs, all indicators presented in this deliverable (unless stated otherwise) are publicly available through [an online dashboard](#) (Figure 9).

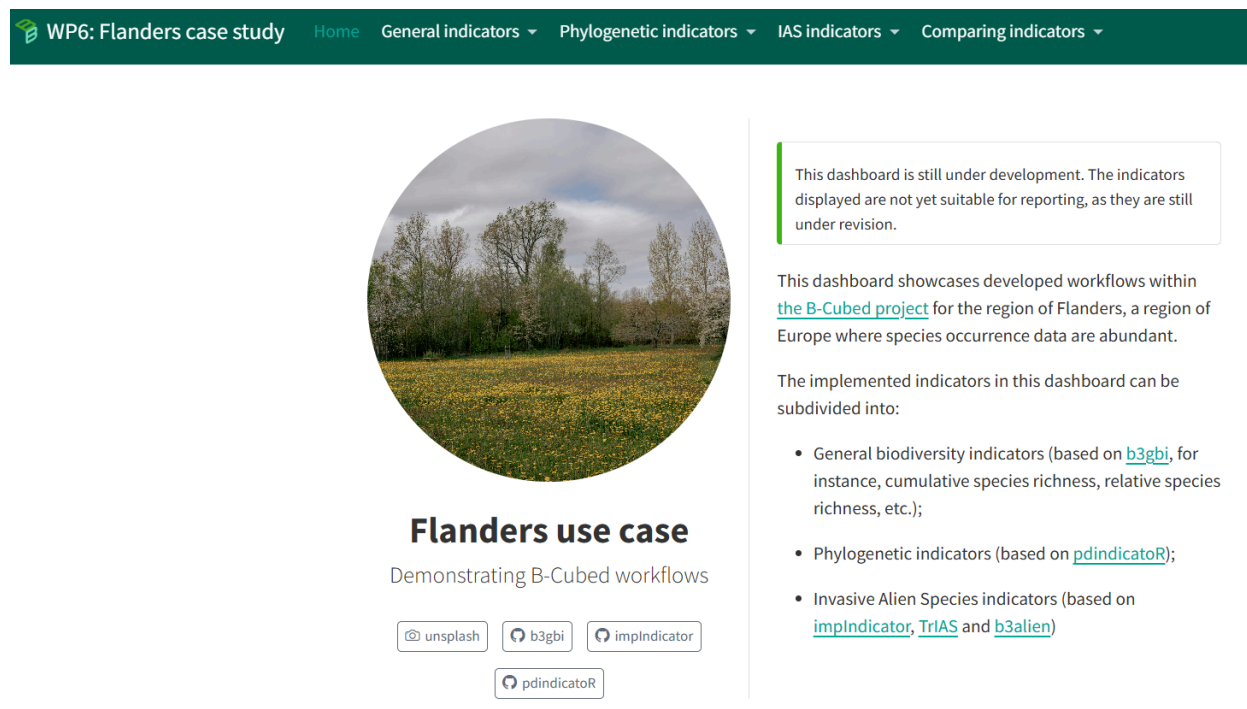


Figure 9: The homepage of the Flanders use case dashboard, accessible via <https://flanders.b-cubed.eu/>. The corresponding GitHub repository is <https://github.com/b-cubed-eu/flanders-use-case>.

The dashboard was developed with four primary objectives. First, the development of this dashboard provided an opportunity to test and refine the existing B3 workflows. At the same time, the added value of incorporating these workflows into existing biodiversity data infrastructures and pipelines is being evaluated. Second, this dashboard is intended to showcase the various indicators that can be calculated by applying the B3 workflows to openly available biodiversity data from Flanders. As such, it facilitates the exploration of GBIF data by researchers considering its use in their work, as well as by policy makers evaluating biodiversity data at the Flemish level for reporting to the EU and other biodiversity monitoring and reporting frameworks. Third, the dashboard provides a platform for demonstrating these indicators to local stakeholders. Finally, the dashboard may serve as a source of inspiration for existing biodiversity dashboards to incorporate the newly developed B3 workflows. This could expand the range of indicators they provide and improve the efficiency of their current analytical workflows. Examples of such platforms include the Flemish Biodiversity Portal (<https://natuurdatabank.be/>), the Belgium GBIF node (<https://gbif.biodiversity.be/>) (Figure 10), and the Alien Species Portal (still under development). Close collaboration was maintained with the managers of the Belgium GBIF node to ensure that the outputs are interoperable and can be readily integrated into their platform.

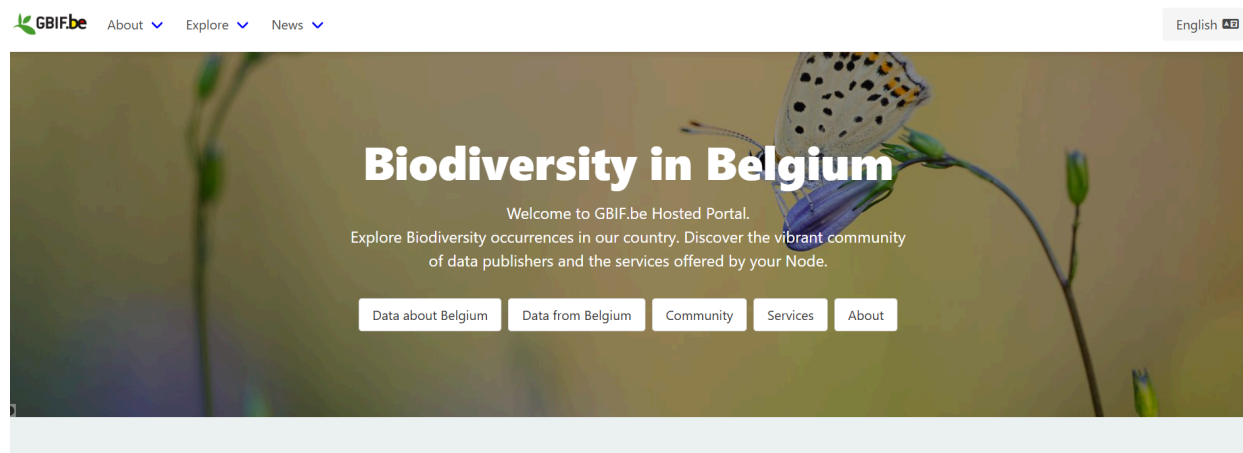


Figure 10: Homepage of the Belgium GBIF node, whose maintainers we are in close contact with to facilitate the incorporation of B3 indicators, such as observed species richness maps for several taxonomic groups (e.g. birds, mammals, plants), into their platform.



2. Methodology

All analyses were performed in R unless mentioned otherwise (R Core Team, 2026).

2.1. Creating cubes with rgbif

As an example, the SQL query used to download the species occurrence cube for native flowering plants in Flanders is provided below ([GBIF.org](https://gbif.org), 2026b).

```
SELECT
  kingdom,
  kingdomkey,
  phylum,
  phylumkey,
  class,
  classkey,
  "order",
  orderkey,
  family,
  familykey,
  genus,
  genuskey,
  species,
  specieskey,
  "year",
  GBIF_EEARGCODE(
    1000,
    decimallatitude,
    decimallongitude,
    COALESCE(coordinateuncertaintyinmeters,1000)
  ) AS eeacellcode,
  MIN(COALESCE(coordinateUncertaintyInMeters, 1000)) AS minCoordinateUncertaintyInMeters,
  MIN(GBIF_TEMPORALUNCERTAINTY(eventDate, eventTime)) AS minTemporalUncertainty,
  IF(ISNULL(orderkey), NULL, SUM(COUNT(*)) OVER (
    PARTITION BY orderkey, GBIF_EEARGCODE(1000, decimallatitude, decimallongitude,
    COALESCE(coordinateuncertaintyinmeters, 1000)), "year")) ordercount,
  IF(ISNULL(familykey), NULL, SUM(COUNT(*)) OVER (
    PARTITION BY familykey, GBIF_EEARGCODE(1000, decimallatitude, decimallongitude,
    COALESCE(coordinateuncertaintyinmeters, 1000)), "year")) familycount,
  IF(ISNULL(genuskey), NULL, SUM(COUNT(*)) OVER (
    PARTITION BY genuskey, GBIF_EEARGCODE(1000, decimallatitude, decimallongitude,
    COALESCE(coordinateuncertaintyinmeters, 1000)), "year")) genuscount,
  COUNT(*) AS occurrences

FROM occurrence

WHERE
  GBIF_Within('POLYGON ((2.327344 51.09871, ..., 2.327344 51.09871))', decimallatitude,
  decimallongitude) = TRUE
AND "year" >= 1975
  AND occurrenceStatus = 'PRESENT'
  AND (coordinateUncertaintyInMeters <= 1000 OR coordinateUncertaintyInMeters IS NULL)
  AND NOT occurrence.basisofrecord IN ('FOSSIL_SPECIMEN', 'LIVING_SPECIMEN')
  AND NOT GBIF_STRINGARRAYCONTAINS(issue, 'ZERO_COORDINATE')
  AND NOT GBIF_STRINGARRAYCONTAINS(issue, 'COORDINATE_OUT_OF_RANGE')
  AND NOT GBIF_STRINGARRAYCONTAINS(issue, 'COORDINATE_INVALID')
  AND NOT GBIF_STRINGARRAYCONTAINS(issue, 'COUNTRY_COORDINATE_MISMATCH')
```



```

AND speciesKey IS NOT NULL
AND decimalLatitude IS NOT NULL
AND decimalLongitude IS NOT NULL
AND occurrence.hasgeospatialissues = FALSE
AND NOT GBIF_STRINGARRAYCONTAINS(occurrence.issue, 'TAXON_MATCH_FUZZY', TRUE)
AND (occurrence.taxonkey IN ('196', '220') OR occurrence.acceptedtaxonkey IN ('196',
'220') OR occurrence.kingdomkey IN ('196', '220') OR occurrence.phylumkey IN ('196',
'220') OR occurrence.classkey IN ('196', '220') OR occurrence.orderkey IN ('196',
'220') OR occurrence.familykey IN ('196', '220') OR occurrence.genuskey IN ('196',
'220') OR
occurrence.specieskey IN ('196', '220'))

GROUP BY
    kingdom, kingdomkey,
    phylum, phylumkey,
    class, classkey,
    "order", orderkey,
    family, familykey,
    genus, genuskey,
    species, specieskey,
    "year",
    GBIF_EEARGCODE(
        1000,
        decimalLatitude,
        decimalLongitude,
        COALESCE(coordinateuncertaintyinmeters,1000)
    )

```

In this SQL query, the following settings are defined, partially based on the guidelines of Langerhaert et al. (2026):

- `coordinateUncertaintyInMeters` is not larger than the grid size of the reference grid used for spatial aggregation (in this case 1000 m). Otherwise, circular artefacts appear in the landscape by spatial randomization (Figure 11).
- The uncertainty radius to apply to a point with missing coordinate uncertainty (`COALESCE(coordinateUncertaintyInMeters, 1000)`) is set at the grid size of the cube (e.g. in this case 1000 m)
- `minTemporalUncertainty` is calculated per grid cell and filtered to be less than 315360000 seconds (ten years). (it is not yet possible to filter directly on this variable, but only after the cube is created).
- Ensure occurrence status is present (`occurrenceStatus = 'PRESENT'`).
- Ensure the origin of the record is not a fossil or living specimen (`NOT occurrence.basisofrecord IN ('FOSSIL_SPECIMEN', 'LIVING_SPECIMEN')`)
- Ensure speciesKey, latitude, and longitude are defined (`AND speciesKey IS NOT NULL AND decimalLatitude IS NOT NULL AND decimalLongitude IS NOT NULL`)
- Ensure there are no geospatial issues (`occurrence.hasgeospatialissues = FALSE`)
- To avoid edge effects, we added a 15 km buffer around the borders of Flanders. Because the original administrative polygon contains internal holes (areas enclosed



within the boundary but not included in the polygon), these were removed prior to analysis to create a continuous study area. We then used the resulting polygon to select occurrence records within the study area (`GBIF_Within('POLYGON ((2.327344 51.09871, ..., 2.327344 51.09871))', decimalLatitude, decimalLongitude) = TRUE`) (see Figure 12).

After cube creation, an extra filtering step is performed. In order to retain only species native to Flanders, we used a published checklist with all observed species in Belgium between January 1, 1643 and December 30, 2023 (Swinen et al., 2025). Only species indicated in this checklist as having `establishmentMeans` equal to native were retained.

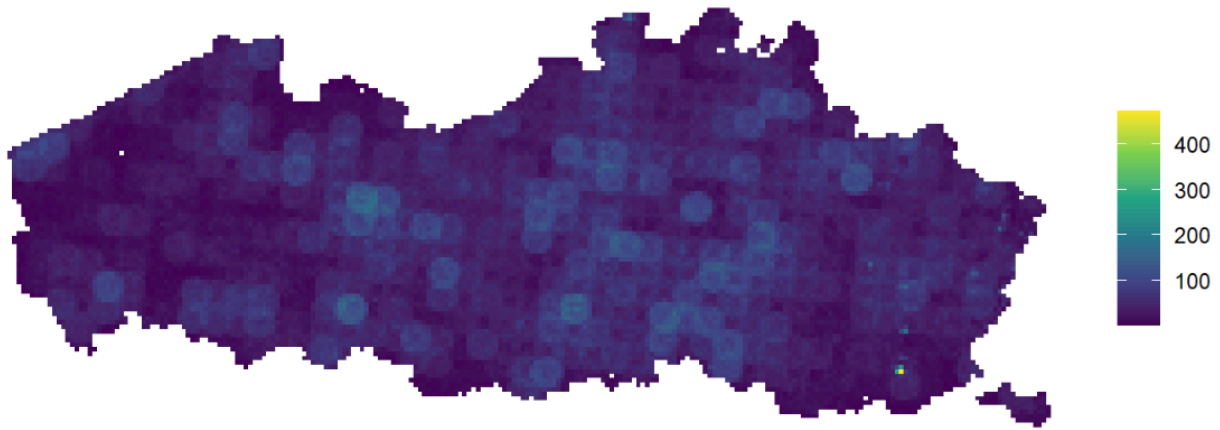


Figure 11. Spatial distribution of observed insect species richness in Flanders at a 1-km resolution when spatial randomisation is permitted and observations with high spatial uncertainty are retained. Distinct circular hotspots of high species richness emerge across the landscape as a result of these settings. Data source: [GBIF.org](https://www.gbif.org) (2026c).

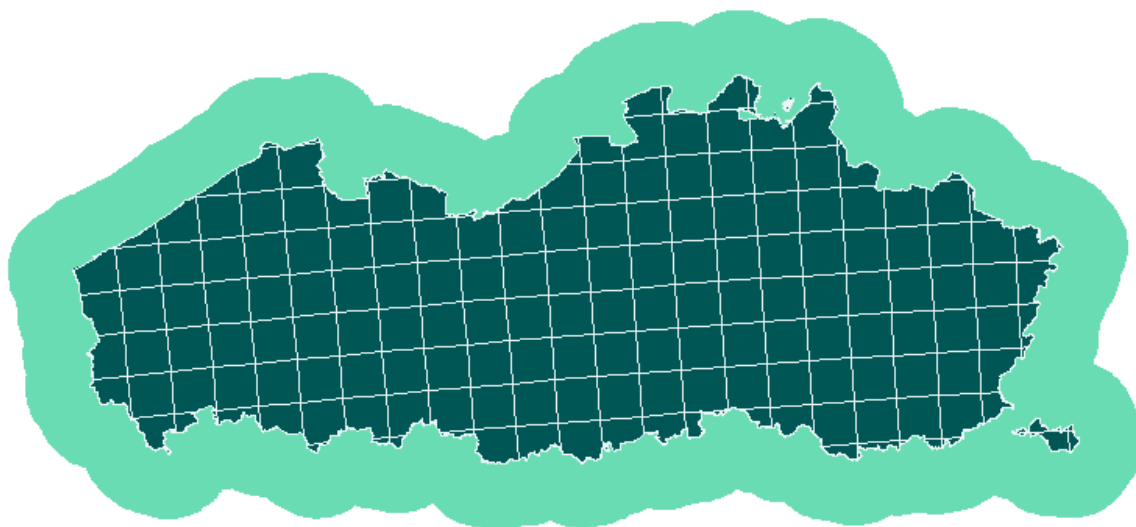


Figure 12: To generate cubes based on GBIF occurrences for the Flanders case study, we first removed all internal holes (areas enclosed within the administrative boundary but not included in the polygon, e.g. Brussels-Capital Region) from the Flanders polygon to create a continuous study area (dark green). Next, we added a 15 km buffer to create the final polygon (light green). A 15 km buffer is recommended when generating cubes at 10 km resolution as the diagonal of a 10 by 10 km grid cell is 14.14 km. After downloading the cube at EEA 1 km or 10 km resolution, only the grid cells that overlapped with Flanders were selected within the cube.

A general introduction to downloading cubes using the GBIF Occurrence SQL Download API can be found [here](#) (Desmet, 2025), together with a complete overview of [all available functions](#).

2.2. Performing (dubi)cube quality checks

The quality of the flowering plants cube at 1 km² resolution was checked using the function `diagnose_cube()` of the **dubicube** R-package (Langerhaert & Van Daele, 2026a). The analysis of a set of diagnostic rules resulted in the following outcome:

```
Data cube diagnostics
-----
● OK - temporal_min_years
  Cube contains observations across 52 years.

● OK - temporal_missing_years
  Cube contains 0 missing years.

● OK - spatial_min_cells
  Cube contains observations across 14365 grid cells.
```



- OK - spatial_max_uncertainty
Cube contains 0 records where the coordinate uncertainty is larger than the grid cell resolution.
- OK - spatial_miss_uncertainty
Cube contains 0 records with missing coordinate uncertainty.
- OK - taxon_min_taxa
Cube contains observations across 1185 taxon keys.
- OK - obs_min_records
Cube contains 3240164 observation records (rows).
- OK - obs_min_total
Cube contains a total of 5198912 observations.

No caveats were identified in the diagnostic assessment. However, a detailed inspection of occurrence distributions revealed that the number of occurrences per grid cell is highly right-skewed (see Figure 13). Since the occurrence data are count data, a log-normal distribution is to be expected. A small number of grid cells contained exceptionally high numbers of occurrences (up to 25 000), reflecting intensive sampling effort associated with targeted research projects. These extreme values were considered ecologically meaningful and were therefore retained rather than removed. To improve interpretability, this indicator is visualised using a logarithmic color scale. After transforming the data (log-normal), this indicator is clearly normally distributed (Figure 14).

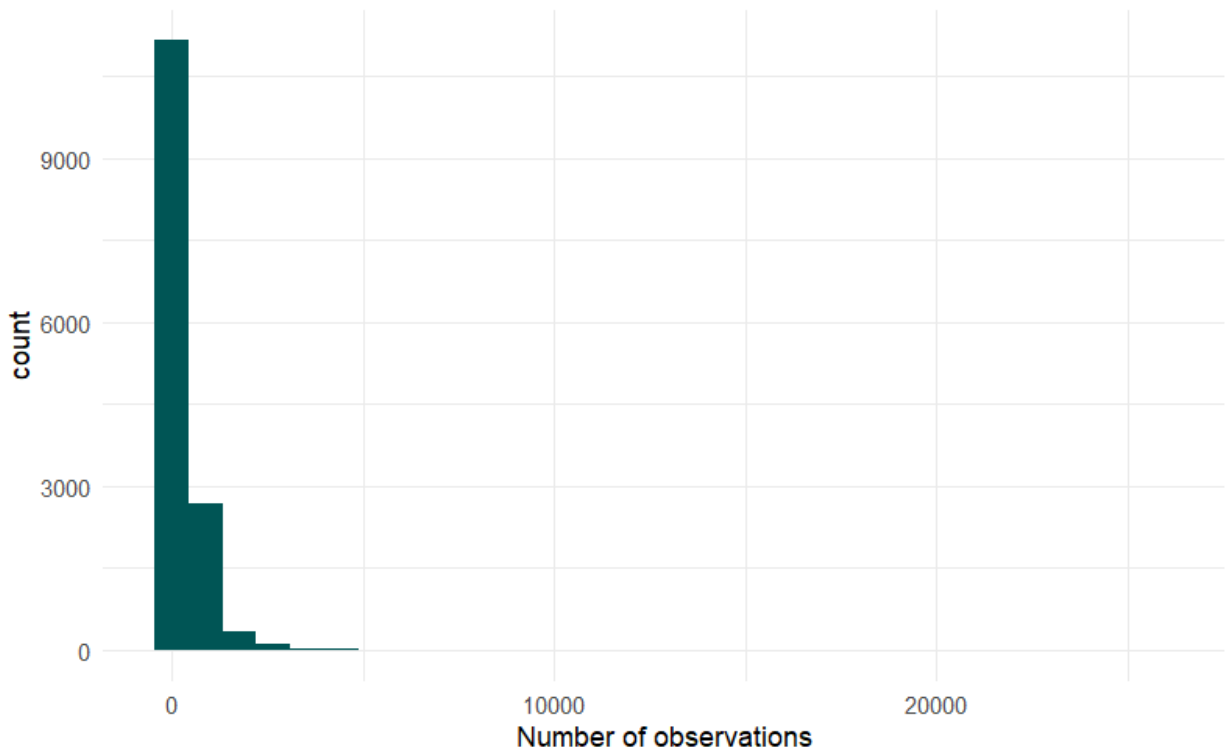


Figure 13: Frequency distribution of non-transformed number of observations for native flowering plants in Flanders at EEA 1-km resolution. This distribution is heavily right skewed.

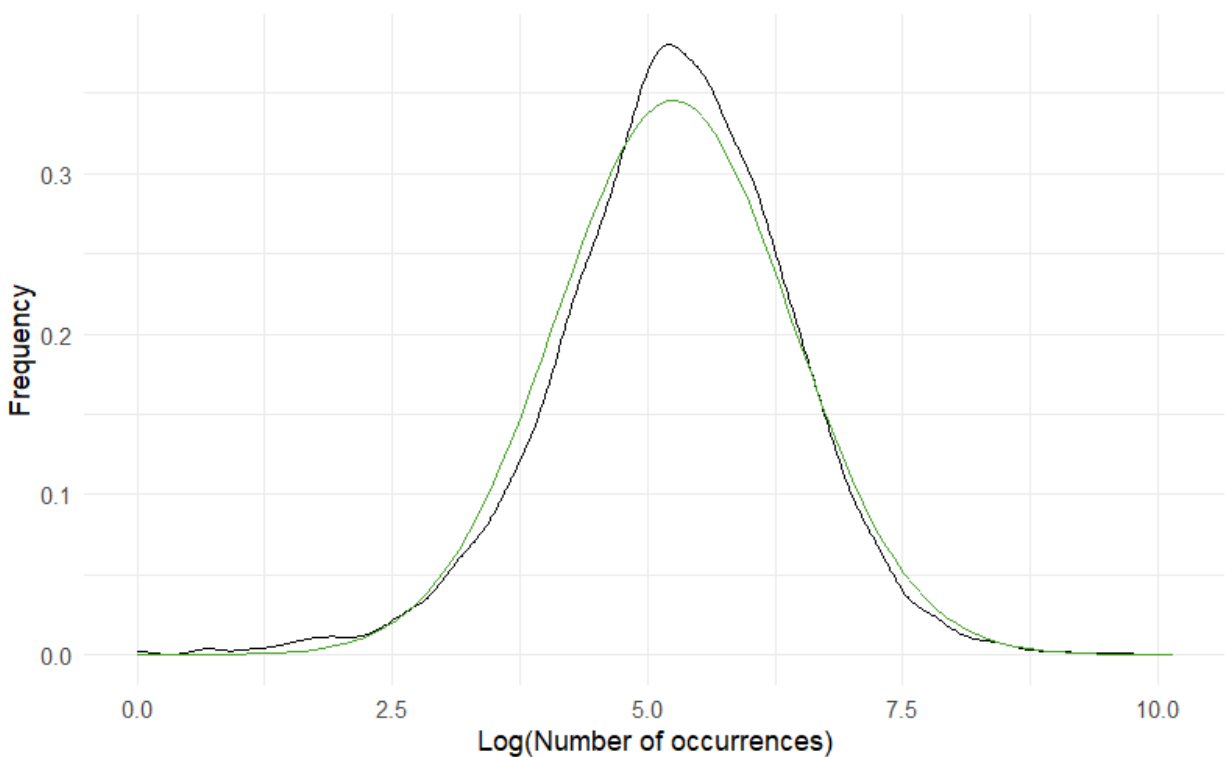




Figure 14: Illustration of frequency distribution after log-transformation of number of observations for native flowering plants in Flanders at EEA 1-km resolution. Green: a normal distribution, black: density function of the actual distribution in the log scale.

2.3. Overview of indicators tested

The tested packages in the Flanders use case for calculating indicators are the following: **b3gbi** (Dove., 2026), **pdindicatorR** (Breugelmans et al., 2025), **trias** (Oldoni et al., 2026), **b3alien** (Trekels & Groom, 2025), **dissmapr** (MacFadyen & Hui, 2026) and **invasimapr** (MacFadyen et al., 2025). **dubicube** (Langerlaert & Van Daele, 2026a) was also applied to perform quality checks and calculate uncertainties (Table 3).

Table 4 provides an overview of all B3 indicators tested per package in this case study, including their corresponding taxonomic focus, the spatial resolutions applied, whether each indicator is included in the dashboard, and whether it is calculated separately for Natura 2000 and non-Natura 2000 areas to enable comparison. By default, data are included from 1975 onward, unless stated otherwise.

Furthermore, the following indicators of **b3gbi** were also tested but were skewed by survey effort bias in ways that were difficult to interpret, and therefore were excluded from the dashboard. For example, Species Richness, Hill-Shannon Diversity, and Hill-Simpson Diversity explicitly correct for survey effort; however, they are still subject to effort bias with the extreme temporal fluctuations of GBIF data. We opted to include less complex metrics like observed richness because the effort bias is more predictable and transparent.

- Abundance-Based Rarity
- Area-Based Rarity
- Taxonomic Distinctness
- Species Richness (Estimated by Coverage-Based Rarefaction)
- Hill-Shannon Diversity (Estimated by Coverage-Based Rarefaction)
- Hill-Simpson Diversity (Estimated by Coverage-Based Rarefaction)
- Pielou's Evenness and Williams' Evenness

The case study proved highly valuable, as it highlighted the need to expand the **b3gbi** package with additional indicators: (i) Completeness, (ii) Species Richness Density, and (iii) Relative Occupancy.

Further, extra indicators related to invasive alien species were also included (see Table 5). These are calculated based on either GBIF data cubes or the Belgium GRIIS checklist (Desmet et al., 2025) and are visualised on [the Alien Species Portal](#).

**Table 3: Overview of tested B3 packages, a short description and their GitHub repository.**

Package	Description	GitHub repository
	Analyze biodiversity trends and spatial patterns from GBIF data cubes, using flexible indicators like richness, evenness, and more.	https://b-cubed-eu.github.io/b3gbi/
	An R package to calculate spatial phylogenetic diversity distribution maps based on occurrence cubes and user-provided phylogenetic trees	https://github.com/b-cubed-eu/pdindicatorR/
	A repository containing pipelines to generate alien species indicators (checklist-based indicators or occurrence-based indicators)	https://trias-project.github.io/indicators/
	A Python package to calculate the Global Biodiversity Framework Target 6 Headline Indicator	https://pypi.org/project/b3alien/
	An R package for Compositional Dissimilarity and Biodiversity Turnover Analysis	https://b-cubed-eu.github.io/dissmapr/
	An R package that quantifies and maps community-level invasion fitness and site-specific invasibility.	https://b-cubed-eu.github.io/invasimapr/
	An R package that includes functions for calculating indicator uncertainty using bootstrapping, as well as tools for interpreting and visualising uncertainty.	https://b-cubed-eu.github.io/dubicube/



Table 4: Overview of tested B3 indicators per package and their attributes. The dashboard default refers to the Flanders use case (<https://flanders.b-cubed.eu/>). The Alien Species Portal is a separate dashboard, accessible by <https://alienspecies.inbo.be>.

Package	Indicator tested	Species group	Included in dashboard	Resolution	Natura2000 comparison
	Observed species richness Total occurrences Cumulative species richness Density of occurrences Mean year of occurrence (newness) Occupancy turnover Relative occupancy	Flowering plants Florabank dataset Insects Birds	✓	1 km 10 km	✓
	Phylogenetic diversity Phylogenetic diversity loss map	Flowering plants	✓	1km 5 km	✓
	Occupancy GAM's	GRIIS checklist	✓ (Alien Species Portal)	1km	✓



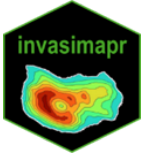
	Target 6	GRIIS checklist	✓	-	-
	Species invasiveness Site invasibility Trait invasiveness	Plants	-	10 km	-

Table 5: Overview of extra indicators included in the [Alien Species Portal](#). Per indicator, the data source is provided.

Indicator name	Data source
Map with occupied grid cells (occupancy)	GBIF data cube
Annual number of occupied grid cells	GBIF data cube
Invasion history	GBIF data cube
Management map	INBO data or GBIF data cube (depending on species)
Management timeline	INBO data or GBIF data cube (depending on species)
Number of new introductions of alien species per year	Belgian GRIIS checklist (Desmet et al., 2025)
Cumulative number of introduced species	Belgian GRIIS checklist (Desmet et al., 2025)
Pathways associated with alien species introductions	Belgian GRIIS checklist (Desmet et al., 2025)
Trend in occupancy between baseline and reporting period	Belgian GRIIS checklist (Desmet et al., 2025)
The number of introduced taxa per introduction pathway	Belgian GRIIS checklist (Desmet et al., 2025)



2.3.1. A closer look at **invasimapr** and **dissmapr**

Since **invasimapr** (WP4) requires a series of preprocessing steps before it can be applied, partly based on the **dissmapr** package (WP4), this section provides a brief overview of the overall workflow.

Invasimapr is an open-source R package designed to quantify and map community-level invasion fitness and site-specific invasibility (MacFadyen et al., 2025). As part of this case study, the **invasimapr** workflow was tested on all plant species observed in Flanders during the last 10 years ([GBIF.org](https://gbif.org), 2026d). The main goal of this part was to test the reproducibility of the workflow. Therefore, the workflow uses plant trait data from the TRY database as its default source when users do not provide their own trait dataset (Kattge et al., 2020).

The Invasibility Cube framework is a way to predict whether a species can successfully establish in a new ecosystem by integrating three things: species identity, environmental conditions, and functional traits. All species are first translated into a “trait space,” where each species is represented by a point based on its characteristics (such as growth rate, size, or stress tolerance). This reduces complex biological differences into a structured geometric space using dimensionality reduction techniques such as Non-metric Multidimensional Scaling (NMDS). Within this trait space, the existing community of species forms a shape called a convex hull. This hull represents the range of trait combinations already occupied by resident species and can be seen as the “competitive boundary” of the ecosystem.

Invasion success is then described in terms of invasion fitness, which is the per-capita growth rate of a rare introduced species in a system at equilibrium. If this value is positive, the species can increase and potentially establish; if negative, it will fail.

This invasion fitness is determined by three main components:

1. **Abiotic suitability:** how well the invader’s traits match the local environmental conditions.
2. **Niche crowding:** how similar the invader is to existing species in trait space, which increases competition when overlap is high.
3. **Resident competition:** how saturated the community already is with established species.

Overall, the framework predicts invasion outcomes by combining environmental filtering and competitive interactions into a unified geometric structure in trait space. Geometrically, species located near the center of the resident convex hull experience strong competition because they resemble many existing species. Species near the edges may face less competition and find unoccupied niche space, while species outside the hull may represent novel strategies but are not necessarily well matched to the environment. The framework can be described as a three-dimensional species × environment × trait cube that is mapped into a reduced trait space using NMDS, with the convex hull delineating the occupied niche space of the resident community (Figure 15).

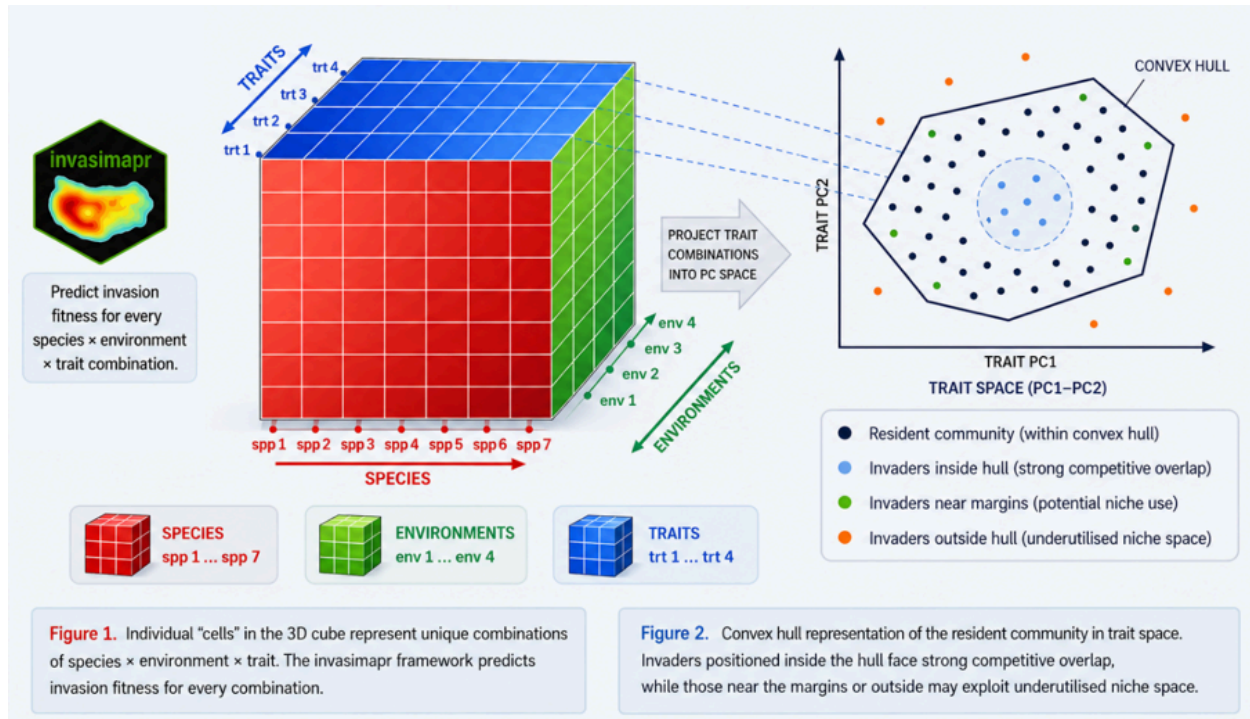


Figure 15: The invasimapr workflow, source: <https://b-cubed-eu.github.io/invasimapr/>.

Environmental raster data for the selected sampling sites or species locations can be retrieved and processed using `dissmapr::get_enviro_data()`, which supports environmental layers obtained through the **geodata** package or supplied as local raster data (MacFadyen & Hui, 2026). The function constructs a buffered area of interest, crops and optionally resamples the environmental rasters, extracts values at each site, fills isolated missing values, and returns the processed raster stack together with a site-level environmental data table (Figure 16).

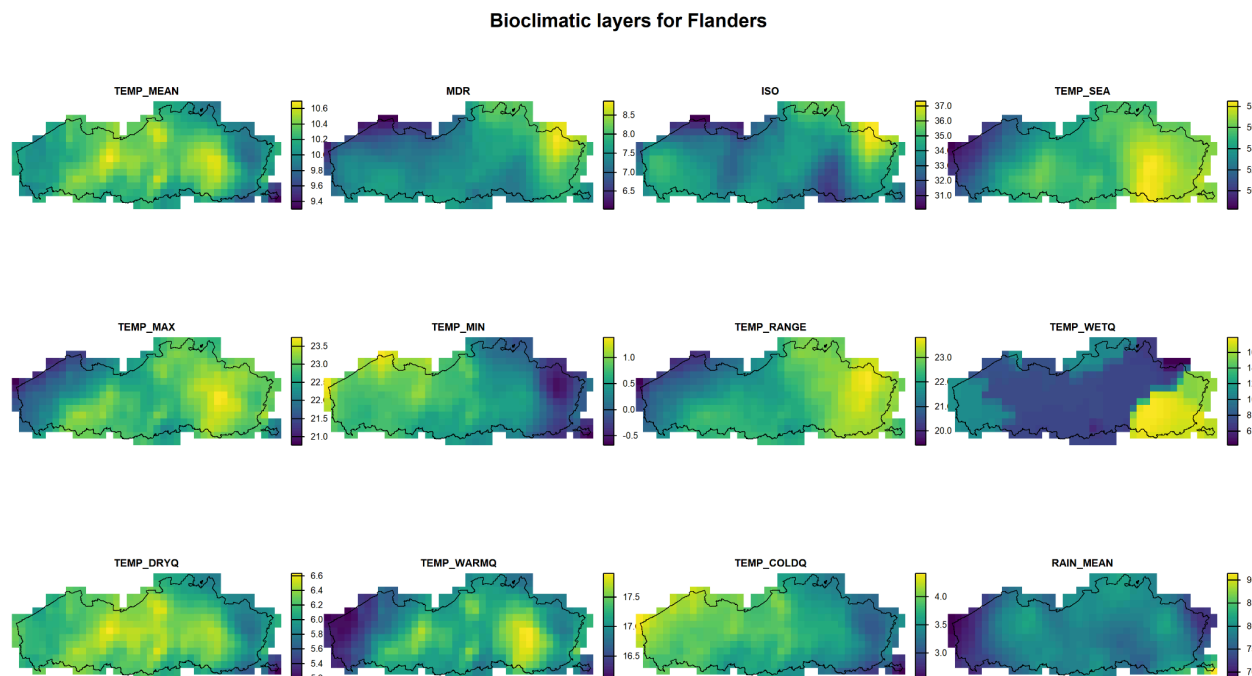


Figure 16: Overview of bioclimatic layers for Flanders, used to derive environmental traits of the invasibility cube. Environmental layers are obtained through the geodata package (Hijmans, 2026). (TEMP_MEAN: Annual Mean Temperature, MDR: Mean Diurnal Range , ISO: Isothermality, TEMP_SEA: Temperature Seasonality , TEMP_MAX: Max Temperature of Warmest Month, TEMP_MIN: Min Temperature of Coldest Month, TEMP_RANGE: Temperature Annual Range , TEMP_WETQ: Mean Temperature of Wettest Quarter, TEMP_DRYQ: Mean Temperature of Driest Quarter, TEMP_WARMQ: Mean Temperature of Warmest Quarter, TEMP_COLDQ: Mean Temperature of Coldest Quarter and RAIN_MEAN: Annual Precipitation)

2.4. Natura 2000 versus non-Natura 2000

2.4.1. Comparison of indicators

Natura 2000 is a network of protected areas across the European Union, established under the Birds and Habitats Directives to ensure the long-term conservation of Europe's most valuable and threatened species and habitats (European Commission, 2025). To compare indicators between Natura 2000 and non-Natura 2000 areas, Natura 2000 areas were intersected with the EEA-1km grid. Grid cells containing any part of a Natura 2000 area were classified as Natura 2000 cells, whereas cells without overlap with Natura 2000 areas were classified as non-Natura 2000 cells (Figure 17). The resolution of 1km for the EEA grid was chosen as data availability and accuracy are limited at finer spatial scales, whereas at coarser scales species occurrences become less clearly associated with the proportion of Natura 2000 area within a grid cell. A similar approach was adopted in the TrIAS Project (Groom et al., 2022). Both areas designated under the Habitats Directive (Agentschap voor Natuur en Bos, 2014) and the Birds Directive (Agentschap voor Natuur en Bos, 2005) were included as Natura 2000 areas.

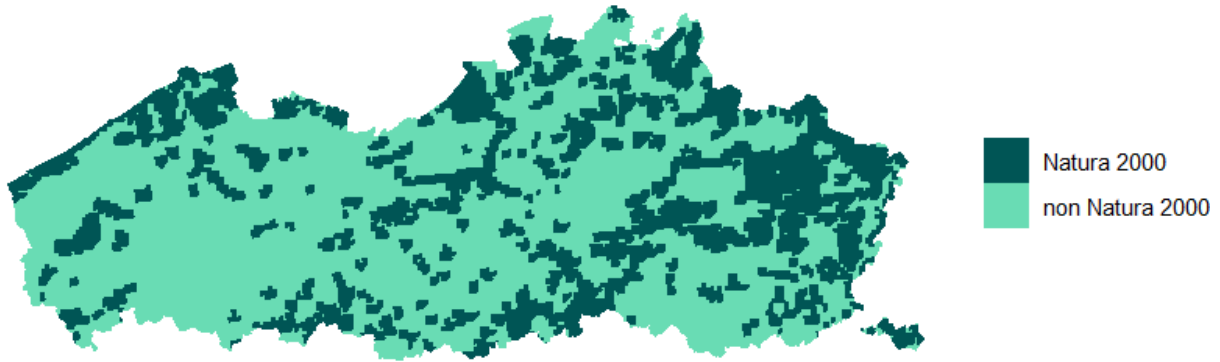


Figure 17: Distribution of Natura 2000 and non-Natura 2000 areas across EEA-1km grid cells for Flanders. Grid cells containing any part of a Natura 2000 area were classified as Natura 2000 cells, whereas cells without overlap with Natura 2000 areas were classified as non-Natura 2000 cells.

Species richness density and occurrence density (respectively, the number of species per km² and the number of occurrence records per km²; see further) were compared between Natura 2000 and non-Natura 2000 areas. These indicators were chosen since they are independent of the difference in surface area between the two areas. We only expect the largest area to have a more precise estimate. The number of species/occurrences was first calculated for each EEA-1km grid cell and year as, respectively, the number of unique plant species recorded, and the number of occurrences within that cell. After that, the mean number $\bar{S}$ was calculated per year t for each area. The means were then subtracted to get a final difference per year:

$$\Delta_t = \bar{S}_{\text{Natura2000},t} - \bar{S}_{\text{non-Natura2000},t}$$

Δ_t will be positive when the species richness/occurrence density is higher in Natura 2000 areas and negative when it is higher in non-Natura 2000 areas.

We adopted the workflow of Langerhaert & Van Daele (2026b) to evaluate the difference between the two area types in a robust and transparent way. This is done by evaluating the limits of confidence intervals against a null hypothesis baseline and pre-defined effect-size thresholds (effect classification). These functionalities are implemented in the **dubicube** R package (Langerhaert & Van Daele, 2026a). The code is available in the [Flanders Use Case](#) GitHub repository in the folder `_preprocessing/R/Natura2000`.

2.4.2. Confidence interval calculation

To obtain confidence limits, two complementary statistical approaches were tested to quantify differences between the two area types. First, a non-parametric bootstrap was used to estimate uncertainty directly from the observed data. Second, a generalized linear mixed model (GLMM) was fitted to estimate annual differences while accounting for the repeated observations of grid cells through time.



2.4.2.1. Bootstrap analysis

Although the **dubicube** package provides bootstrapping functionality, we could not use it here since the Natura and non-Natura 2000 area datasets have different sizes (different number of grid cells). Stratified bootstrapping could potentially be used but is not implemented in the package. Therefore, we coded the bootstrap ourselves and resampling was performed independently within each area type.

The bootstrap analysis was performed separately for each year. Within Natura 2000 and non-Natura 2000, grid cells were sampled independently with replacement from the observed set of grid cells. Each bootstrap sample contained the same number of grid cells as the original dataset for the corresponding habitat and year. Mean local species richness was calculated for both habitat types, and the bootstrap statistic Δ_t was calculated. This procedure was repeated 1,000 times for each year, yielding an empirical distribution of annual differences.

Percentile-based 95 % confidence intervals were obtained from the 2.5th and 97.5th percentiles of the bootstrap distribution using the **dubicube** package.

2.4.2.2. Model-based analysis

As a complementary analysis, annual differences in local species richness were estimated using a Gaussian generalized linear mixed model fitted with the **glmmTMB** package (Brooks et al., 2017; McGillicuddy et al., 2025). Although species richness represents count data, exploratory analyses showed substantial overdispersion relative to a Poisson distribution. A Gaussian model was therefore preferred, as it directly models mean species richness and yields annual habitat contrasts that are straightforward to interpret.

The model was specified as

$$Y_{ijt} \sim \mathcal{N}(\mu_{ijt}, \sigma^2)$$

where Y_{ijt} denotes the observed number of species/occurrences in grid cell i , area type j , and year t (as a factor variable). The expected richness was modelled as

$$\mu_{ijt} = \beta_0 + \beta_{\text{area}(j)} + \beta_{\text{year}(t)} + \beta_{\text{area} \times \text{year}(j,t)} + b_i$$

where

$$b_i \sim \mathcal{N}(0, \sigma_b^2)$$

is a random intercept accounting for repeated observations of the same grid cell. The interaction between area type and year allows the difference between Natura 2000 and non-Natura 2000 to vary through time.



Estimated marginal means were obtained for each habitat within each year, and annual differences between habitats were calculated as pairwise contrasts using the **emmeans** package (Lenth & Piaskowski, 2026). Approximate 95 % confidence intervals were derived using the normal approximation.

2.4.2.3. Differences between methods

In contrast to the bootstrap approach, which treats each year independently, the mixed-effects model estimates all years simultaneously and explicitly accounts for the repeated observation of the same grid cells through time. This allows the analysis to incorporate the hierarchical structure of the data and to obtain more stable annual estimates.

A random intercept was included for each 1-km grid cell. Ecologically, this accounts for the fact that some grid cells are consistently richer or poorer in plant species or number of occurrences than others, for example because of persistent differences in habitat quality, soil conditions, land use or other environmental characteristics. By modelling this baseline variation explicitly, annual comparisons between Natura 2000 and non-Natura 2000 are less influenced by intrinsically rich or poor cells.

In principle, the model-based framework also allows the inclusion of additional explanatory variables such as survey effort. We explored the use of total occurrence counts as a proxy for effort, but this variable was strongly confounded with the year $\times$ area type structure in the data. As a result, the effect of effort could not be estimated independently from the main temporal and habitat effects. The final model was therefore fitted without an explicit effort correction. The random intercept for grid cell also accounts for any consistent differences in recording intensity. However, it does not account for temporal variation in survey effort, such as changes in recording intensity between years or between habitats over time. Such variation requires an explicit measure of survey effort, which was not available for the present analysis.

The non-parametric bootstrap forms the primary basis for statistical inference in this study because it requires relatively few assumptions and quantifies uncertainty directly from the observed data. The mixed-effects model was included as a complementary, exploratory analysis to illustrate the potential of a hierarchical modelling framework that accounts for repeated observations of grid cells through time. This model based approach can be refined in future work through improved distributional assumptions, explicit modelling of spatial dependence, and the inclusion of additional explanatory variables, such as survey effort, where suitable data are available.

2.4.3. Effect classification

For both the bootstrap and model-based analyses, annual differences and their confidence intervals were interpreted using effect classification. Confidence intervals were evaluated against a reference value of zero and predefined effect-size thresholds (5 species or 15 occurrences), allowing annual differences to be classified as higher richness/occurrence density



in Natura 2000 compared to non-Natura 2000, lower density in Natura 2000, no meaningful difference, or uncertain difference (Table 6). This approach emphasizes the magnitude and uncertainty of the estimated effects rather than relying solely on statistical significance.

Table 6: Effect classification rules used. See also [this dubicube vignette](#).

Effect	Rule
Higher in Natura 2000	Confidence interval above reference value
No meaningful difference	Confidence interval between thresholds and contains reference value
Lower in Natura 2000	Confidence interval below reference value
Uncertain	Confidence interval contains reference value and at least one of the thresholds

2.5. How to deal with survey effort bias?

Dealing with the large bias in survey effort when interpreting and calculating indicators is challenging. We tested two possible ways of handling this bias: (i) calculating completeness and adding this as a mask when visualising other indicators and (ii) restricting the analysis to well-surveyed cells in the case of the Florabank dataset. The first approach was recommended based on the findings of Langerlaert et al. (2026).

2.5.1. Completeness

This indicator estimates how complete our knowledge of an ecological community is by calculating "Sample Coverage" (a concept originally developed by Turing and Good (1953), and popularised by Chao and Jost (2012)). Instead of just counting the raw number of observations, it estimates the proportion of all individuals in the actual community that belong to the species we have already detected in our sample. For maps, the value is calculated for each spatial cell using all historical data for that cell. For time-series analyses, the value is calculated independently for each time step over the complete data cube.

- High value (Value of 1.0 or 100%): The sample is almost complete, suggesting that further fieldwork or additional searches are unlikely to reveal new species. Although other species might be present, they are likely very rare and/or have extremely low detectability.
- Low value (Value of 0.2 or 20%, for example): The sample is incomplete. This suggests that while you have found some species, an estimated 80% of all individuals in that community belong to species that have not yet been detected or logged in the database.

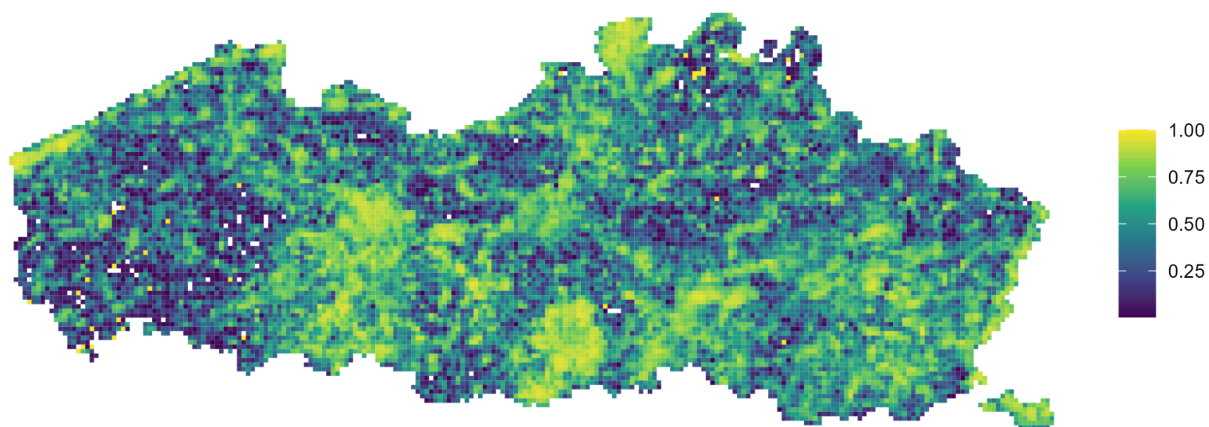


Figure 18: This figure illustrates the completeness of GBIF data for flowering plant communities in Flanders from 1975 to the present. Data source: [GBIF.org](https://www.gbif.org) (2026e).

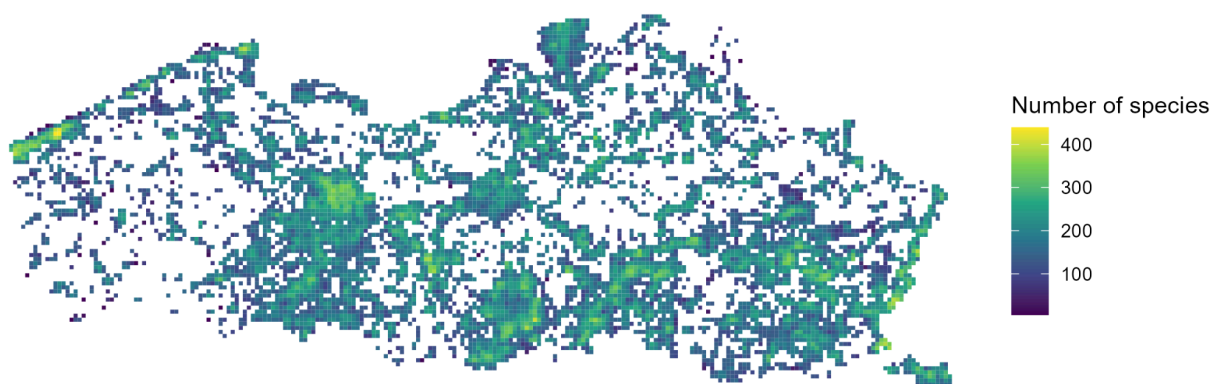


Figure 19: This figure demonstrates how completeness can be used as a mask when displaying other indicators. Here, the number of flowering plant species in Flanders since 1975 is shown, restricted to locations with completeness greater than 0.5. Data source: [GBIF.org](https://www.gbif.org) (2026e).

When calculating the completeness of each grid cell for native flowering plants (Figure 18), the mean completeness across Flanders was 0.49, while the median was 0.51. This value was used as the threshold for visualising other indicators derived from the same cube in the Flanders use-case dashboard. The higher this value is set, the fewer grid cells are included in the analysis. Conversely, the lower this threshold is set, the higher the likelihood that the observed patterns are still primarily driven by sampling effort.

Each indicator from this cube is presented both with and without a completeness-based mask, allowing users to assess the potential influence of survey effort on the results (Figure 19).

2.5.2. Well-surveyed cells



The disadvantage of the method described above is that a cell with high completeness across all years in the datacube may still have very low completeness in a specific year, which can lead to misleading results when visualising time trends. Therefore, we also tested the methodology applied by Florabank to investigate changes in relative occupancy over time. In this approach, only cells with more than 100 species observations within a given year are retained in the datacube. The advantage is that the remaining data originate from well-surveyed cells. Consequently, if a species is not recorded in a well-surveyed cell, it can be considered absent.



3. Results

3.1. General indicators

For each indicator, the indicator type, a general explanation, its policy relevance, and its limitations are described. These elements, together with an example, are presented in a dedicated factsheet for each indicator. The factsheets have been designed by Pensoft Publisher and will be made available through [Zenodo](https://zenodo.org/) (Hillaert et al., 2026). To ensure scientific accuracy and clarity, the factsheets were reviewed by researchers at INBO with a broad range of expertise.

In addition, the factsheets will be printed and distributed at several international conferences, including BEeS 2026 (<https://www.lifewatch.eu/bees-2026/>), Neobiota 2026 (<https://www.neobiota2026.org/>), and TDWG (<https://www.tdwg.org/>), to raise awareness of the B3 workflows developed for indicator calculation using openly available GBIF data.

The indicators displayed below are based on an occurrence cube of native flowering plants recorded from 1975 to the present, at either 1 km or 10 km resolution. For the production of indicator maps for Flanders, a 5 km resolution is considered most appropriate (Langeraert et al., 2026). However, this functionality was not yet available in **b3gbi** when the analyses were performed. In Flanders much of the long-term mapping and monitoring of the flora has been conducted using 4 km grid cells using the Lambert 1972 (EPSG: 31370) grid system. Therefore, most of the available data would be unsuitable for grids of 1km.

The other general indicators calculated for native flowering plants of the Florabank database, insects and birds can be accessed through [the dashboard](#) and will not be displayed in this section.

3.1.1. Total Occurrences

Indicator type

Data quality indicator.

Explanation

The total number of occurrences is calculated by summing the occurrences of all species observed for each cell or year. For maps, the value is calculated for each spatial cell. For time-series analyses, the value is calculated per time step over the complete data cube.

- High value: A high number of occurrences has been submitted, signalling a hotspot for human outdoor activity or nature reporting.
- Low value: Relatively few occurrence records have been submitted, potentially indicating data "blind spots" where few people pass by and report observations. These 'blind spots' are also not part of long-term monitoring programs.



Applicability for policy

This metric indicates the degree and locations of citizen science engagement. It also helps to interpret other metrics more correctly by highlighting biases in survey effort over time and space. While not itself an indicator, it is vital for interpreting the results of calculated indicators.

Limitations

It indicates which locations are preferred for outdoor recreation and nature recording, rather than which locations are of conservation value. Though you will tend to get lower numbers of observations in an area with fewer species.

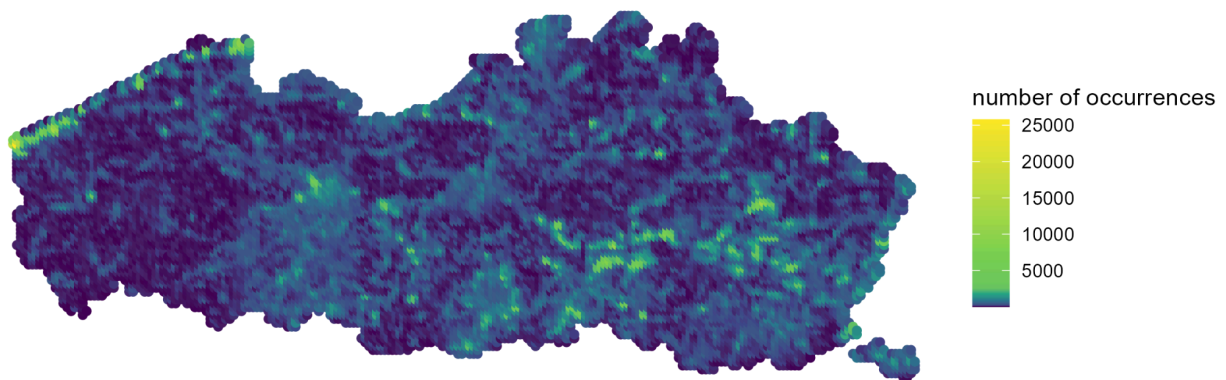


Figure 20: Spatial distribution of the total number of occurrence records of native flowering plants in Flanders reported since 1975. Data source: [GBIF.org](https://www.gbif.org) (2026e).

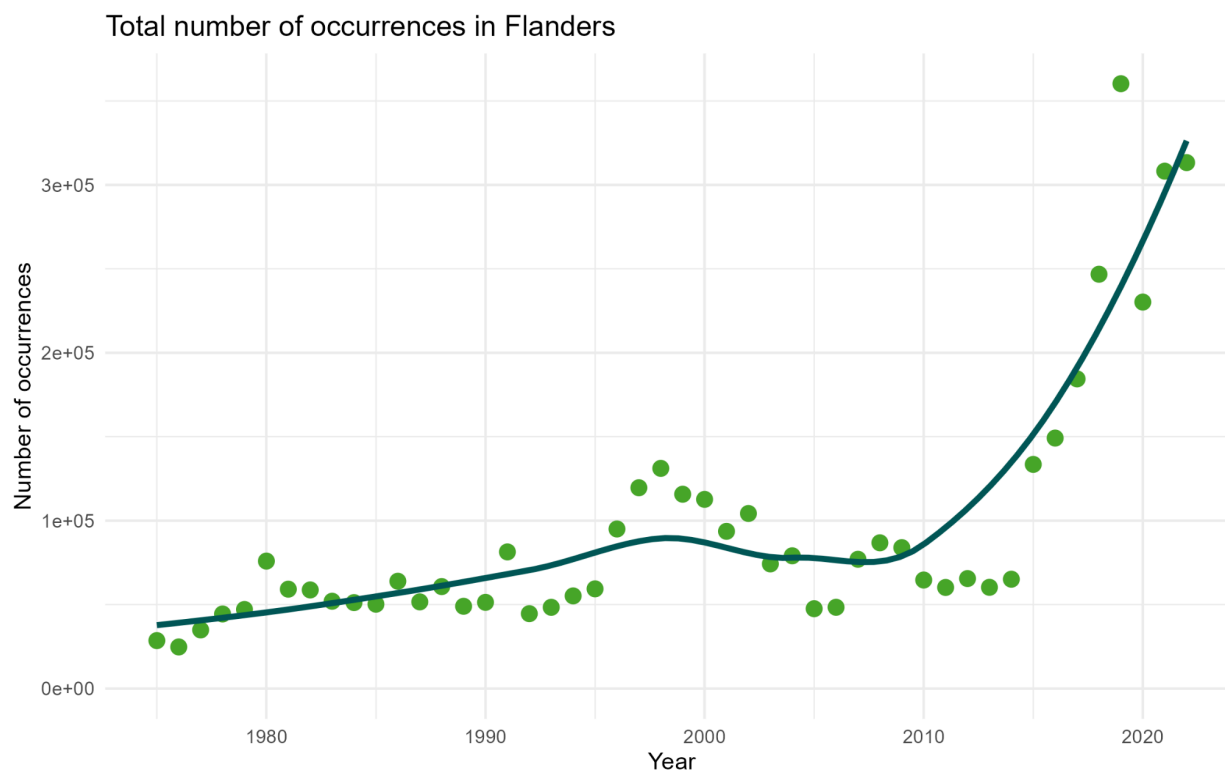


Figure 21: Annual total number of occurrence records of native flowering plants in Flanders, including a LOESS smoother. Data source: [GBIF.org](https://gbif.org) (2026e).

Interpretation

A few locations in Flanders have an exceptionally high number of occurrences in GBIF. These outliers may reflect intensive monitoring efforts and are therefore retained in the analyses. Overall, the spatial distribution of occurrences is consistent with known patterns in Flanders: high numbers of occurrences are found in well-known biodiversity hotspots, such as the Belgian dunes and Nationaal Park Hoge Kempen, whereas relatively few occurrences are recorded in central West Flanders, a region characterised by intensive agriculture and generally lower biodiversity.

From 1975 to 2010, the number of records shows a steadily increasing trend. Just before 2000, a modest increase in plant occurrence records is observed, likely reflecting intensified volunteer recording efforts in preparation for the publication of the Atlas of the Flora of Flanders (Van Landuyt, 2006). Since around 2008, there has been a sharp increase in the number of records, coinciding with a more active Florabank volunteer community and the establishment and expansion of Waarnemingen. After Waarnemingen stopped publishing data to GBIF in 2018, no clear decline in occurrence numbers is observed, as data collected by Florabank volunteers continue to be published to GBIF (Van Landuyt & Brosens, 2026).

3.1.2. Occupancy turnover



Indicator type

Trend / Pressure Indicator.

Explanation

Occupancy turnover measures changes in species composition over time, reflecting the rate at which species appear or disappear from a given area between two time steps in the data cube. The metric is calculated using the Jaccard dissimilarity index (Jaccard, 1901), which compares the species present in two time steps and quantifies their dissimilarity on a scale from 0 to 1.

- High values (closer to 1): indicate a complete turnover in species composition between the two time periods.
- Low values (closer to 0): indicate little or no change in species composition between the two time periods.

This indicator is available only as a time series.

Applicability for policy

- Ecosystem stability: This is a vital tool for assessing how ecosystems have responded to large-scale pressures like climate change, habitat fragmentation, land use change or nitrogen deposition.
- Evaluating effectiveness of management strategies: It allows managers to evaluate whether restoration policies are actually working. For example, if a province creates a new wetland, a controlled increase in turnover might indicate that the targeted wetland species are successfully moving in.
- Invasive species: Biological invasions can result in high turnover rates.

Limitations

This measure is largely affected by survey effort. Further, it must always be combined with rarity or taxonomic indices to understand the true nature of the turnover shift (positive or negative for local biodiversity). Turnover timescales also differ across organism types. Mobile, short-lived species such as birds and insects exhibit faster turnover rates than sessile, long-lived tree species.

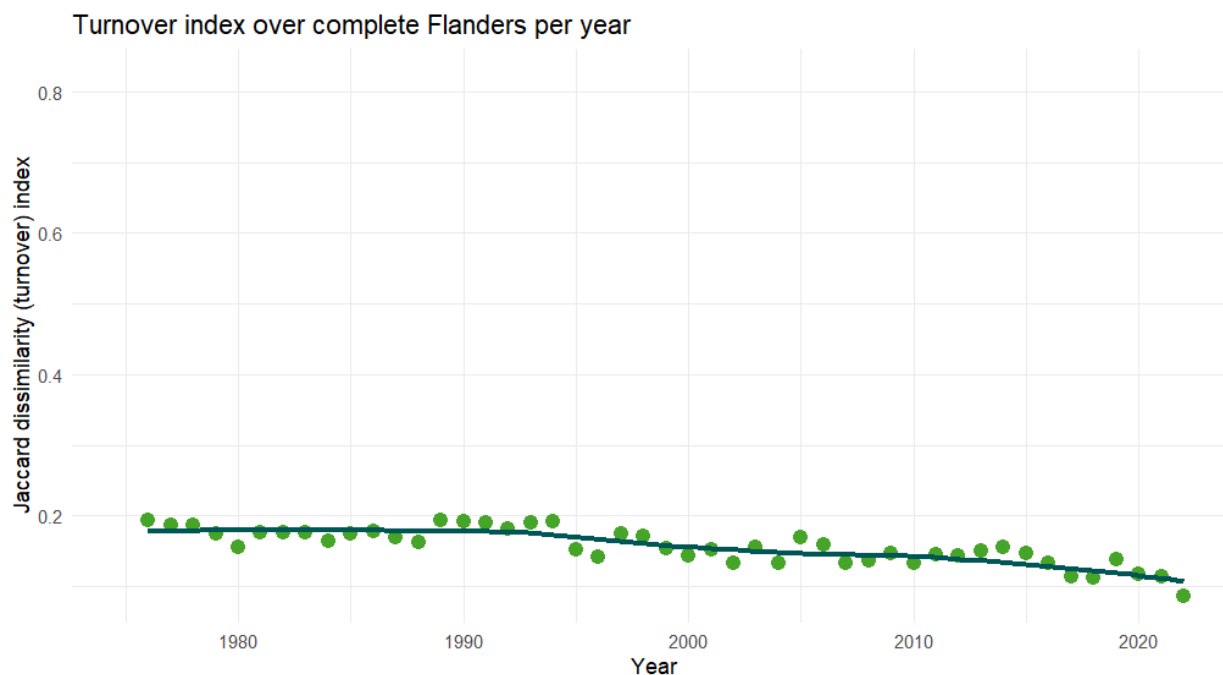


Figure 22: Annual turnover index of native flowering plants in Flanders, including a LOESS smoother. Data source: [GBIF.org](https://gbif.org) (2026e).

Interpretation

At the level of Flanders, this indicator is hard to interpret. It might be more informative when applied at smaller scales (i.e. a Ramsar site or a restoration site).

3.1.3. Observed species richness

Indicator type

State indicator.

Explanation

Observed species richness represents the total number of unique species discovered and recorded. For maps, the value is calculated for each spatial cell. For time-series analyses, the value is calculated per time step over the complete data cube.

- High value: A high number of unique species, indicating a diverse local community.
- Low value: A low number of unique species, signalling a species-poor habitat.



Applicability for policy

This metric is necessary for determining biodiversity hotspots on a map or tracing potential declines in species richness over time. With taxonomic filtering, one can quantify the number of species in a particular cohort, such as known alien species, tree species, or carnivores.

Limitations

Observed species richness is highly sensitive to sampling bias and human observation effort. If certain grid cells are studied more carefully or visited more frequently by volunteers than others, they will artificially appear to have more species. Therefore, this metric should be interpreted within the context of a measure of survey effort, such as total occurrences or occurrence density. Furthermore, results are skewed by species detectability. Rare or widely dispersed species are easily missed by citizens, whereas highly conspicuous or easy-to-find species will often be reported (Hillebrand et al., 2018).

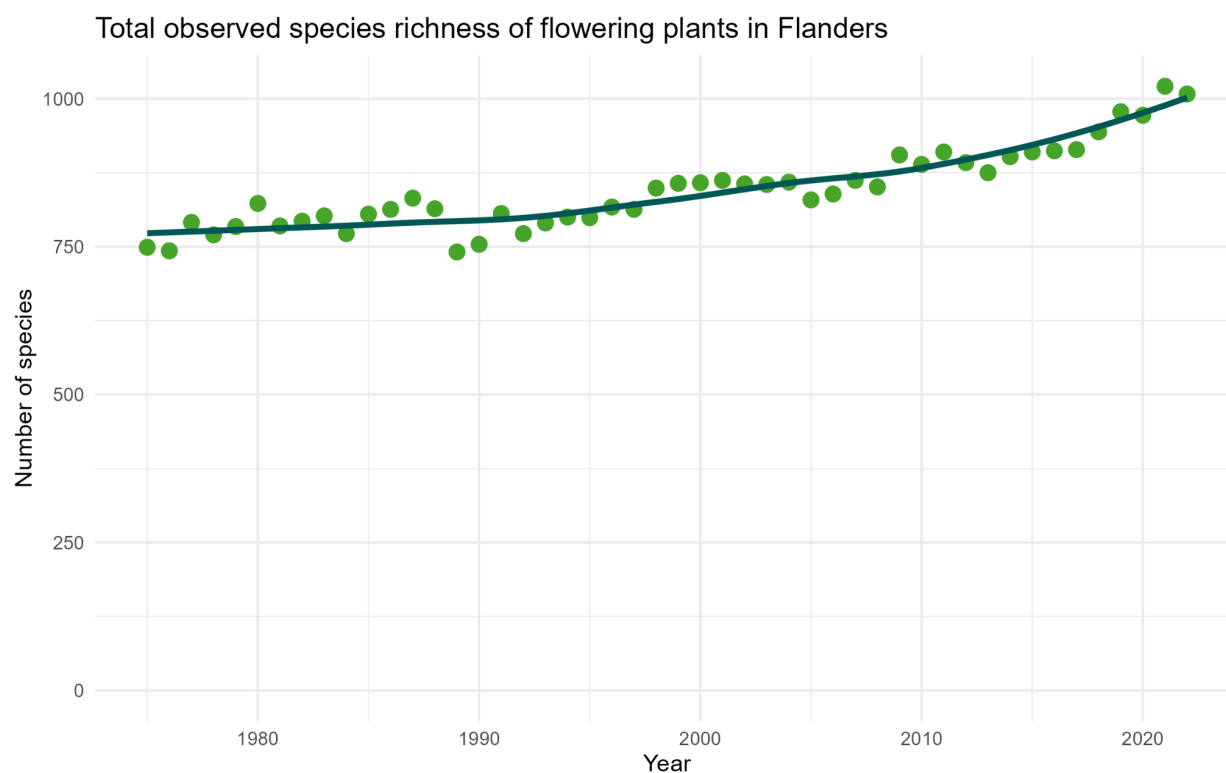


Figure 23: Total observed species richness of native flowering plants in Flanders reported since 1975, including a LOESS smoother. Data source: [GBIF.org](https://www.gbif.org) (2026e).

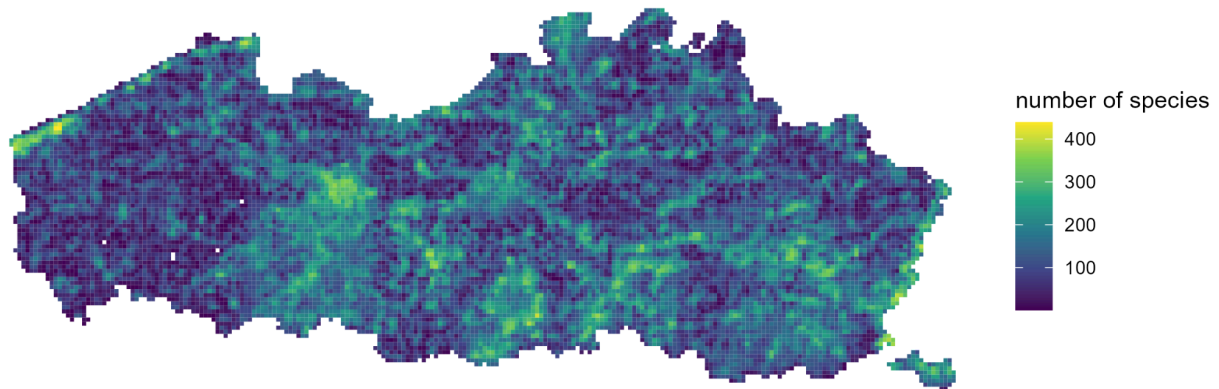


Figure 24: Spatial distribution of observed species richness of native flowering plants in Flanders reported since 1975. Data source: [GBIF.org](https://gbif.org) (2026e).

Interpretation

Species richness is highest in the Belgian dunes and other protected nature reserves, as well as in the area around Ghent and at the periphery of Brussels. The lowest species richness is observed in central West Flanders, a region characterised by intensive agriculture.

When examining the temporal trend in the number of angiosperm species, a gradual increase is observed over time. This pattern is somewhat unexpected, as regional species richness is generally assumed to be relatively stable over time. A likely explanation is incomplete and uneven sampling during the period 1975–2010. Data collected for Florabank during this period were not sampled homogeneously across space and time; for example, not all ecoregions were included each year. As a result, some species were likely missed in individual years (Van Landuyt et al., 2006).

In the last 11 years, volunteers have been encouraged to sample more systematically across Flanders, resulting in more homogeneous spatial coverage and a higher number of species recorded per year. This improvement is facilitated by the introduction of maps highlighting undersurveyed locations, which guide volunteer effort, as well as the establishment of a structured flora monitoring network in Flanders (Van Landuyt & Westra, 2015).

3.1.4. Newness

Indicator type

Data quality indicator.

Explanation



This indicator measures the age distribution of records within a dataset. For spatial maps, it calculates the average (mean) year of all species occurrence records within a specific grid cell, indicating how recent or historical the data in that cell is. For time-series analyses, the value is calculated as a running historical average. It calculates a cumulative mean at each time step based on all records collected from the beginning of the dataset until that time point. This tracks how the 'center of mass' of the dataset changes over time as new records are added, revealing whether historical or newer records are dominant.

- High value (recent mean year): The majority of the data in the grid cell was collected recently. This is often driven by citizen science data over the last few years.
- Low value (older mean year): This indicates that a large portion of the species occurrence records consists of older museum collections or long-running monitoring programs.

Applicability for policy

- Monitoring survey history: The indicator helps visualise how survey effort has evolved across space and time, highlighting areas that have been surveyed recently as well as regions with limited recent monitoring.
- Data recency assessment: Grid cells with lower mean observation years rely more heavily on older records and may require re-surveying to ensure that management decisions are informed by up-to-date biodiversity data.

Limitations

A recent mean year is not necessarily an indication of quality, as some countries or regions have been conducting comprehensive biodiversity monitoring for many years and will therefore reflect an older mean year of occurrence, while others may show a recent mean year due to, e.g., the sudden availability of large amounts of citizen science data.

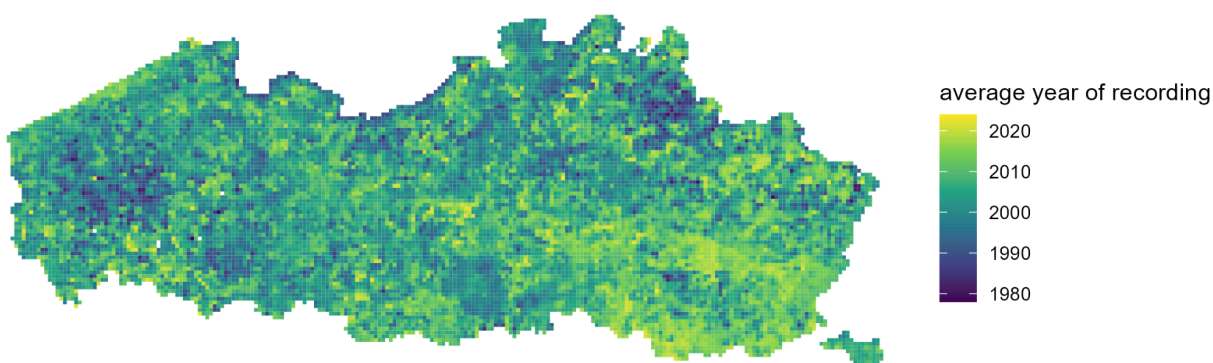


Figure 25: This figure shows the spatial distribution of the average year of recording for native flowering plants in Flanders, based on records collected since 1975. Data source: [GBIF.org](https://www.gbif.org) (2026e).

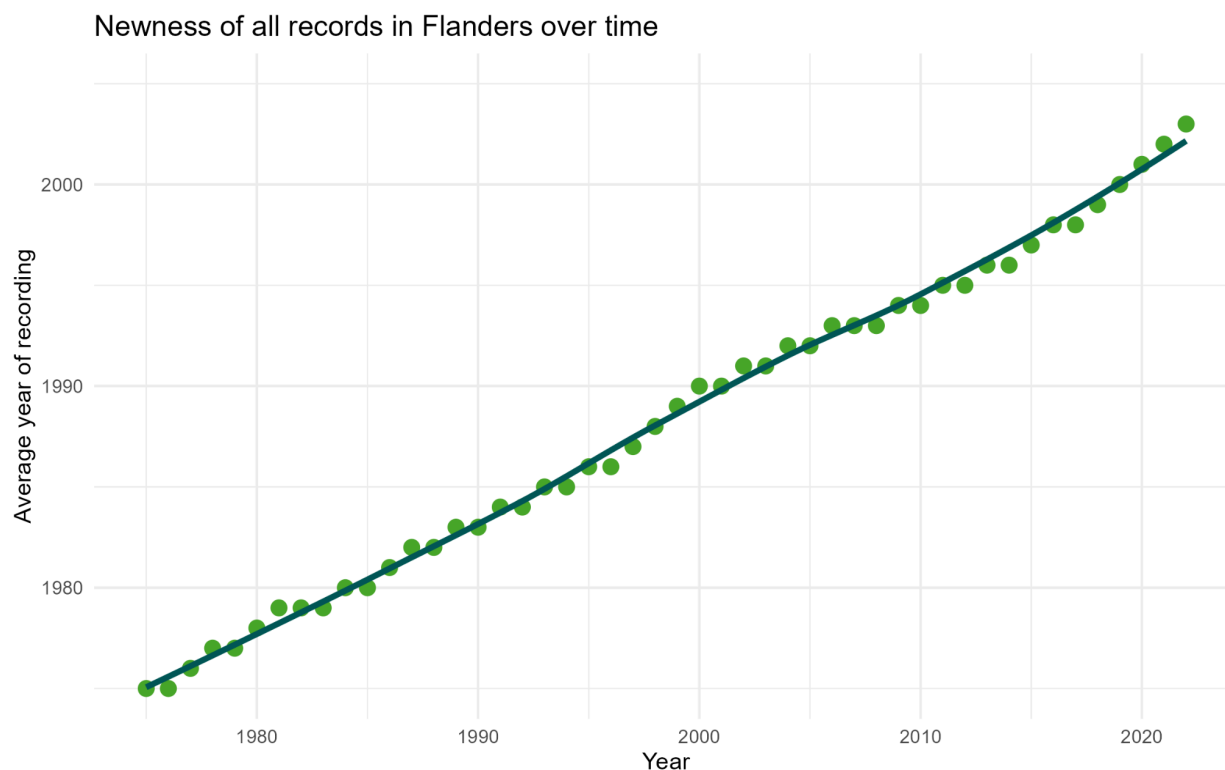


Figure 26: This figure shows the annual average recording year for all available records of native flowering plants in Flanders since 1975, including a LOESS smoother. Data source: [GBIF.org](https://gbif.org) (2026e).

Interpretation

There is a large variation in newness across Flanders. The area in central West Flanders has a lower newness than Zwin national park, Hoge Kempen or Hart van Haspengouw. Logically, the average year of recording is steadily increasing over time.

3.1.5. Cumulative species richness

Indicator type

Data quality indicator / pressure

Explanation

Cumulative richness is calculated by adding only newly observed unique species each year to a cumulative sum. For maps, the value is calculated as a snapshot for each spatial cell. For time-series analyses, the value is calculated as a running cumulative sum for each sequential time step over the complete data cube.



- Steep increase of the cumulative curve: Each timestep, new species are still being discovered for the first time in that region.
- Shallow increase or flat cumulative curve: Almost no new species are being added to the historical list anymore.

Applicability for policy

- Data completeness check: If the cumulative richness line is steep, it indicates that the data is still incomplete and not all species are represented in the dataset.
- Tracking new arrivals: In highly monitored regions where data is already complete, a sudden increase in cumulative richness can serve as a leading indicator for the arrival of invasive alien species or climate-driven species shifts.

Limitations

This metric is poorly suited for measuring rapid, negative biodiversity changes or current habitat degradation. Since it is a cumulative sum, the value never decreases. If a local extinction occurs or if common species crash in numbers, cumulative richness will ignore it and remain at its maximum. It may also be influenced by taxonomic changes, which can increase the number of recognised species without any additional species actually establishing in the area. In addition, gradual increases in cumulative richness may reflect increased survey effort rather than new species arrivals.

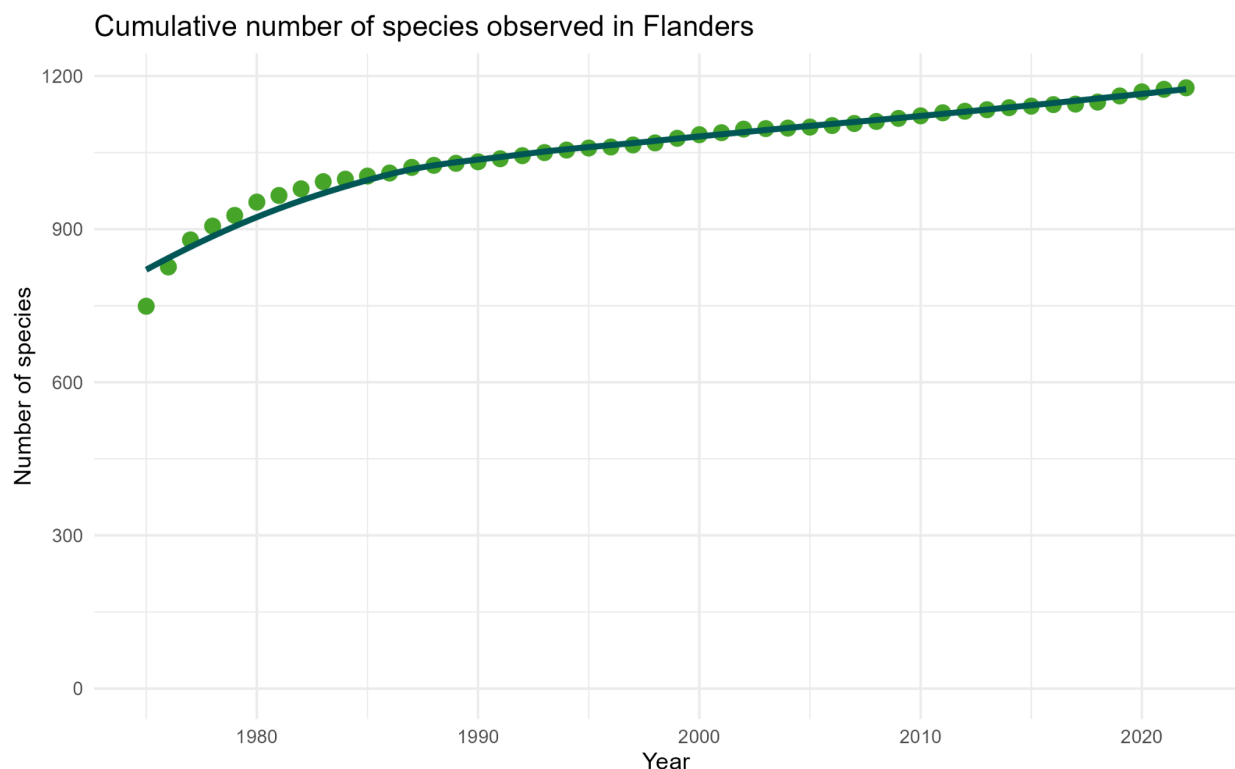




Figure 27: This figure shows the temporal trend in cumulative species richness of native flowering plants in Flanders since 1975, including a LOESS smoother. Data source: [GBIF.org](https://gbif.org) (2026e).

Interpretation

Although the number of native angiosperms is considered to have been stable across time, the cumulative number of species is steadily increasing. This is to be expected when you have a completeness below 1 and increasing sampling effort. To a minor degree this might result from changes in taxonomic backbone resulting in more total species and rediscoveries of species (e.g. *Rubus*, *Taraxacum*, *Hieracium*).

3.1.6. Relative occupancy of well surveyed cells based on Florabank.

Indicator type

State/ Trend indicator

Explanation

First, records are retained for a given year only if a minimum number of species (100) has been recorded within a grid cell during that year. Otherwise, records from that cell are removed from the data cube. Subsequently, relative occupancy per year is calculated as the proportion of well-surveyed grid cells in which a species has been observed.

Applicability for policy

By considering only well-surveyed cells and calculating the proportion of occupied cells per year, this indicator reduces the influence of changes in survey effort on observed trends. It is therefore recommended when assessing biodiversity trends while aiming to minimize the impact of observation bias.

Limitations

Although this indicator reduces the influence of variation in survey effort, it does not fully remove observer bias. The identification of well-surveyed cells depends on the selected threshold for the minimum number of recorded species, which may influence the resulting trends. In addition, changes in survey practices, observer expertise, or spatial coverage within well-surveyed cells may still affect occupancy estimates. The indicator also excludes poorly surveyed areas, potentially reducing spatial representativeness and limiting its applicability for species with restricted distributions or low detectability.

Examples

The relative occupancy within well surveyed cells was calculated as a test for three species showing clear increasing or decreasing trends over time (personal communication): *Centaurea*



cyanus, *Arnoseris minima*, and *Ophrys apifera*. The predicted trends were largely confirmed by the relative occupancy estimates (Figures 28-30).

Centaurea cyanus has experienced declines associated with agricultural intensification, particularly increased herbicide use. However, the species has shown signs of recovery in some regions, potentially linked to its frequent inclusion in wildflower seed mixtures and agri-environmental measures (Storkey et al., 2012). *Arnoseris minima* has also declined, likely reflecting the broader decline of arable plants associated with agricultural intensification and increased agrochemical use (Storkey et al., 2012). Finally, *Ophrys apifera* shows increasing occupancy, consistent with the observed northward range expansion of the species, which has been linked to climate change (Naczek et al., 2025).

These examples further highlight the need to combine multiple measures when interpreting biodiversity trends. For example, in the case of *Ophrys apifera*, the increase in absolute occupancy coincides with an increase in overall recording effort and could therefore partly reflect sampling artefacts. However, the increasing trend in relative occupancy provides additional support that this pattern reflects a genuine ecological trend.

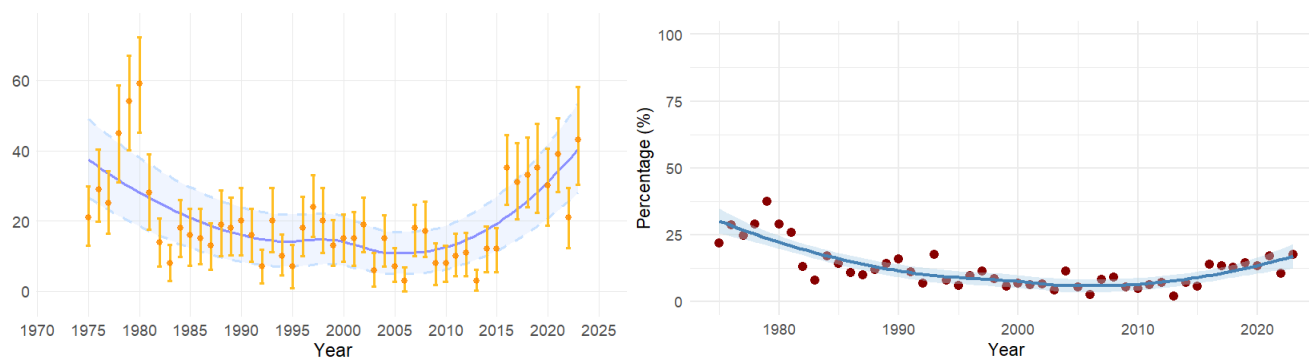


Figure 28: Temporal trends in absolute occupancy (left; number of occupied well-surveyed 5×5 km grid cells) and relative occupancy (right; percentage of occupied well-surveyed 5×5 km grid cells) for *Centaurea cyanus*. The line represents a LOESS fit. Data source: Florabank via [GBIF.org](https://gbif.org) (2026f).

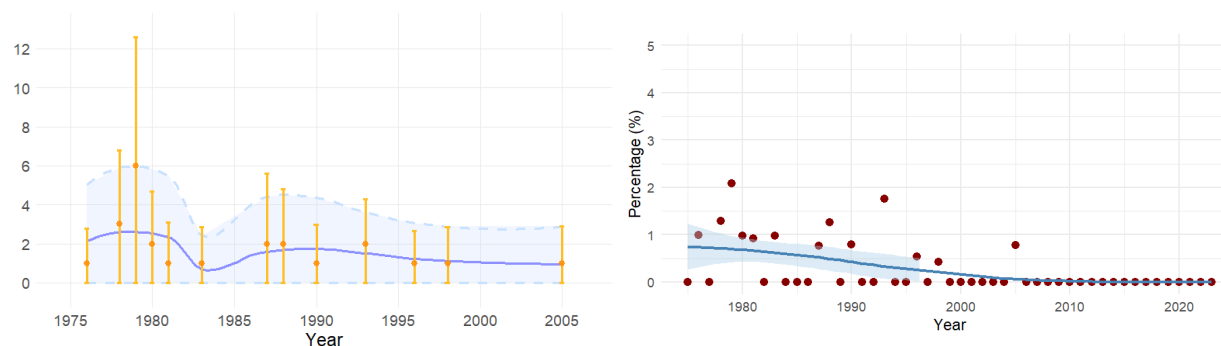


Figure 29: Temporal trends in absolute occupancy (left; number of occupied well-surveyed 5×5 km grid cells) and relative occupancy (right; percentage of occupied well-surveyed 5×5 km grid cells) for *Arnosseris minima*. The line represents a LOESS fit. Data source: Florabank via [GBIF.org](https://gbif.org) (2026f).

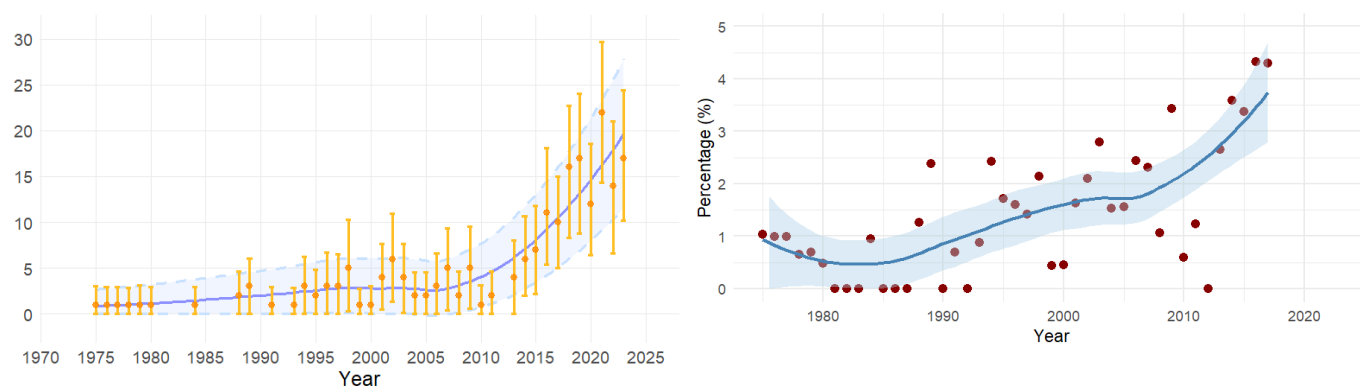


Figure 30: Temporal trends in absolute occupancy (left; number of occupied well-surveyed 5×5 km grid cells) and relative occupancy (right; percentage of occupied well-surveyed 5×5 km grid cells) for *Ophrys apifera*. The line represents a LOESS fit.



Data source: Florabank via [GBIF.org](https://gbif.org) (2026f).

3.2. Phylogenetic indicators

The indicators displayed below are based on an occurrence cube of only native flowering plants, observed between 1975 up to now at either 1 km or 5 km resolution.

3.2.1. Phylogenetic Diversity (PD)

Indicator type

State indicator

Explanation

This indicator takes into account the evolutionary relatedness of species by calculating the sum of the branch lengths (representing the accumulated genetic changes) of the evolutionary tree of life that connect all species found in a specific area (Faith, 1992). Instead of just counting the number of different species (species richness), it calculates the minimum path length needed to span across those species on a phylogenetic tree.

- High value: The area contains species that are far apart on the tree of life (for example, it includes species of ancient lineages that have evolved separately from most other modern species). This means a wide variety of evolutionary features and distinct genetic traits are probably present in that ecosystem.
- Low value: The species in the area are very closely related and clumped together on the same or a few related branches of the tree of life, meaning the ecosystem has a comparatively low variety of distinct evolutionary features.

Applicability for policy

This metric is highly valuable for long-term conservation planning and climate change adaptation policies. Because we can only guess which biological traits and features will be crucial for species survival in changing future environments, protecting a wide variety of features (rather than just the number of species) maximizes the overall adaptability and resilience of the ecosystem (Robuchon et al., 2023).

Limitations

The indicator is highly dependent on having an accurate and complete phylogenetic tree with reliable branch lengths for the taxonomic group being analyzed. If the underlying evolutionary relationships are poorly mapped, the calculated tree-spanning path can become skewed,



potentially over- or underestimating the true evolutionary value of the area.

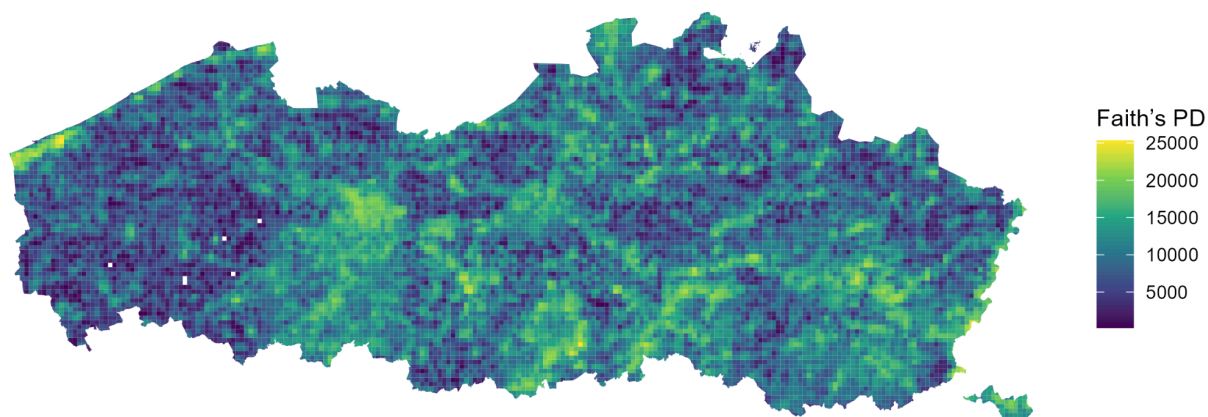


Figure 31: This map shows the spatial distribution of phylogenetic diversity among native flowering plants in Flanders, based on records collected since 1975. Data source: [GBIF.org](https://gbif.org) (2026e).

Interpretation

Areas with high phylogenetic diversity generally correspond well with areas showing high observed species richness (Figure 24). The highest values for this metric are found in the Belgian dunes, other protected nature reserves, and areas surrounding Ghent and the outskirts of Brussels. In contrast, the lowest species richness is observed in central West Flanders, a region largely characterised by intensive agricultural land use. A more detailed comparison is needed to determine whether and where differences occur between patterns of phylogenetic diversity and observed species richness across Flanders.

3.2.2. Expected Phylogenetic Diversity (PD) Loss Map

Indicator type

Pressure/ecosystem indicator.

Explanation

This indicator calculates exactly how much evolutionary history we stand to lose if threatened species disappear from a specific grid cell. Species are considered 'threatened' when classified as Critically Endangered (CR), Endangered (EN) or Vulnerable (VU) according to the local IUCN Red List. The total Phylogenetic Diversity is then once calculated *with* all species present; and second, *without* the threatened species. The difference between these two scores equals the Expected PD Loss (Faith, 2015; Gumbs et al., 2023).



- High value (High Expected Loss): If the threatened species in this grid cell go extinct, a substantial chunk of the evolutionary tree of life and the feature diversity associated with it for this region will vanish. This happens because these endangered species have no close relatives in these cells that share their evolutionary history and evolved traits.
- Low value (Low Expected Loss): The threatened species in this cell are either close cousins to other common species (meaning their disappearance won't completely delete that branch of the tree of life) or there are few threatened species present in the area.

Applicability for policy

This indicator provides a "triage map" for policymakers and conservation agencies that highlights the areas where a large amount of phylogenetic diversity is represented by threatened species. Rather than just looking at the decline in species richness alone, it also considers the variety of features across the evolutionary tree of life. Resources can then be allocated towards high-priority areas, for example, to expand protected areas to include hotspots of evolutionary heritage (Robuchon et al., 2023).

Limitations

This indicator treats all "threatened" categories (CR, EN, VU) with equal weight when calculating the expected loss, and therefore, it does not differentiate between a species that is on the brink of extinction versus one that is vulnerable. Furthermore, it relies on the up-to-date accuracy of regional Red Lists, and therefore, the potential loss of a species that has become threatened but is not officially assessed yet will not be factored in.

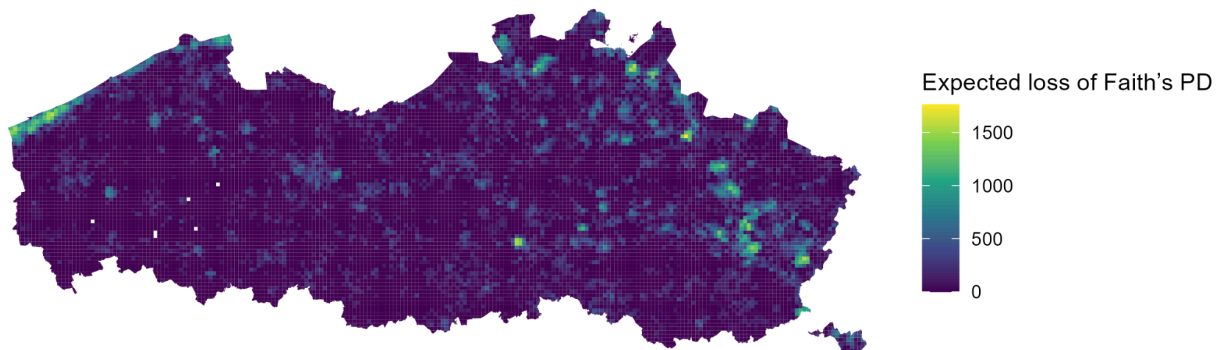


Figure 32: This map represents the 'Expected Phylogenetic Diversity (PD) Loss Map' of native flowering plants in Flanders, based on records collected since 1975. Data source: [GBIF.org](https://gbif.org) (2026e).

Interpretation



The Belgian dunes and areas in Limburg (e.g. Hoge Kempen National Park) show a higher potential loss of phylogenetic diversity among threatened species compared to other regions in Flanders.

3.3. Invasive alien species indicators

3.3.1. Alien Species Portal

The Alien Species Portal has been developed by the Institute for Nature and Forest Research (INBO). The portal showcases occurrence, distribution and management data (if available) for all alien species included in the Belgium GRIIS checklist (Desmet et al., 2025).

The platform contributes to the monitoring and reporting of invasive alien species in accordance with the [European legislation \(No 1143/2014\)](#). As such, it allows investigation of invasion dynamics, and eventually supports research-based management of IAS in Belgium, including Flanders.

The platform was initially developed with a budget from “OmgevingsInformatie Samenwerkingsverband” (OIS) and INBO (budget IAS), but because its timeline aligned with the B3 project, all software workflows were developed based on either occurrence cubes generated by **rgbif** (applying the B3 workflow) or the Belgium GRIIS checklist (Desmet et al., 2025). The platform is also fully automated with daily or monthly updates of indicators based on GBIF data.

Below is an overview of all indicators that are available via the Alien Species Portal (<https://alienspecies.inbo.be/>) among which there are general indicators, the emerging status per species, checklist-based indicators and the Target 6 indicator. Currently, almost all indicators are incorporated in the Alien Species Portal except for Target 6. At the moment, the B3 budget is used to finalize the dashboard and to add the Target 6 indicator to the platform.

3.3.2 General information per species

In the section ‘species information’, the user can select a species name and the following indicators are visualised:

- Map with occupied grid cells (occupancy) (Figure 33)
- Annual number of occupied grid cells
- Invasion history (Figure 34)
- Management map (Figure 35)
- Management timeline
- Emerging status (see 3.3.3)

The occupancy map, invasion history and emerging status of a species are based on GBIF data cubes, generated by **rgbif** (Table 5). Besides occurrence data, there is also management data for some species (e.g. *Xenopus laevis* or *Ondatra zibethicus*). The management data is either internally hosted by INBO or also published on GBIF in specific datasets (e.g. Van Moer et al., 2026).

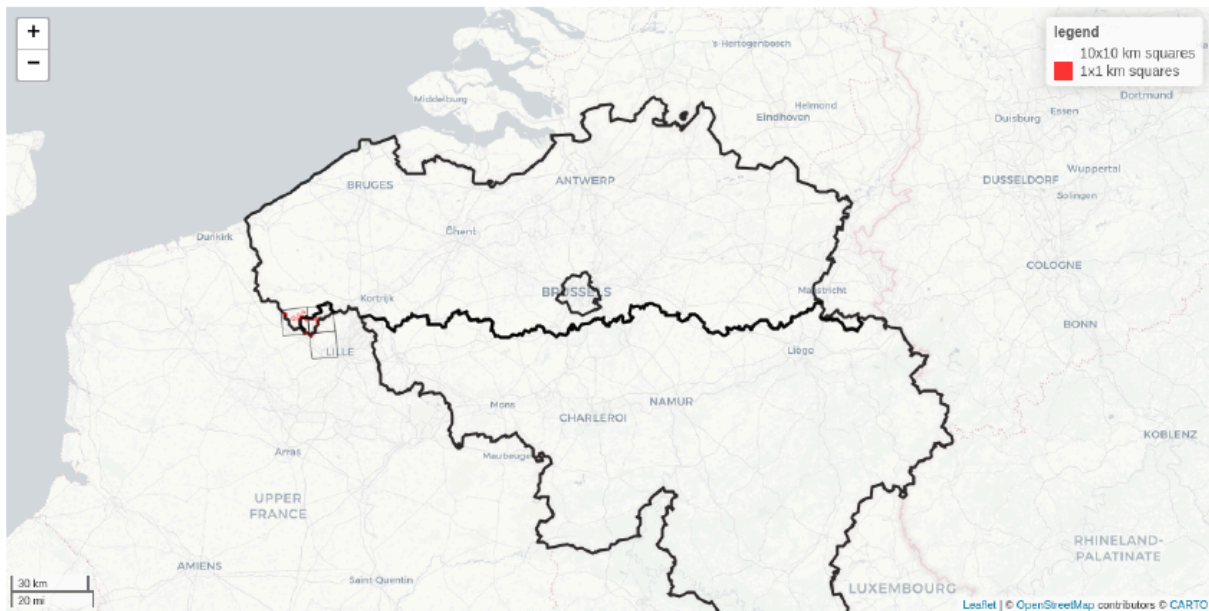


Figure 33: Map showing the occupancy of *Xenopus Laevis* from 2006 to 2025 in Belgium. Black squares represent 10 km x 10 km areas where the species has been recorded. Red squares highlight more precise 1 km x 1 km locations within these areas where sightings have been reported. (Source: alienspecies.inbo.be)

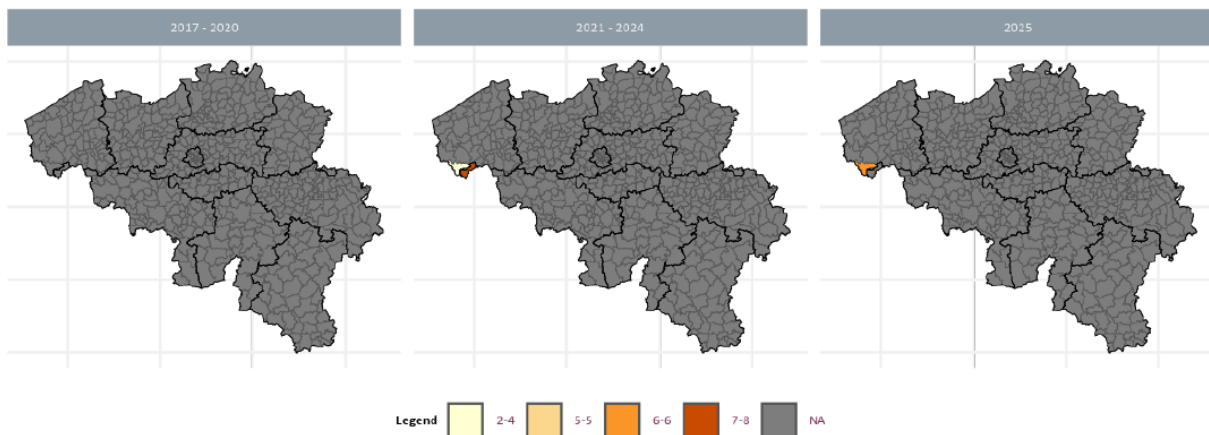


Figure 34: This figure illustrates the invasion history of *Xenopus Laevis*. The color code represents the number of reported occurrences per municipality per time period of *Xenopus Laevis*. (Source: alienspecies.inbo.be)

+

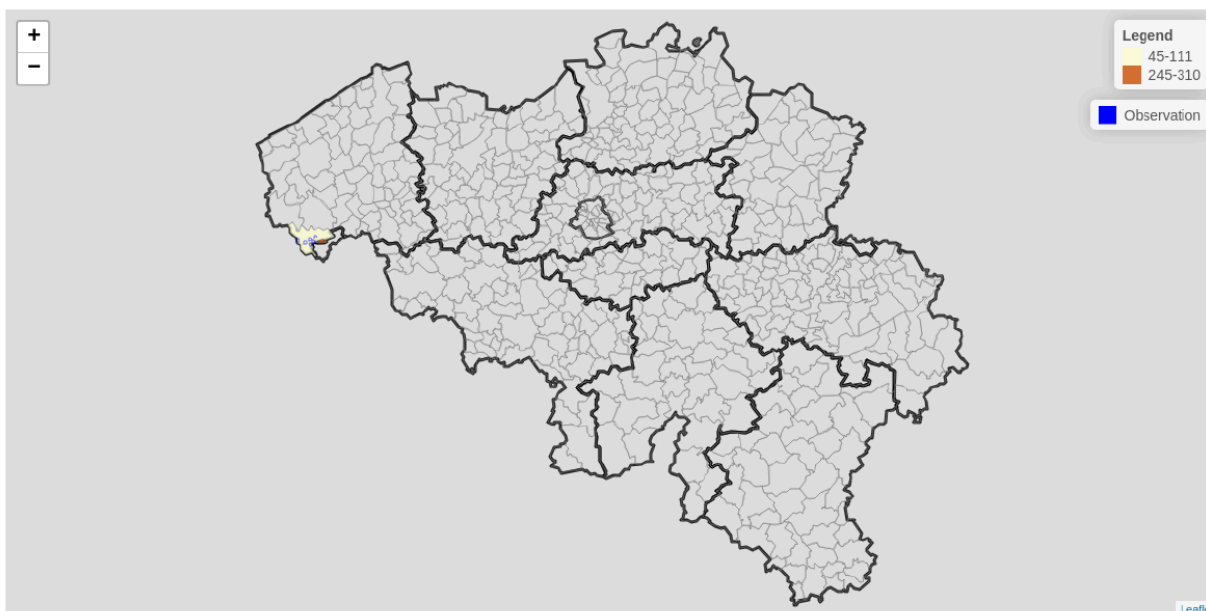


Figure 35: The absolute number of managed individuals of *Xenopus Laevis* in Flanders.
(Source: alienspecies.inbo.be)

3.3.3. Emerging Status of Invasive Alien Species

Indicator type

Pressure indicator / monitoring management

Explanation

The emergence status of species is statistically assessed by fitting Generalised Additive Models (GAMs) to a timeseries of either the number of observations or the number of occupied grid cells. By analysing the sign of the first and second derivatives (the slope and acceleration) of the model curve over the previous three years, species are classified into specific categories:

- Emerging: Both the first and second derivatives are positive, indicating a statistically significant and accelerating increase in observations or occupancy.
- Potentially emerging: The first derivative is positive, but the acceleration is not strictly positive.
- Unclear: Both derivatives are equal to zero, showing a flat or ambiguous trend.
- Not emerging: The first derivative is negative or the first derivative is equal to zero while the second derivative is negative. *(Note: If a lack of data prevents the model from converging, a standardised decision tree is applied to assign a status. To prevent data lags from skewing the results, the current calendar year is excluded from the calculation.)*

Applicability for policy



This indicator is a key operational tool for rapid response and priority management under the [EU Regulation \(No 1143/2014\)](#).

Limitations

The indicator is highly sensitive to data availability on GBIF. For species with very few records, the GAM may fail to converge. Then, the classification is shifted to a simplified decision tree. Furthermore, because the indicator intentionally omits data from the current year to ensure completeness, it functions as a “recent historical assessment” (the past three complete years) rather than an instantaneous real-time alert system. Furthermore, it is possible to correct this indicator for observer bias. However, this should be done with caution, as it requires the consistent availability of records published to GBIF at the class level.

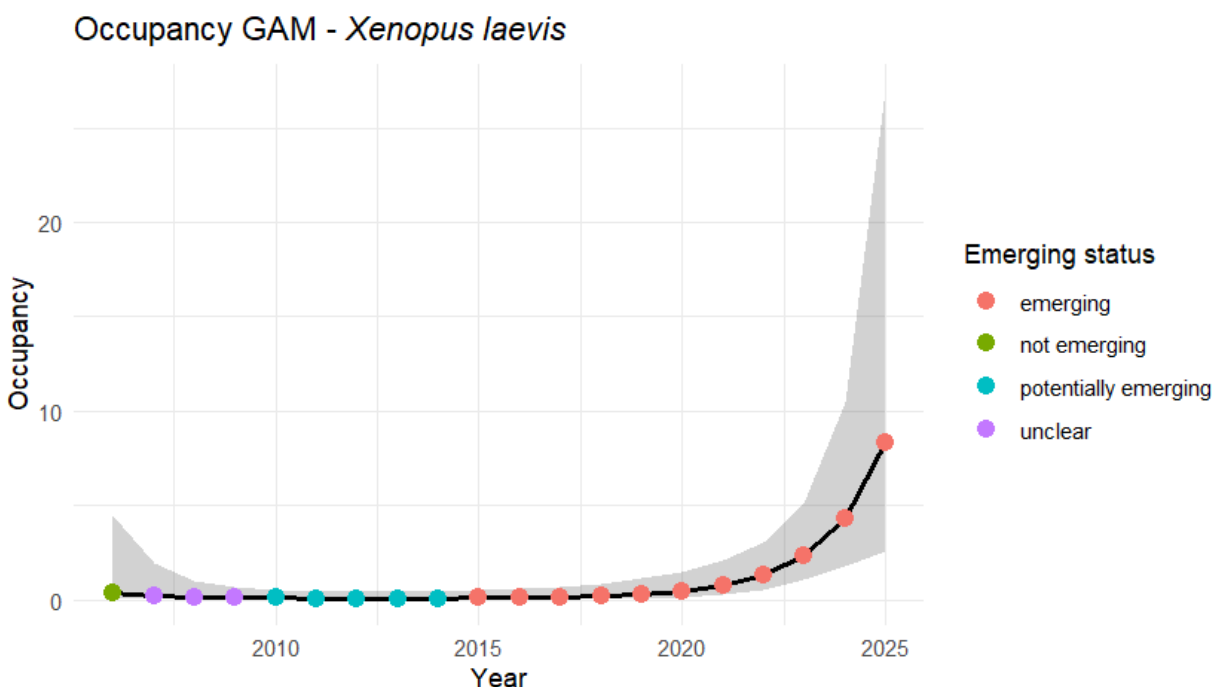


Figure 36. Overview of the annual emerging status of *Xenopus laevis* since 2005 based on Generalised Additive Models (GAMs) following the TrIAS methodology. (Data source: alienspecies.inbo.be)

3.3.4. Checklist-based indicators

Besides indicators at the species level, some indicators are calculated based on the GRIIS checklist (Desmet et al., 2025). These are based on [the TrIAS workflow](#) (Groom et al., 2022) and are the following:

- Number of new introductions of alien species per year (largely resembling the Target 6 indicator, see 3.3.5).
- Cumulative number of introduced species present in Belgium as function of time (Figure 37).



- Pathways associated with alien species introductions.
- Trend in occupancy between baseline and reporting period (also based on GBIF data cubes) (Figure 38).
- The number of introduced taxa per introduction pathway (Figure 39).

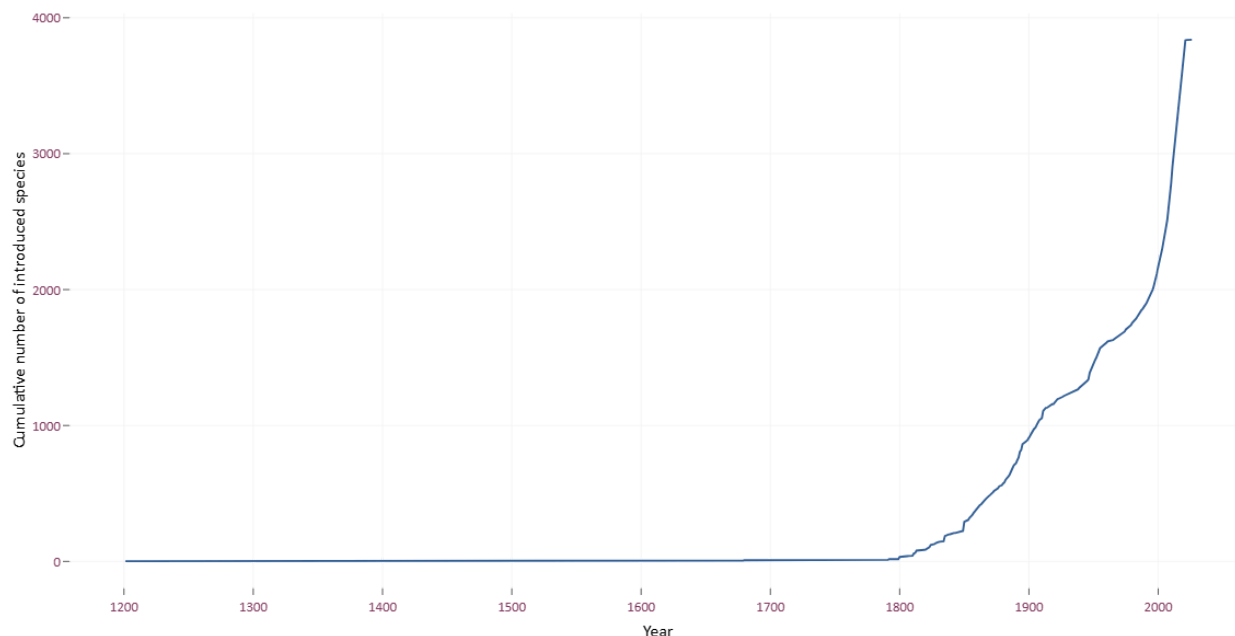


Figure 37: Cumulative number of introduced alien species in Belgium. (Source: alienspecies.inbo.be)

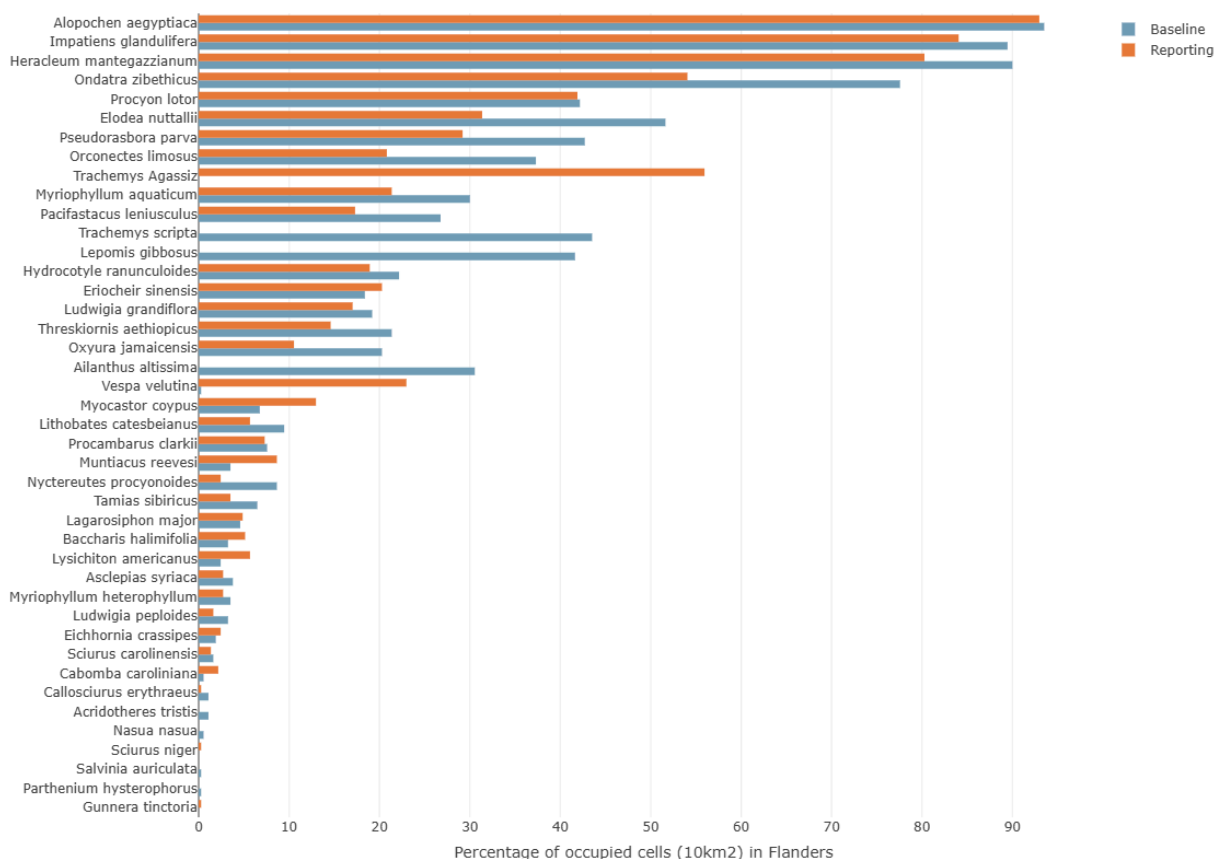


Figure 38: This graph displays the trend in occupancy (i.e., the percentage of occupied 10 km² cells) in Flanders for all species that are included in the list of invasive alien species of Union concern ([Regulation \(EU\) No 1143/2014](#)). For each species, the occupancy at the time of its inclusion on the list is indicated in blue (baseline), while the occupancy during the most recent three-year reporting period is represented in orange (reporting). (Source: [alienspecies.inbo.be](#))

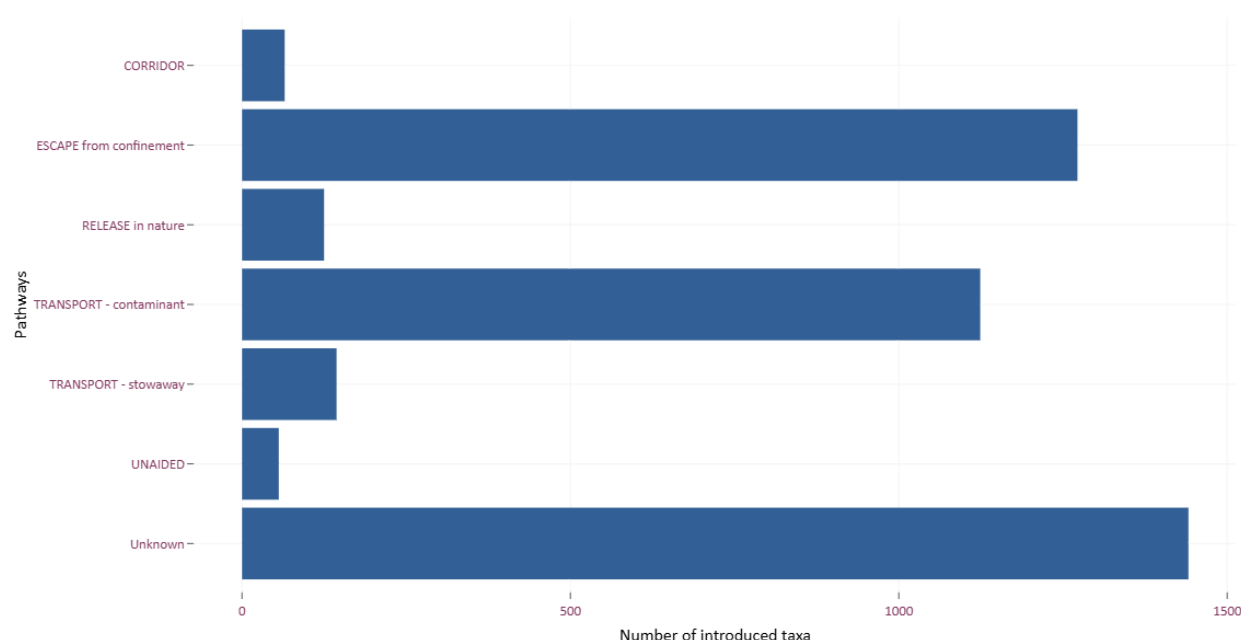


Figure 39: This graph shows the number of introduced taxa per introduction pathway category (CBD, level1 (IUCN, 2017)). (Source: alienspecies.inbo.be)

3.3.5. Target 6 indicator: rate of invasive alien species establishment

Indicator type

Pressure Indicator / monitoring management

Explanation

This indicator quantifies the number of new invasive alien species (IAS) that successfully establish within a country or region per year. The package **b3alien** screens occurrence records against international alien checklists to infer the establishment rate of IAS species (Trekels & Groom, 2026). Furthermore, the package corrects for detection delays by applying the approach proposed by Solow and Costello (2004), resulting in a more reliable basis for estimating IAS establishment rates.

- An increasing rate indicates that more IAS species are introduced and established in the region.
- A declining rate suggests that fewer IAS are being introduced, possibly due to effective prevention efforts targeting key introduction pathways.

Applicability for policy

This metric serves as the official CBD (Convention on Biological Diversity) Headline Indicator 6.1, “Rate of invasive alien species establishment”, to monitor global and national progress toward Target 6 of the Kunming-Montreal Global Biodiversity Framework. This target aims to



reduce the introduction and establishment rates of known invasive species by at least 50% by 2030. Examining this indicator over longer time periods, as recommended in the CBD guidance, can provide valuable insights into the effectiveness of policies and management interventions.

Limitations

This metric is highly sensitive to changes in survey effort over space and time, and differences in detectability among taxonomic groups. Consequently, this indicator should always be assessed alongside local distribution maps and expert ecological knowledge. Measures of observation effort and spatial bias are available through the **b3alien** package and should be considered when interpreting the results.

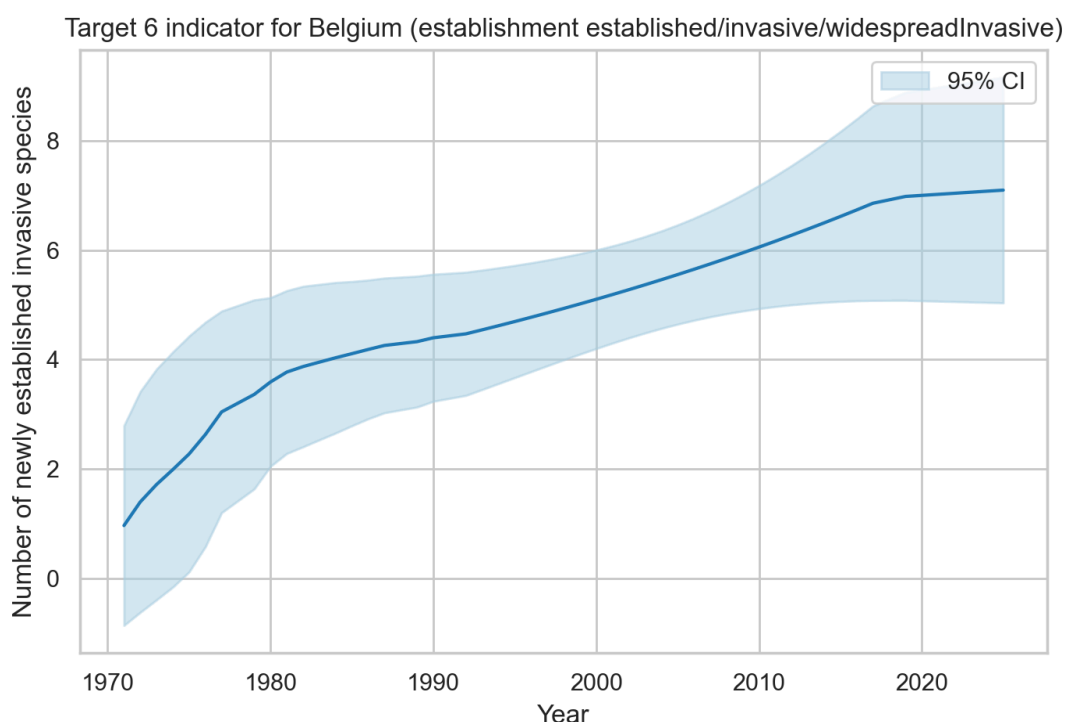


Figure 40: Number of newly established invasive species per year in Belgium since 1970, calculated using the b3alien package to correct for detection lags. This indicator is based on the Belgian checklist of the Global Register of Introduced and Invasive Species (GRIIS) (Desmet et al., 2025) and includes only species classified as established, invasive, or widespread invasive.

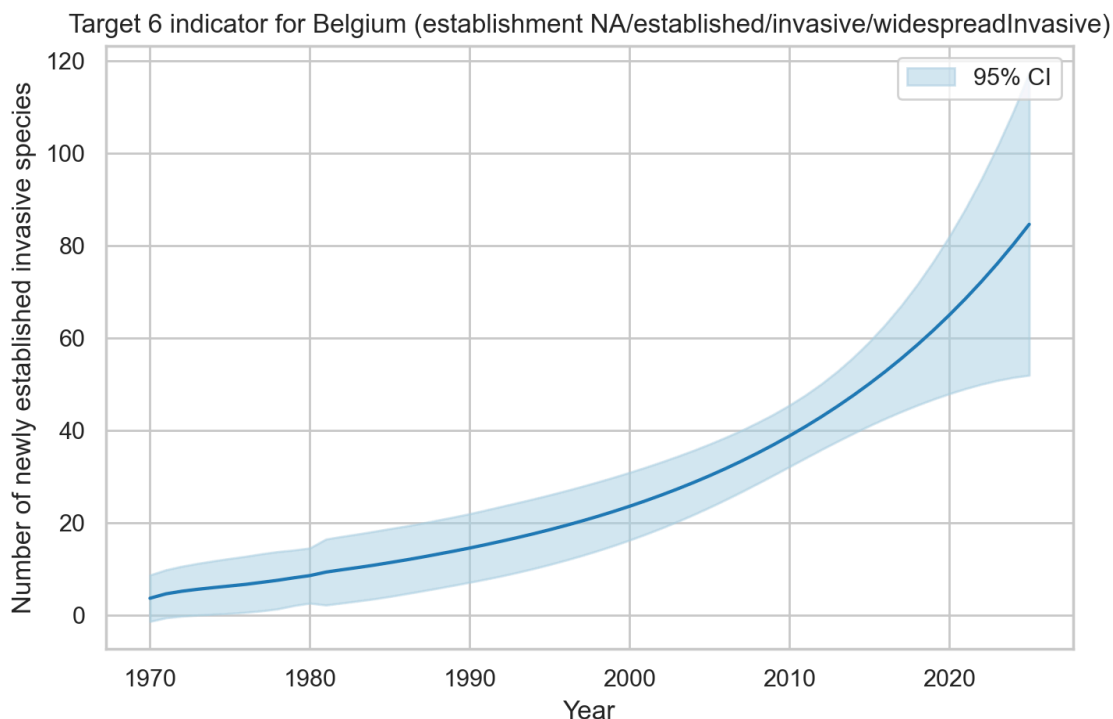


Figure 41. Number of newly established invasive species per year in Belgium since 1970, calculated using the *b3alien* package to correct for detection delays. This indicator is based on the Global Register of Introduced and Invasive Species (GRIIS) checklist for Belgium (Desmet et al., 2025), including a filter for species classified as established, invasive, or widespread invasive, or with an undefined establishment status.

3.3.6. Invasimapr application

Overall, the **invasimapr** workflow was successfully applied. Within Flanders, West Flanders exhibited the lowest trait diversity and, consequently, the highest estimated invasibility. However, this result should be interpreted with caution, as West Flanders is also the most intensively cultivated agricultural region in Flanders. The high degree of agricultural land use may therefore help explain the observed pattern. Currently, the TRY database (Kattge et al., 2020) was consulted as the main input source of the available traits as a pilot test of the workflow being developed in B3. However, in the future, this workflow should be rerun with a carefully selected trait database to offer more insightful results.

3.3.6.1. Data exploration with *dissmapr*

The **invasimapr** workflow (MacFadyen et al., 2025) was applied to a GBIF occurrence cube of vascular plants in Flanders, restricted to species for which functional-trait records were available in the TRY database and aggregated to a 0.1° grid (roughly 7 km cells). This produced a site × species × environment cube covering 3,173 trait-linked plant species. Before invasion fitness was modelled, the data were explored with **dissmapr** (MacFadyen & Hui, 2026) to characterise



the alien and invasive component of this flora, the intensity and spatial structure of sampling, and the observed species richness on which the resident-community model is built (Figures 42–45).

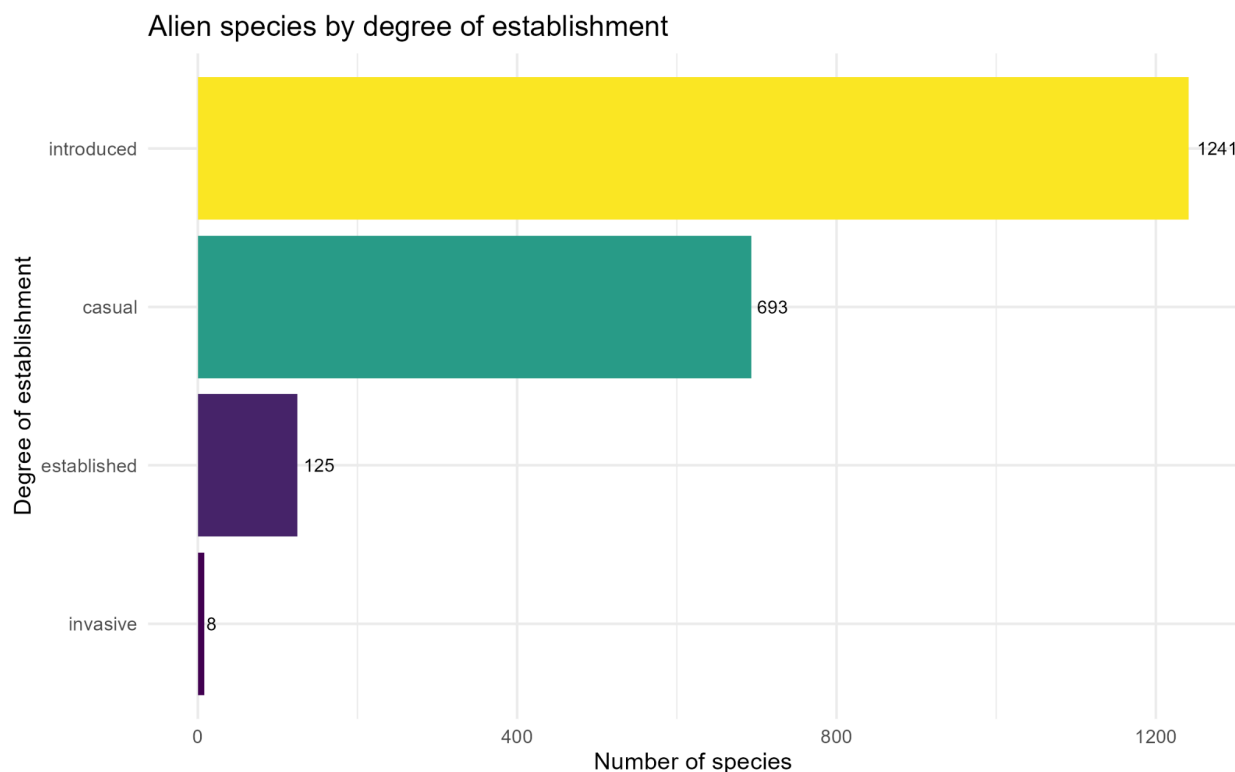


Figure 42: Degree of establishment of the alien plants recorded in the Flanders occurrence cube, based on the Manual of the Alien Plants of Belgium (alien-plants-belgium checklist, published via the Belgian Biodiversity Platform). Each species is assigned to the most advanced establishment stage reported for Flanders.

Of the alien species matched to the checklist, the large majority are classified as merely introduced (1,241) or casual (693), whereas far fewer have become established or naturalised (125) and only eight are explicitly flagged as invasive. This steep decline from introduced to invasive mirrors the classic introduction–naturalisation–invasion continuum, in which only a small subset of introduced taxa progress to persistent, spreading populations. For policy it provides a baseline of how far the alien flora currently sits along that continuum and highlights the comparatively small pool of species that have already crossed into the invasive stage.

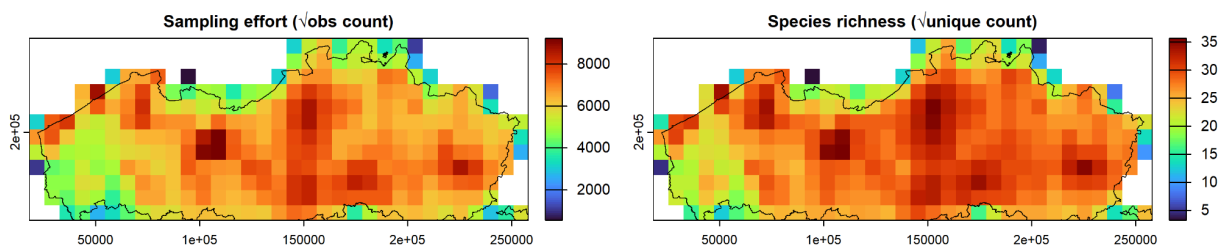


Figure 43: Sampling effort (number of occurrence records) and species richness (number of distinct trait-linked plant species) per grid cell in Flanders, computed with dissmapr. Both quantities are square-root transformed to dampen the influence of a few very heavily recorded cells.

Sampling effort and species richness show closely matching spatial patterns, both peaking in a broad central band of Flanders and thinning towards the periphery, except for the coastal region. Because richer cells are also the most intensively surveyed, part of the apparent richness gradient reflects observation effort rather than ecology alone. This is a well-known property of opportunistic GBIF data that the survey-effort and completeness indicators elsewhere in this deliverable are designed to address. This coupling is an important context for the downstream invasibility maps, which are built on the same underlying grid.

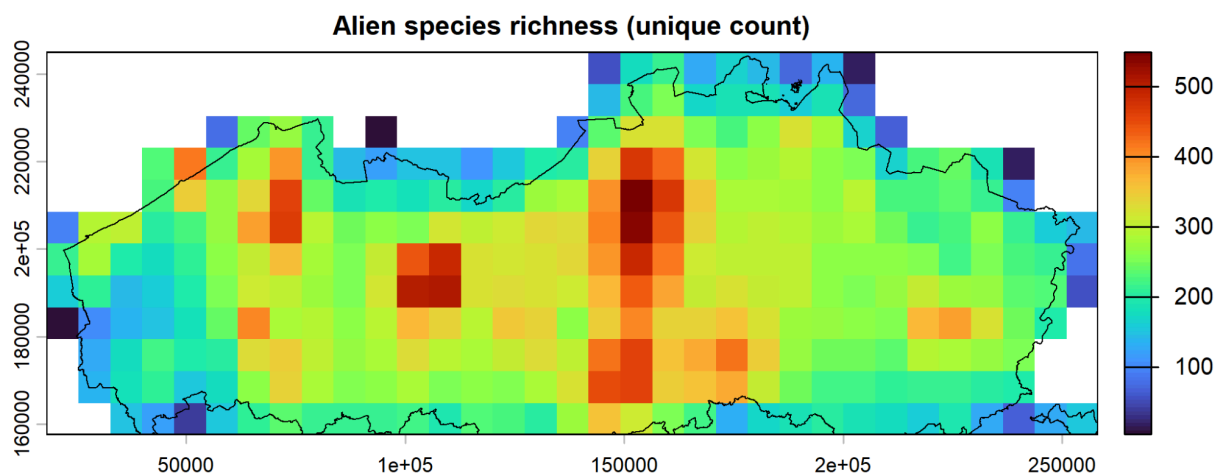


Figure 44: Spatial distribution of observed species richness or the number of distinct trait-linked plant species recorded per grid cell across Flanders over the last 11 years (2016 - 2026) of GBIF records.

Per-cell richness of the trait-linked flora ranges from a couple of hundred species in sparsely sampled peripheral cells to roughly 1,600 in the richest cell, with most of Flanders falling in an intermediate band of several hundred species. The pattern is comparatively even across the region, with isolated hotspots that again coincide with intensively recorded localities. This observed richness is the raw material for the resident-community model: each cell's set of



co-occurring species defines the resident assemblage whose trait structure and saturation later determine how open that cell is to newcomers.

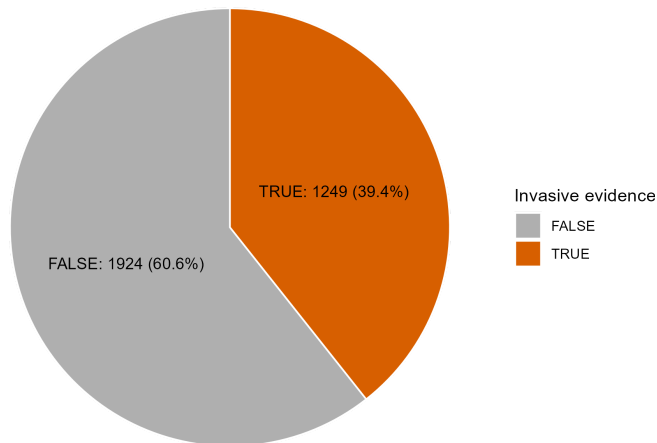


Figure 45: Proportion of the trait-linked species in the cube that are listed as invasive in Belgium on the BFIS/HARMONIA checklist (Harmonia database, 2026) or on the DVW (Flemish Waterways) priority list (Adriaens et al., 2023). Species with such evidence are treated as candidate invaders; the remainder form the resident pool.

About two-fifths of the trait-linked species (1,249 of 3,173; 39.4%) are listed as invasive in Belgium and are therefore taken forward as candidate invaders, while the remaining 60.6% constitute the resident community against which invasion is tested. This split is deliberately inclusive: because species are matched to the checklists by a tolerant genus-plus-epithet rule, the candidate-invader pool is broad and includes many taxa that are cultivated, casual or only occasionally recorded. It should therefore be read as 'species with any listed invasive association' rather than as a curated set of active invaders. We discuss this point further in the discussion, since it directly shapes the invasiveness results below.

3.3.6.2. Shared trait space

To quantify how similar potential invaders are to the species already present, all species were placed in a common functional trait space. Eleven traits with high measurement coverage ($\geq 70\%$ of species) were used: dispersal syndrome, seed dry mass, vegetative plant height, plant lifespan, woodiness, growth form, leaf type, leaf compoundness, photosynthetic pathway and nitrogen-fixation capacity. Species were ordinated with non-metric multidimensional scaling (NMDS) on a Gower distance, and residents and candidate invaders were then compared within the resulting space (Figure 46).

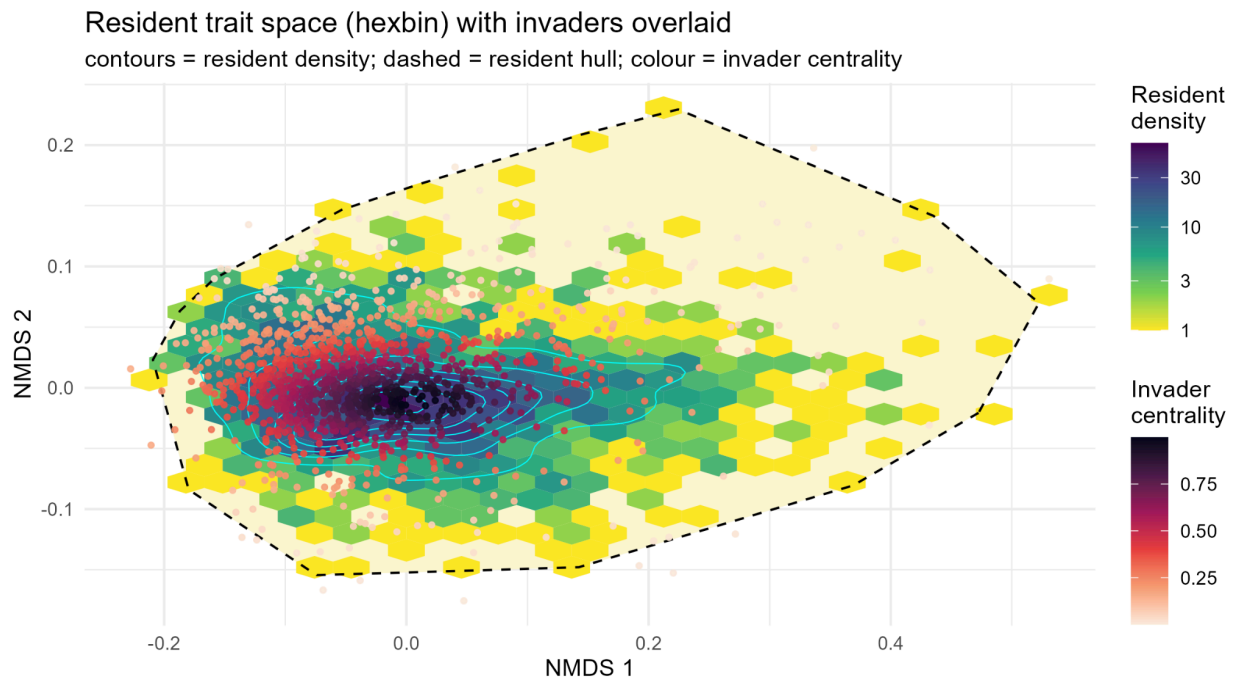


Figure 46: Shared functional trait space of residents and candidate invaders (NMDS ordination). Shading and cyan contours show the density of resident species; the dashed line is the convex hull of the resident trait space; points are candidate invaders, coloured by their centrality within the resident cloud (dark = deep inside the resident core).

The candidate invaders fall almost entirely within the convex hull of the resident community and cluster towards its dense core rather than towards the edges or outside it. In other words, the species most likely to arrive in Flanders are functionally very similar to those already established. This means they represent variations on existing strategies rather than genuinely novel trait combinations. Under the invasibility-cube framework this has a clear consequence: because invaders overlap strongly with residents in trait space, their fate is governed less by trait novelty than by how climatically suitable a site is and how saturated the resident community already is. It also means niche crowding is highest for the most central invaders, which face the stiffest competition from trait-similar residents.

3.3.6.3. Invasion fitness analysis

Invasion fitness (λ) is the modelled per-capita growth rate of a rare invader introduced into a site's resident community; a positive value means the species is expected to be able to increase and potentially establish. λ combines three components estimated from the fitted resident model namely: i) abiotic suitability, ii) trait-similar niche crowding and iii) resident saturation. These are computed for every candidate invader at every site. Aggregating λ across invaders gives a site-level picture of where Flanders is most open to invasion (Figures 47-48); aggregating establishment across sites gives species- and trait-level pictures of what makes a successful invader (Figures 49-50).

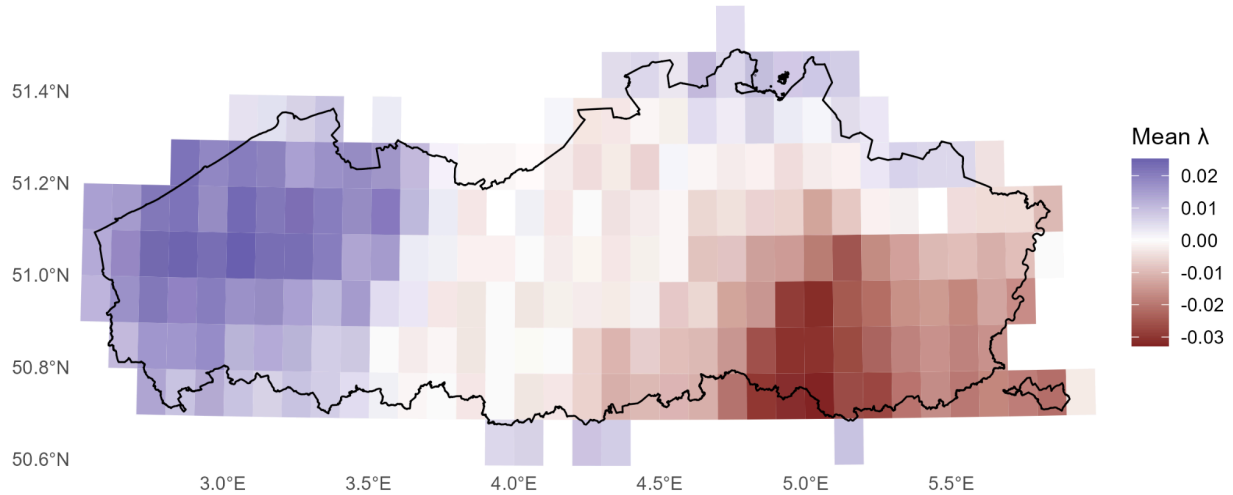


Figure 47: Mean invasion fitness (λ) per site, averaged across all candidate invaders. The diverging scale is centred on the establishment threshold $\lambda = 0$ (blue = net-suitable, red = net-unsuitable on average).

Mean invasion fitness is highest in the west of Flanders where the average invader has a positive growth rate, and declines eastwards, turning negative in the south-east. On average, then, the western tip of Flanders is the most permissive to newcomers and the south-east the least. This west–east gradient is reinforced by the explicit invasibility map shown below.

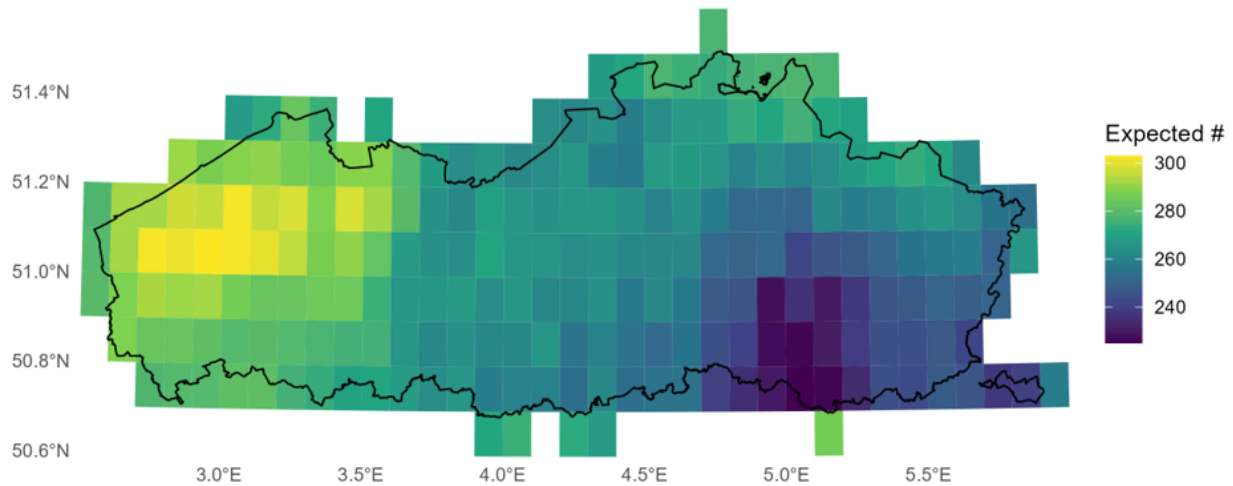


Figure 48: Invasibility map i.e. the expected number of candidate invaders able to establish ($\lambda > 0$) in each grid cell. Higher values (yellow) indicate sites open to a larger share of the invader pool.

Translating fitness into an establishment count sharpens the same pattern: the expected number of establishing invaders peaks at close to 300 species in West Flanders and falls to around 230 in the south-east (roughly the Hageland–Limburg area). West Flanders therefore emerges as the most invasible part of the region. This is consistent with it holding a comparatively species-poor resident community that leaves more of the trait space unoccupied



and more room for functionally similar newcomers. As noted in the application summary, this should be read alongside the fact that West Flanders is also the most intensively farmed part of Flanders; low resident diversity there is at least partly a signature of agricultural land use, so the high modelled invasibility reflects an opportunity structure rather than an observed wave of invasions.

For the most broadly successful invaders, establishment is predicted across large, spatially patchy portions of Flanders rather than in tight range cores, echoing the region-wide suitability seen in the fitness maps. Notably, many of these top-ranking taxa are ornamental, garden or crop species. For example: squashes and pumpkins (*Cucurbita moschata*, *C. pepo*), watermelon (*Citrullus lanatus*), rice (*Oryza sativa*), morning glory (*Ipomoea purpurea*) and several woody ornamentals (*Rhododendron luteum*, *Rhus typhina*). Their prominence is a direct consequence of the inclusive, checklist-driven definition of the invader pool: species that are frequently recorded and functionally average score highly whether or not they are genuine ecological threats. The maps are therefore best read as showing where trait-and-climate conditions would permit establishment, not as forecasts that these particular species will spread.

Ranking invaders by their site-averaged establishment probability shows that predicted invasiveness spans a fairly narrow band - from roughly 45% to 54% mean establishment probability across the pool. The absence of a long high-invasiveness tail is itself informative: because candidate invaders are functionally similar to residents and to one another (Figure 50), the model separates them only weakly, and no small group stands out as dramatically more invasive than the rest. In practice this means that under the current trait-only pilot configuration the species-level ranking is less discriminating than the site-level invasibility maps, and it should be refined with a more targeted trait set and a stricter invader definition before being used to prioritise individual species.



Per-invader establishment ($\lambda > 0$, random subset of 21)

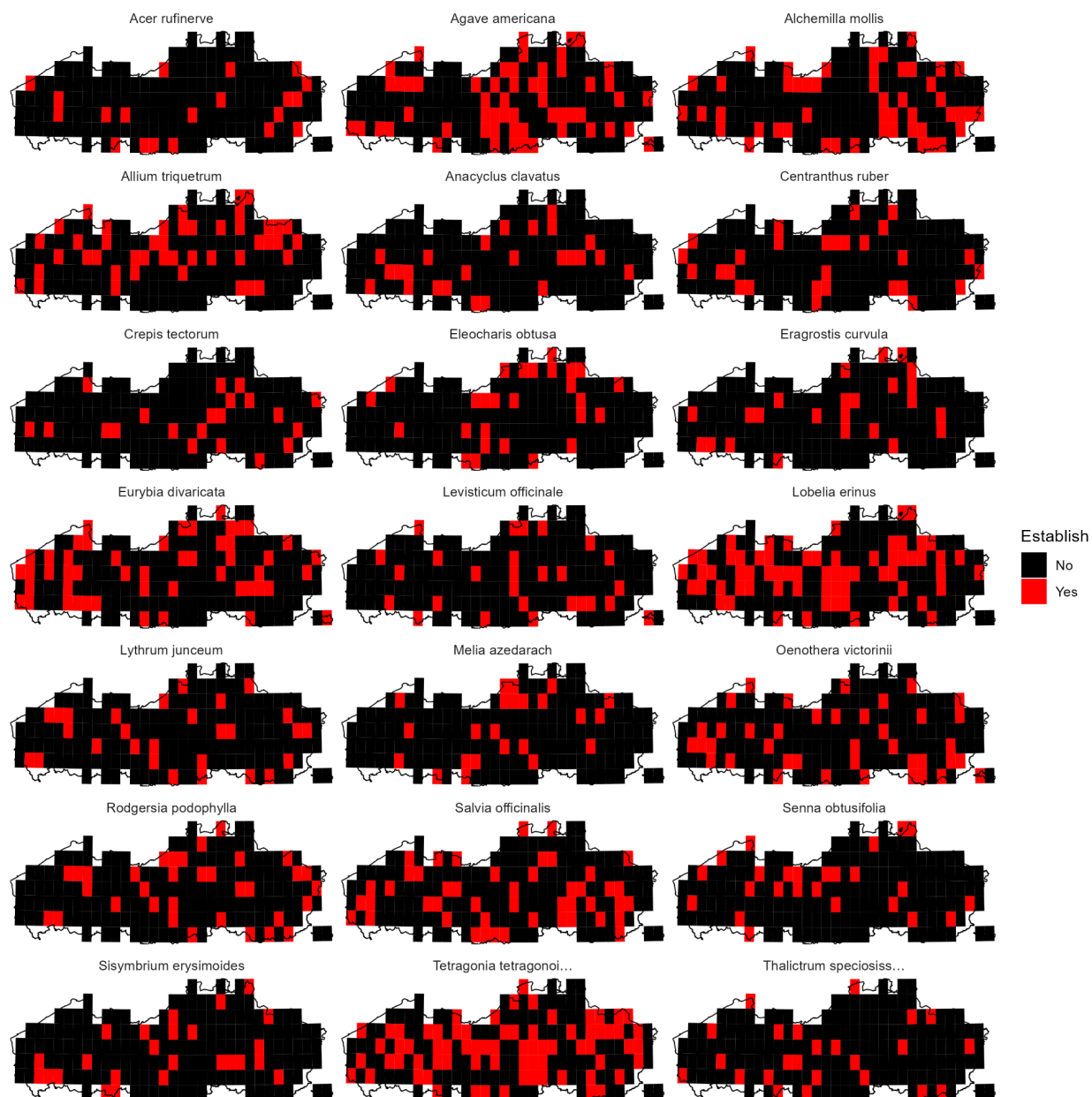


Figure 49: Predicted establishment maps for the 21 random selected species. Red = establishment predicted, black = not predicted.

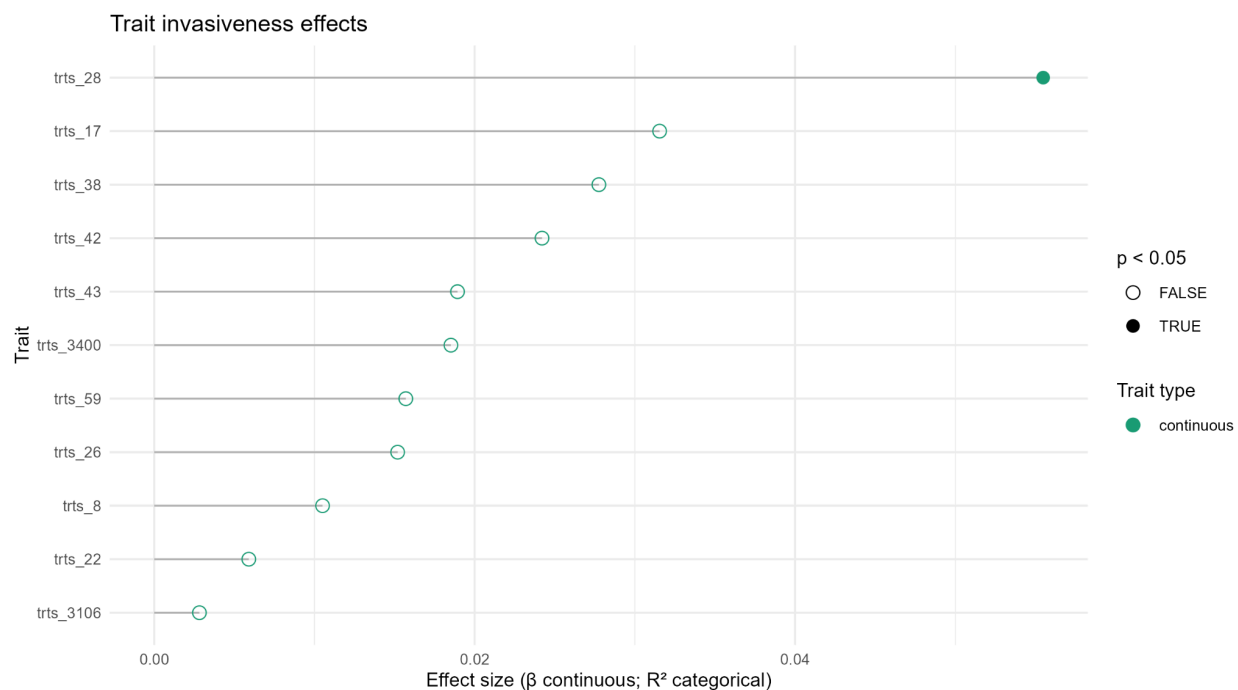


Figure 50: Effect of each functional trait on species invasiveness. Filled points denote effects significant at $p < 0.05$; the horizontal axis is the standardised effect size (β for continuous traits, R^2 for categorical traits). (trts_28: Dispersal syndrome, trts_17: Leaf compoundness, trts_38: Plant woodiness, trts_42: Plant growth form, trts_43: Leaf type, trts_3400: Plant growth form simple consolidated, trts_59: Plant lifespan (longevity), trts_26: Seed dry mass, trts_8: Plant nitrogen (N) fixation capacity, trts_22: Photosynthesis pathway, trts_3106: Plant height vegetative (from Kattge et al., 2020))

Only one trait, dispersal syndrome (trts_28), has a statistically significant association with invasiveness ($p < 0.05$); all other traits, including seed mass, plant height, woodiness, growth form and photosynthetic pathway, show small, non-significant effects. That dispersal syndrome is the single significant driver is ecologically intuitive, since the ability to disperse propagules underpins a species' capacity to reach and colonise new sites. At the same time the effect sizes are modest across the board, consistent with the trait-space result: when invaders and residents occupy overlapping trait space, individual traits carry limited power to discriminate against the most invasive species, and establishment is driven more by site conditions than by any single trait.

3.3.6.4. Cluster analysis

Finally, the species pool was summarised by clustering species on their trait values into 11 functional groups, to check whether the flora organises into distinct trait syndromes (Figure 51).

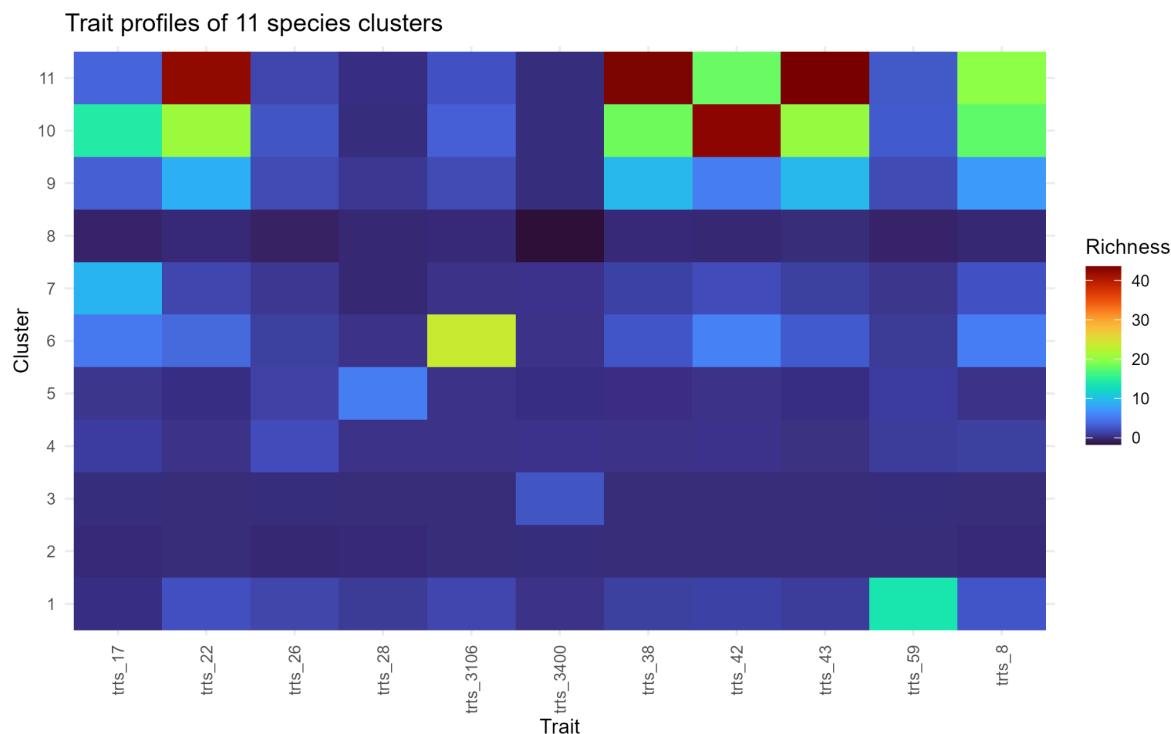


Figure 51: Mean trait profiles of 11 species clusters. Each cell shows a cluster's average (standardised) value for a given trait, so rows are functional 'syndromes' and columns are the eleven traits used in the analysis. (trts_28: Dispersal syndrome, trts_17: Leaf compoundness, trts_38: Plant woodiness, trts_42: Plant growth form, trts_43: Leaf type, trts_3400: Plant growth form simple consolidated, trts_59: Plant lifespan (longevity), trts_26: Seed dry mass, trts_8: Plant nitrogen (N) fixation capacity, trts_22: Photosynthesis pathway, trts_3106: Plant height vegetative (from Kattge et al., 2020))

The clustering resolves the flora into a handful of contrasting trait syndromes rather than a smooth continuum. Most clusters have muted, near-average profiles, but several are defined by one or two standout traits. For example, particular groups are distinguished by their dispersal syndrome, growth form and leaf type, another by plant height, and another by nitrogen-fixation capacity. This structure confirms that the eleven retained traits capture real functional differentiation among species, while also showing that much of the pool is functionally central, which again reinforces why invaders, drawn from this same pool, sit largely inside the resident trait core.

3.4. Natura 2000 versus non-Natura 2000.

For each of the general indicators calculated on native flowering plants in Flanders (see 3.1), we additionally made a comparison of that indicator between Natura 2000 and non-Natura 2000 habitat. All indicators are available via the Flanders Use Case dashboard. However, we describe in detail below how occurrence density and species richness density are particularly interesting indicators to apply when comparing cubes with differing total surface areas.



3.4.1. Occurrence density

3.4.2.1. General info

Indicator type

Data quality indicator.

Explanation

This metric measures the density of occurrence records by summing the number of occurrences per square kilometre within each grid cell. For maps, the value is calculated for each spatial cell. For time-series analyses, the density is summed across all grid cells in the data cube for each time step.

- High value: the grid cell contains a high concentration of occurrences.
- Low value: the occurrence records in this grid cell are sparse.

Applicability for policy

- Allows comparison of survey effort across scales: cubes with different grid sizes can easily be compared.
- Estimating citizen science engagement: it helps evaluate the true density of citizen science records.

However, occurrence density is linked to both survey effort and ecological diversity, which will be difficult to disentangle. They should therefore be interpreted with caution and ideally in combination with other indicators or expert knowledge to ensure correct interpretation.

Limitations

In case of citizen science data, this metric primarily reflects localised human search effort (like a popular walking path or a nature centre).

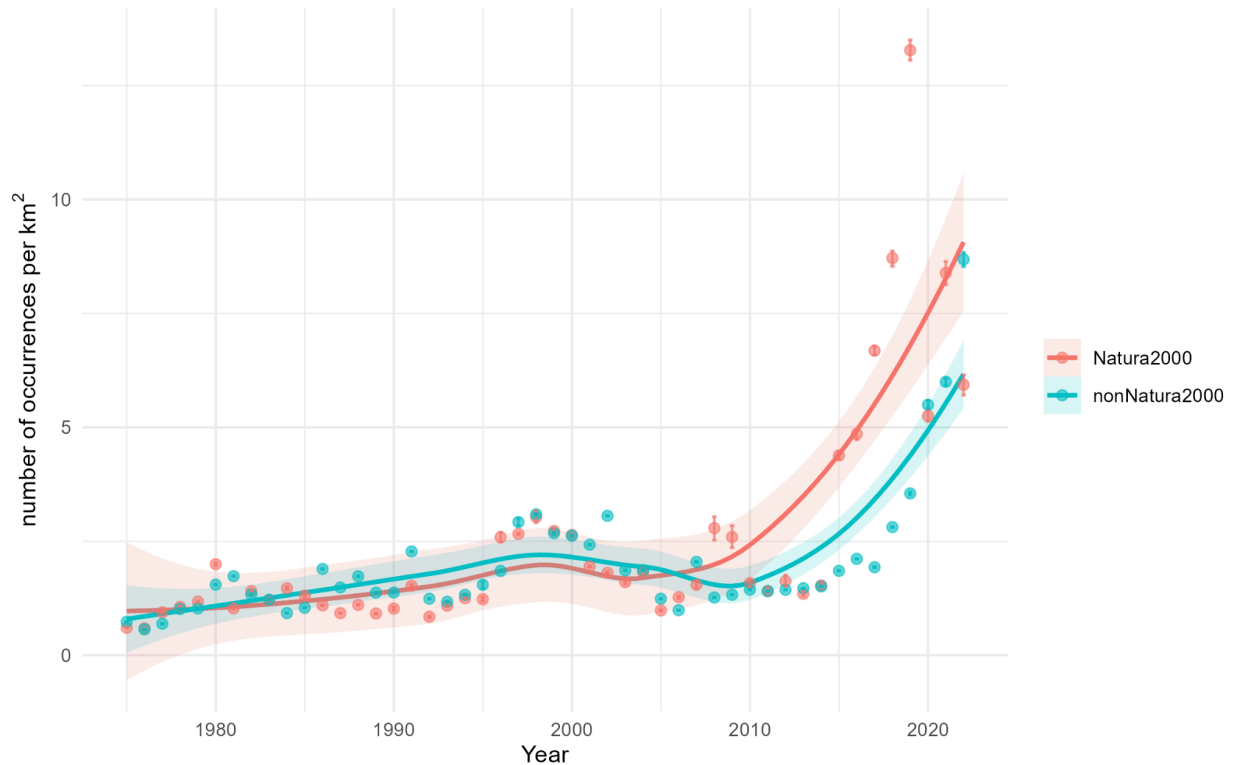


Figure 52: Calculating occurrence density allows comparisons of occurrence numbers between areas of different sizes. In this example, the density of native flowering plant occurrences recorded since 1975 is compared between Natura 2000 and non-Natura 2000 areas in Flanders. Data source: [GBIF.org](https://gbif.org) (2026e).

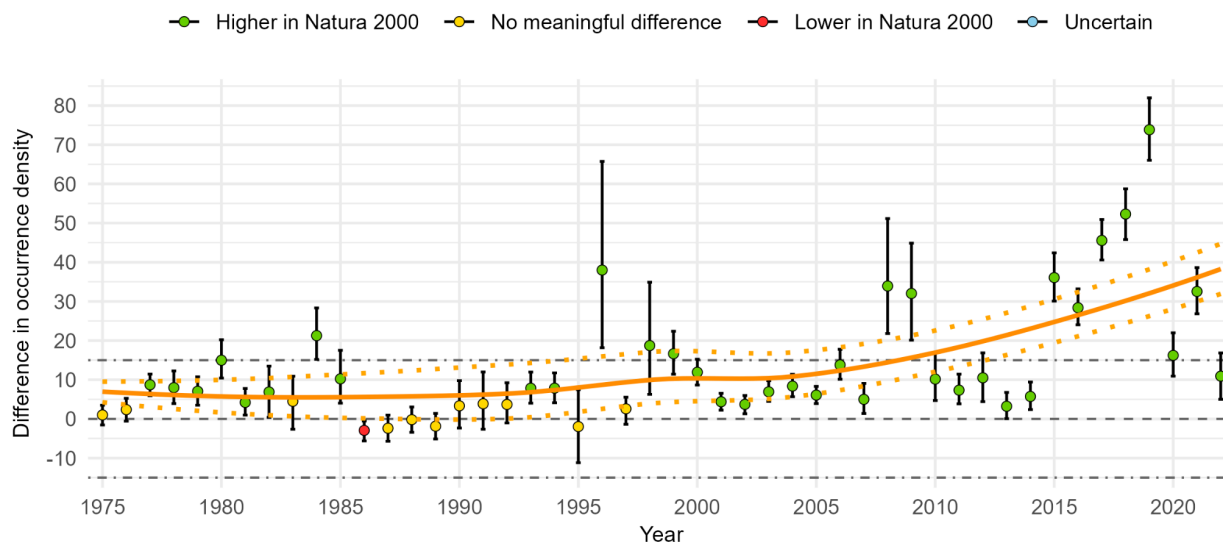
3.4.2.2. Comparison of indicators

The difference in occurrence density between Natura 2000 and non-Natura 2000 areas showed on average a similar pattern for the bootstrap and model-based analyses (Figure 53). Confidence intervals from the mixed-effects model were generally more homogeneous across years than those obtained from the bootstrap analysis. Unlike the bootstrap, which estimates uncertainty independently for each year, the mixed-effects model estimates common variance components from the complete dataset. In particular, the random intercept accounts for persistent differences in species richness among grid cells, while a common residual variance is used for all years. As a result, the estimated uncertainty is less sensitive to year-specific variability in the observed data.

In general, Natura 2000 grid cells show a higher number of occurrences than non-Natura 2000 grid cells throughout most of the study period. Interpretation of temporal trends should, however, be made with caution because no suitable measure of survey effort was available. While the mixed-effects model accounts for persistent differences among grid cells through a random intercept, it cannot correct for temporal changes in recording intensity or systematic differences in survey effort between Natura 2000 and non-Natura 2000 within individual years.



A. Bootstrap



B. Model-based

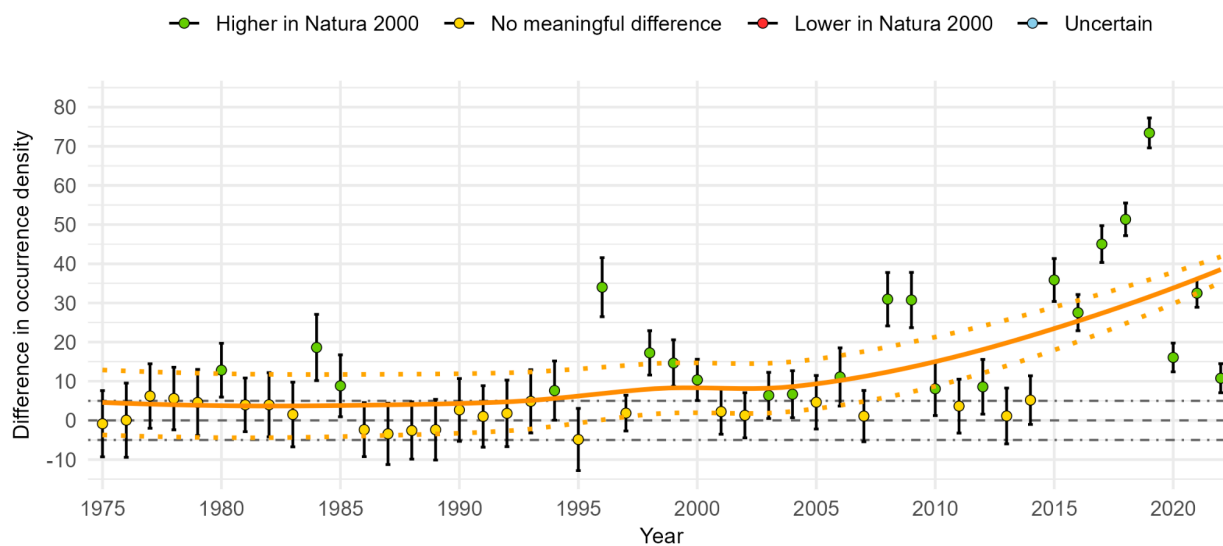


Figure 53: Difference in occurrence density between Natura 2000 and non-Natura 2000 areas based on bootstrapping (A) and a Gaussian GLMM (B). The trend is a LOESS smoother through the estimates and confidence limits. Effect classification is done with reference 0 and a threshold of a difference of 15 occurrences per km² (Table 6). Data source: [GBIF.org](https://gbif.org) (2026e).



3.4.2. Species richness density

3.4.2.1. General info

Indicator type

State Indicator.

Explanation

This indicator measures local biodiversity by dividing the total number of unique species by the surface area, representing the number of unique species found per square kilometre. For maps, the value is calculated for each spatial cell. For time-series analyses, species richness density is obtained by dividing the total number of unique species per year by the surface area of the entire cube.

- High value: The grid cell contains a high concentration of unique species.
- Low value: The area contains a low concentration of unique species.

Applicability for policy:

- Allowing comparison across scales: cubes with different grid sizes can easily be compared.
- Identifying hotspots of biodiversity.

By including a filter at a higher taxonomic rank, richness can be compared between different orders or kingdoms.

Limitations

Observed richness is highly dependent on the comprehensiveness of the dataset it is being applied to. If some regions are more intensively, carefully or systematically sampled than others, this will likely reflect as higher observed richness. Observed richness also depends on the relative abundance and spatial aggregation of each species, with less abundant and less aggregated species less likely to be discovered during surveys (Hillebrand et al., 2018), as well as the detectability of each species.

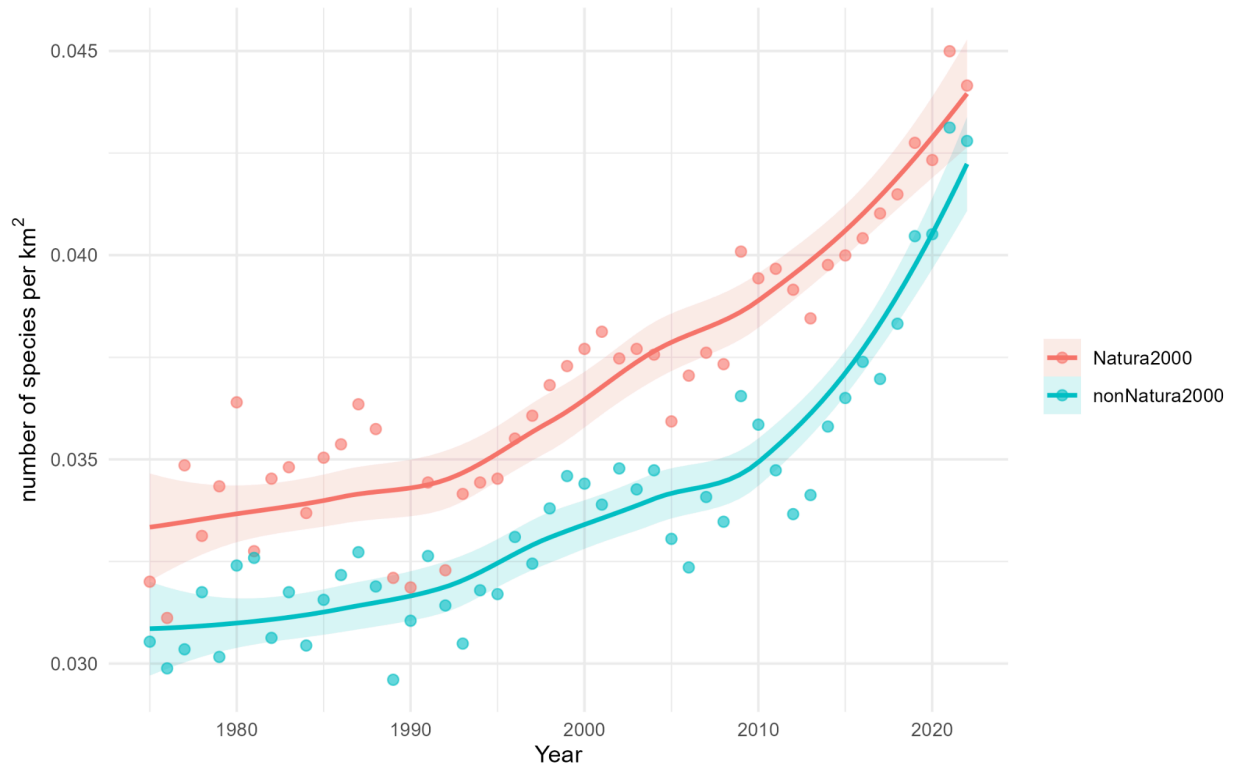


Figure 54: Calculating species richness density allows comparisons of species richness between areas of different sizes. In this example, the species richness density of native flowering plants recorded since 1975 is compared between Natura 2000 and non-Natura 2000 areas in Flanders. Data source: [GBIF.org](https://gbif.org) (2026e).

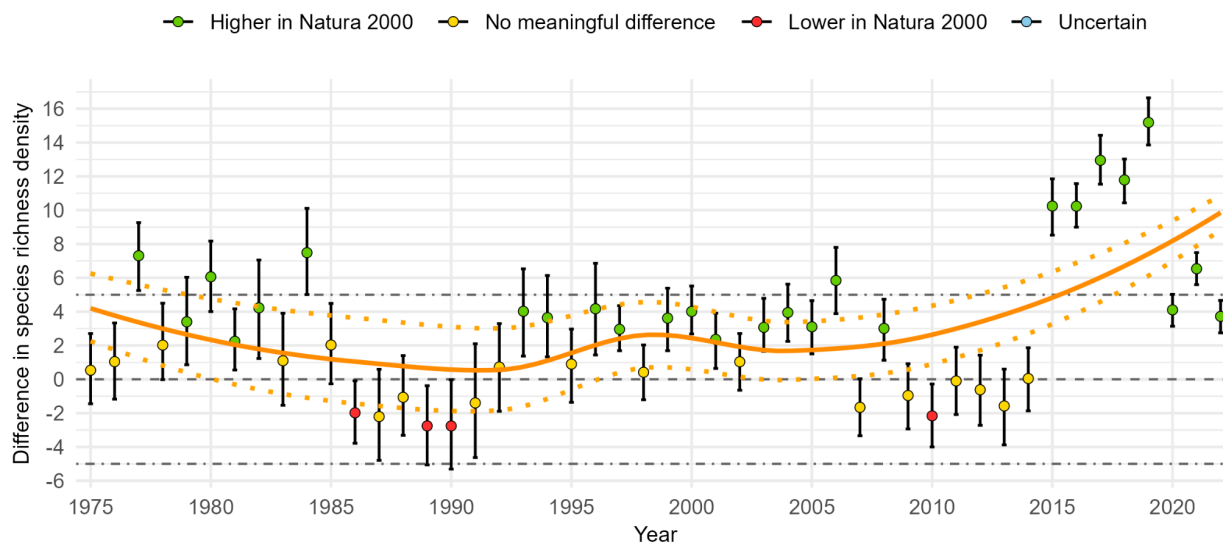
3.4.2.2. Comparison of indicators

The difference in mean local species richness between Natura 2000 and non-Natura 2000 areas showed similar patterns for the bootstrap and model-based analyses (Figure 55). Confidence intervals from the mixed-effects model were generally more homogeneous across years similar as with the occurrence density, although this difference is less pronounced for species richness density.

In general, Natura 2000 grid cells supported a higher local plant species richness than non-Natura 2000 grid cells throughout most of the study period. Interpretation of temporal trends should, however, be made with caution because no suitable measure of survey effort was available. While the mixed-effects model accounts for persistent differences among grid cells through a random intercept, it cannot correct for temporal changes in recording intensity or systematic differences in survey effort between Natura 2000 and non-Natura 2000 within individual years.



A. Bootstrap



B. Model-based

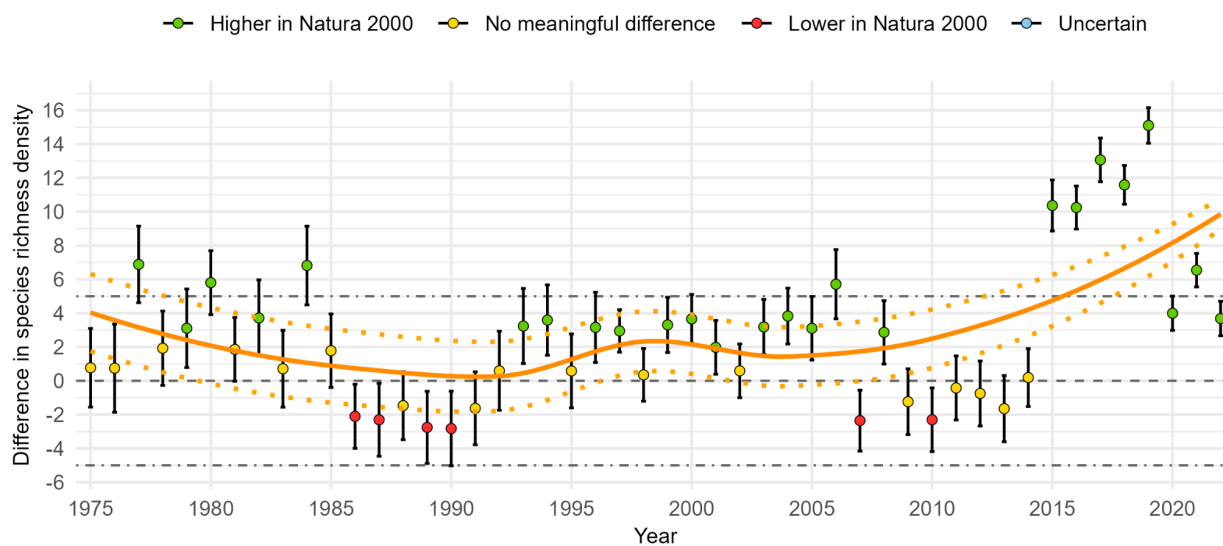


Figure 55: Difference in species richness density between Natura 2000 and non-Natura 2000 areas based on bootstrapping (A) and a Gaussian GLMM (B). The trend is a LOESS smoother through the estimates and confidence limits. Effect classification is done with reference 0 and a threshold of a difference of 5 species per km² (Table 6). Data source: [GBIF.org](https://gbif.org) (2026e).



3.5. WiSDM Flow and cubes

The WiSDM workflow is a framework developed to streamline the generation of species distribution models for invasive alien species (Davis et al., 2024). By automating the integration of species occurrence data, environmental variables, and modelling algorithms, WiSDM enables reproducible prediction of species distributions across space and time.

The workflow supports the production of current and future projections, allowing researchers and policymakers to assess potential species responses to climate change and other environmental pressures. Through the standardisation of modelling procedures and outputs, WiSDM facilitates transparency, comparability, and reuse of results across studies and applications.

As part of recent developments, data cubes have been incorporated as an additional output of the WiSDM workflow. They integrate taxonomic information (GBIF taxonKeys), temporal information (time periods and relevant climate scenarios), and spatial information (EEA grid cells) with the corresponding habitat suitability projected by WiSDM (Table 7). By transforming model predictions into data cubes, the workflow supports advanced analyses and additional visualisation options, such as the visualisation of trends in predicted habitat suitability over time and across climate scenarios (see Figure 56 as an example). This development represents an important step towards operationalising biodiversity data cubes as a component of data-driven risk assessments and conservation actions. This new version of WiSDM including cubes as output is available [here](#).

Table 7: Example of WiSDM output represented as a cube.

Species	TaxonKey	CELLCODE	Time	Scenario	Combined_favourability
Vespa_velutina	1311477	1kmE3938N3166	Current	Baseline	0.6778565049
Vespa_velutina	1311477	1kmE3939N3166	Current	Baseline	0.4840580821
Vespa_velutina	1311477	1kmE3940N3166	Current	Baseline	0.4273465872
Vespa_velutina	1311477	1kmE3955N3166	Current	Baseline	0.7715364695
Vespa_velutina	1311477	1kmE3956N3166	Current	Baseline	0.8720250726
Vespa_velutina	1311477	1kmE3957N3166	Current	Baseline	0.7240606546
Vespa_velutina	1311477	1kmE3958N3166	Current	Baseline	0.8665957451
Vespa_velutina	1311477	1kmE3959N3166	Current	Baseline	0.6290605068
Vespa_velutina	1311477	1kmE3960N3166	Current	Baseline	0.5101902485
Vespa_velutina	1311477	1kmE3934N3165	Current	Baseline	0.7953375578
Vespa_velutina	1311477	1kmE3935N3165	Current	Baseline	0.8402297497

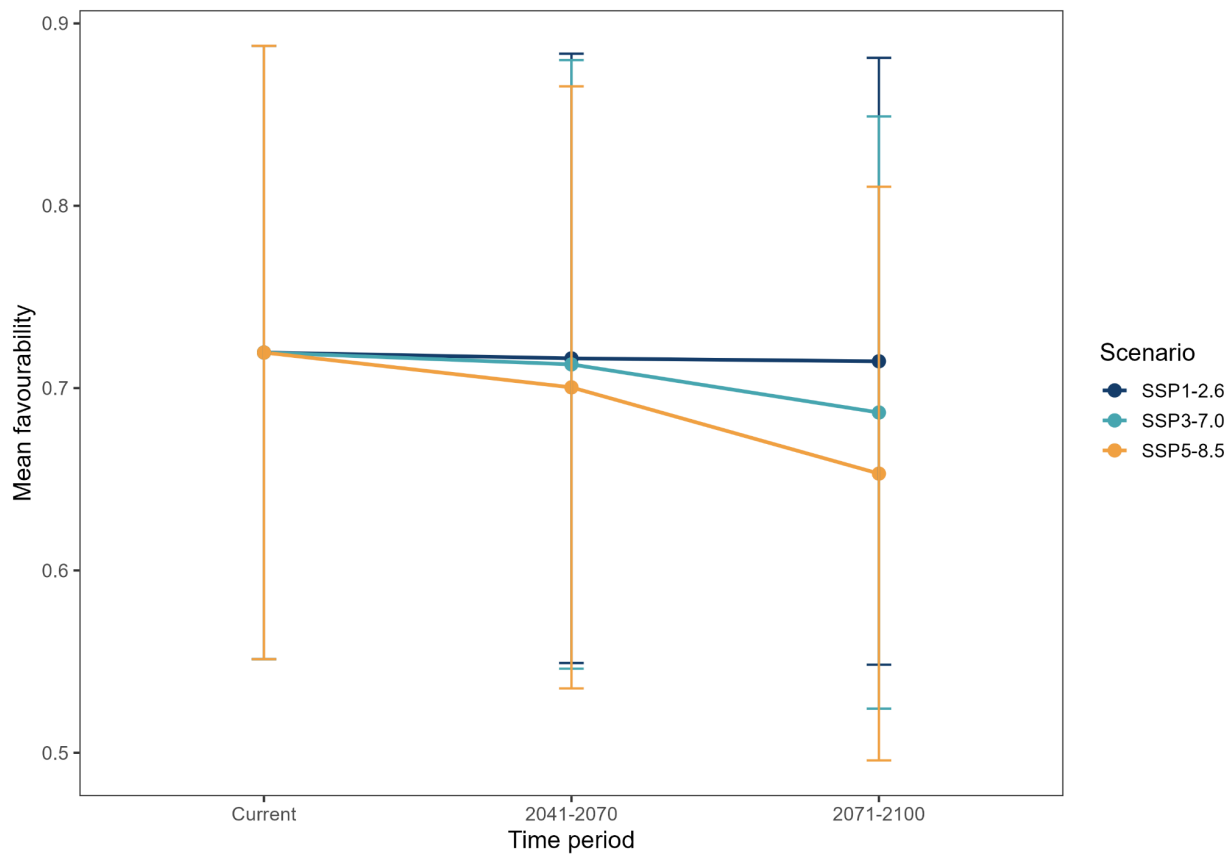


Figure 56: A WiSDM prediction represented as a data cube enables the visualization of trends in predicted habitat suitability over time and across climate scenarios (indicated by different colours). This example shows predicted favourability for *Vespa velutina* in Belgium across three time periods under three climate scenarios. Dots represent mean favourability, while error bars indicate the standard deviation.



4. Discussion and future perspectives

This study demonstrates that B3 indicators can effectively transform large biodiversity datasets into useful information for evidence-based policymaking. The indicator factsheets and public dashboard improve the practical use of B3 workflows.

4.1. General findings

A major strength of this work is the integration of different types of indicators, including general biodiversity, phylogenetic and invasive alien species indicators. Together, these indicators provide complementary information on biodiversity status, ecological change and data quality. In addition, the comparison between Natura 2000 and non-Natura 2000 areas illustrates how the developed workflows can support the evaluation of conservation measures in protected areas.

The use of interoperable data products (geojson files shared on Zenodo) and close collaboration with existing biodiversity platforms such as the Belgium GBIF, are critical in supporting long-term maintenance of developed workflows.

This study highlights several challenges associated with the use of GBIF data. Most importantly, variation in survey effort remains a major source of uncertainty. Spatial and temporal variation in recording intensity, particularly in citizen science datasets, can strongly influence indicator values and apparent trends. In Flanders the number of occurrence records per km² is higher in Natura 2000 than in non-Natura 2000 area, making it difficult to disentangle genuine ecological patterns from sampling bias. Although approaches such as estimating completeness, focusing on specific datasets (e.g. only using Florabank data), and restricting analyses to well-surveyed cells can improve the interpretation of results, they cannot completely eliminate sampling bias. Consequently, biodiversity indicators should always be interpreted alongside measures of data quality and expert knowledge.

Large differences in data availability among taxa are observed in GBIF. These differences largely depend on whether long-term monitoring programmes are conducted in the study area and whether their data have been published. For instance, GBIF contains extensive data on birds and plants in Flanders, whereas data on insects are much less abundant.

The results also illustrate that not all indicators are equally suitable for every application. Some indicators proved robust and informative, whereas others were strongly affected by uneven sampling effort and were therefore excluded from the dashboard. This underlines the importance of critically evaluating indicators before applying them in policy or management. The case study also stimulated the development of new indicators, such as completeness, relative occupancy and species richness density. Some indicators would also benefit from having a spatial component, allowing them to be visualized on maps, rather than being analyzed only over time in a time series (e.g. the turnover index). This highlights the importance of testing developed workflows in practical use cases.

4.2. Modifiable areal unit problem



The results of spatially aggregated biodiversity indicators may be affected by the modifiable areal unit problem (MAUP), whereby summary values and statistical relationships change according to the size, shape and alignment of the spatial units used. This includes a scale effect, where results differ among 1-km, 5-km and 10-km grids, and a zoning effect, where results change if a grid of the same resolution is shifted or replaced by differently shaped units.

In this case study, MAUP may influence mapped patterns of species richness, occurrence density, completeness, phylogenetic diversity and other indicators. Fine-resolution cells reveal more local variation but contain fewer records and are more sensitive to survey effort and coordinate uncertainty. Coarser cells produce smoother and often more stable patterns, but may conceal local ecological differences. Grid alignment may also affect which observations and proportions of Natura 2000 habitat fall within each cell, potentially altering comparisons between Natura 2000 and non-Natura 2000 areas.

The indicators should therefore be interpreted as conditional on the spatial framework used. For example, because nearby species occurrences tend to be spatially autocorrelated, a square grid cell would be expected, on average, to contain more species than a hexagonal cell of the same area, because a square has a longer perimeter and encompasses locations that are more widely separated. We did not conduct a formal sensitivity analysis across alternative grid sizes, shapes or alignments, so the robustness of the results to MAUP remains uncertain. Spatial resolution and grid geometry should therefore always be reported with the indicator, and comparisons among studies should preferably use the same spatial framework.

4.3. **invasimapr** as a regional invasion indicator

Within Deliverable 6.3, **invasimapr** occupies a distinctive niche among the tested indicators. The general, phylogenetic and invasive-alien-species indicators presented earlier are largely descriptive: they summarise what has been observed, namely occurrences, richness, newness, checklist status and emerging-status trends. **invasimapr** instead turns the same GBIF occurrence cubes into a predictive, mechanistically framed indicator by combining occurrences with functional traits and climate to estimate where the region is open to invasion (site invasibility) and which species are best placed to exploit that opening (species invasiveness). It therefore complements the checklist-based and occurrence-based invasive-alien-species indicators of Section 3.3 by looking forward to potential establishment rather than backwards at realised presence.

The pilot application to Flanders produced a coherent and interpretable signal. The candidate invaders are functionally similar to the resident flora (Figure 46), so invasion outcomes are governed mainly by environmental suitability and community saturation rather than by novel traits; dispersal syndrome is the only trait significantly linked to invasiveness (Figure 50); and both the mean-fitness and invasibility maps identify West Flanders as the most invasible part of the region (Figures 47 and 48). The last result is ecologically sensible under the invasibility-cube framework, since a species-poor, trait-depauperate community offers more unoccupied niche space, but it is confounded with intensive agricultural land use and with sampling effort, and so should be read as an opportunity structure rather than a prediction of imminent invasion.

As a policy tool the indicator's value is primarily spatial and preventative. Invasibility maps can help target surveillance and early-warning effort towards the areas most likely to admit new



establishments, support horizon-scanning and watch-list prioritisation, and feed risk screening for protected areas (for example the Natura 2000 comparisons in section 3.4 in this deliverable) and reporting against instruments such as the EU Invasive Alien Species Regulation (Regulation (EU) No 1143/2014) and Target 6 on invasive species. Because it is built entirely on openly available GBIF, TRY and WorldClim data through the reproducible **disssmapr/invasimapr** workflow, it can in principle be re-run for any region with comparable data coverage.

Several limitations temper these conclusions and define the work still needed. First, this run was an explicit reproducibility pilot that used the default TRY trait database with a coverage-based trait selection; the retained traits are broad functional descriptors rather than traits chosen for their relevance to invasion, which likely explains their weak, largely non-significant effects. Second, the candidate-invader pool is defined inclusively through tolerant fuzzy matching to broad checklists, so it contains many cultivated and casual species and overstates the number of genuine invaders, which inflates establishment counts and compresses the invasiveness ranking. Third, the indicator inherits the biases of opportunistic occurrence data: richness and, through it, modelled invasibility partly track survey effort, and the analysis is correlative rather than a mechanistic population model. Finally, the 0.1° grid is coarse for operational targeting.

Future iterations should therefore (i) replace the default trait set with a curated, invasion-relevant selection and improve trait coverage; (ii) tighten the invader definition to a validated list of species of concern; (iii) correct for survey effort and completeness using the (dubi)cube quality checks applied elsewhere in this deliverable; and (iv) validate the predicted invasibility and invasiveness against independent records of known invasions in Flanders before the outputs are used to guide management. With these refinements, **invasimapr** has the potential to become a transferable, occurrence-cube-based indicator of ecosystem invasibility that adds a forward-looking dimension to the B3 indicator suite.

4.4. Limitations and recommendations

Using cubes to calculate biodiversity indicators has multiple advantages. However, it is important to bear in mind the limitations of this approach. These limitations relate to spatial, temporal and dataset aspects. Some of the issues identified in the data quality analysis (Langerhaert et al., 2026) also arise in this case study. In this section, we discuss a number of general considerations and limitations. Specific limitations linked to individual indicators are discussed within the respective chapters above.

4.4.1. Dominance of a single or a few datasets

The GBIF data from Flanders is characterised by the dominance of a single dataset. The number of occurrences fell abruptly in 2018 because observations of native species from this dataset have not been submitted to GBIF since then. This decline significantly limits the usability of the dataset as a whole. For temporal indicators it hinders trend analysis and for spatial indicators, calculated over a longer period, considerable weight is given to relatively outdated data.

This bottleneck was known from the start of the Flanders case study. Species groups were selected that are provided from other datasets in GBIF and for which this fallback does not



apply. When there is little familiarity with the individual datasets that feed into GBIF, it is strongly recommended that a sensitivity analysis is carried out to determine the influence of each dataset in terms of space, time and species. This can help to inform decisions about excluding certain datasets, restricting the analysis to a suitable time period or filtering out specific overrepresented species. The **dubicube** package includes a number of built-in functions for carrying out such a sensitivity analysis as well as visualising and interpreting the results.

4.4.2. Publication delay

The analysis takes publication into account. While observations published via citizen science platforms, such as iNaturalist, are available almost immediately, it typically takes two or more years for validated datasets of the monitoring data from research institutions to be published on GBIF. The publication delay also varies from dataset to dataset. To avoid misinterpretation of trends, it is advisable to filter out the most recent years. For this case, a margin of three years was applied.

4.4.3. Observation bias

One major limitation stems from the uncertainty about the survey effort or the underlying sampling design. Three complementary aspects of observation can be distinguished: survey effort, survey completeness, and species detectability. Some of the indicators in the case study are proxies for these aspects and help to distinguish between genuine biodiversity patterns and effects from the survey methods, technology or the behaviour of surveyors. A number of initially proposed indicators were not included in the Flanders case due to their high sensitivity to observation bias.

4.4.4. Spatial uncertainty

This case study also emphasises the limitations due to spatial uncertainty. A large number of measurements with low spatial precision can blur or distort spatial patterns. The resolution of the grid should not be smaller than the spatial accuracy of the large majority of the observations. The higher the spatial resolution of the data cube, the greater the effect of spatial precision. It is recommended to filter out the records with low or unknown spatial accuracy. It is a trade-off between data loss due to filtering and the level of detail provided by the higher resolution. In the case of temporal indicators that cover a large enough area, priority may be given to the completeness of the data.

4.5. Future prospects

Future work should focus on improving methods to account for sampling effort and further expanding the set of available indicators. In addition, the publication of long-term monitoring datasets and biodiversity data in accordance with the FAIR principles (Wilkinson et al., 2016) should be encouraged. This will increase the value of B3 workflows for biodiversity monitoring and conservation.



Besides occurrence data, data on invasive alien species management should also be published to GBIF following the Darwin core standard. Work to extend the Darwin core standard for management data is currently ongoing (Reyserhove et al., 2022). This is critical in order to develop indicators specifically focusing on IAS management based on GBIF data.

Close collaboration between users and developers of the B3 workflow has proven to be highly valuable and should continue throughout the afterlife of the project.

The Flanders Use Case dashboard is scheduled to be updated in August 2026 and will subsequently be updated manually on an annual basis. The code supporting the dashboard, including the preprocessing workflow and raw indicator files, will also be deposited on Zenodo and updated annually.

Discussions with the maintainers of the [Belgian GBIF node](#) are ongoing. Once it has been determined which indicators should be included for the different taxonomic groups, these indicators will be generated and made available as GeoJSON files on Zenodo.

Finally, the [Alien Species Portal](#) will be finalised as part of B3 and is expected to be launched at the end of 2026.



Acknowledgements

We would like to thank Shawn Dove, Katelyn Faulkner, and Tsungai Zengeya for their thorough review and valuable input. Furthermore, we would like to thank Shawn Dove and Maarten Trekels for their support during the testing of the packages **b3gbi** and **b3alien**, respectively. Finally, we would like to thank all colleagues and collaborators who contributed to the development of this deliverable through discussions, feedback, and practical support. In particular, our gratitude goes to Wouter Van Landuyt, Toon Spanhove, Fleur Petersen and Wouter Van Reeth for their valuable feedback.

The authors acknowledge the use of ChatGPT (OpenAI) for support with language editing and improving the clarity and structure of parts of this document. The authors remain fully responsible for the content, interpretation, and conclusions presented.



References

- Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, 16, 3–9.
- Adriaens, T., D'hondt, B., Denys, L., Packet, J., Scheers, K., Oldoni, D., & Steen, F. (2023). *DVW - priority species list of the Flemish Waterways Authority* [Checklist dataset]. Research Institute for Nature and Forest (INBO). <https://doi.org/10.15468/52b8h9>
- Agentschap voor Natuur en Bos. (2005). *Vogelrichtlijngebieden* [Data set]. <https://www.vlaanderen.be/datavindplaats/catalogus/vogelrichtlijngebieden>
- Agentschap voor Natuur en Bos. (2014). *Habitatrichtlijn(deel)gebieden* [Data set]. <https://www.vlaanderen.be/datavindplaats/catalogus/habitatrichtlijndeelgebieden>
- Baltes, P. B., & Kunzmann, U. (2003). Wisdom. *The Psychologist*, 16(3), 131–132.
- Breugelmans, L., Trekels, M., & Hendrickx, L. (2025). *pdindicatorR: Calculate and visualize phylogenetic diversity indicators based on species occurrence data cubes* (Version 0.1.0) [R package]. <https://doi.org/10.5281/zenodo.14237551>
- Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9(2), 378–400. <https://doi.org/10.32614/RJ-2017-066>
- Chamberlain, S., Barve, V., McGlinn, D., Oldoni, D., Desmet, P., Geffert, L., & Ram, K. (2026). *rgbif: Interface to the Global Biodiversity Information Facility API* (Version 3.8.5.1) [R package]. <https://CRAN.R-project.org/package=rgbif>
- Davis, A. J. S., Groom, Q., Adriaens, T., Vanderhoeven, S., De Troch, R., Oldoni, D., Desmet, P., Reyserhove, L., Lens, L., & Strubbe, D. (2024). Reproducible WiSDM: A workflow for reproducible invasive alien species risk maps under climate change scenarios using standardized open data. *Frontiers in Ecology and Evolution*, 12, Article 1148895. <https://doi.org/10.3389/fevo.2024.1148895>
- Desmet, P. (2025). *Download a species occurrence cube from GBIF.org* [Tutorial]. B-Cubed. <https://docs.b-cubed.eu/tutorials/download-a-cube-from-gbif/>
- Desmet, P., Reyserhove, L., Oldoni, D., Groom, Q., Adriaens, T., Vanderhoeven, S., & Pagad, S. (2025). *Global Register of Introduced and Invasive Species - Belgium* (Version 1.15) [Checklist dataset]. Invasive Species Specialist Group ISSG. <https://doi.org/10.15468/xoidmd>
- Dove, S. (2026). *b3gbi: General biodiversity indicators for biodiversity data cubes* (Version 0.9.3) [R package]. <https://github.com/b-cubed-eu/b3gbi>
- European Commission. (2025). *Natura 2000: The EU nature protection network*. https://environment.ec.europa.eu/topics/nature-and-biodiversity/natura-2000_en



European Parliament, & Council of the European Union. (2014). Regulation (EU) No 1143/2014 of 22 October 2014 on the prevention and management of the introduction and spread of invasive alien species. *Official Journal of the European Union*, L 317, 35–55. <https://eur-lex.europa.eu/eli/reg/2014/1143/oj>

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61(1), 1–10. [https://doi.org/10.1016/0006-3207\(92\)91201-3](https://doi.org/10.1016/0006-3207(92)91201-3)

Faith, D. P. (2015). Phylogenetic diversity, functional trait diversity and extinction: Avoiding tipping points and worst-case losses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1662), Article 20140011. <https://doi.org/10.1098/rstb.2014.0011>

Gari, S. R., Newton, A., & Icely, J. D. (2015). A review of the application and evolution of the DPSIR framework with an emphasis on coastal social–ecological systems. *Ocean & Coastal Management*, 103, 63–77. <https://doi.org/10.1016/j.ocecoaman.2014.11.013>

GBIF.org (28 July 2026a) GBIF Occurrence Download <https://doi.org/10.15468/dl.pwky78>

GBIF.org (16 March 2026b) GBIF Occurrence Download <https://doi.org/10.15468/dl.knsn44>

GBIF.org (3 June 2026c) GBIF Occurrence Download <https://doi.org/10.15468/dl.pp4qw6>

GBIF.org (10 June 2026d) GBIF Occurrence Download <https://doi.org/10.15468/dl.v6qjgj>

GBIF.org (16 March 2026e) GBIF Occurrence Download <https://doi.org/10.15468/dl.knsn44>

GBIF.org (26 March 2026f) GBIF Occurrence Download <https://doi.org/10.15468/dl.uk8t4g>

Groom, Q., Vanderhoeven, S., Eggermont, H., Coupremagne, M., Adriaens, T., Desmet, P., Reyserhove, L., Oldoni, D., Strubbe, D., Davis, A., Lens, L., Monty, A., Mahy, G., Jacquemart, A.-L., Verleye, T., Van Hoey, G., Van De Kerchove, R., Backeljau, T., Dispas, M., ... Termonia, P. (2022). *Tracking invasive alien species: Building a data-driven framework to inform policy* (Final report, Contract BR/165/A1/TrIAS). Belgian Science Policy Office. https://www.belspo.be/belspo/brain-be/projects/FinalReports/TrIAS_FinRep.pdf

Gumbs, R., Chaudhary, A., Daru, B. H., Faith, D. P., Forest, F., Gray, C. L., Kowalska, A., Lee, W.-S., Pellens, R., Pipins, S., Pollock, L. J., Rosindell, J., Scherson, R. A., & Owen, N. R. (2023). Indicators to monitor the status of the tree of life. *Conservation Biology*, 37(6), Article e14138. <https://doi.org/10.1111/cobi.14138>

Harmonia database. (2026). *Harmonia database*. Belgian Forum on Invasive Species. <http://ias.biodiversity.be>

Hijmans, R. (2026). *geodata: Access geographic data* (Version 0.6-10) [R package]. <https://github.com/rspatial/geodata>



Hillaert, J., Breugelmans, L., Langerhaert, W., Adriaens, T., Delva, S., MacFadyen, S., Oldoni, D., Yovcheva, N., Van Daele, T., & Groom, Q. (2026). *Factsheets for B-Cubed indicators* (Version 1.0) [Data set]. Pensoft Publishers. <https://doi.org/10.5281/zenodo.21644242>

Hillebrand, H., Blasius, B., Borer, E. T., Chase, J. M., Downing, J. A., Eriksson, B. K., Filstrup, C. T., Harpole, W. S., Hodapp, D., Larsen, S., Lewandowska, A. M., Seabloom, E. W., Van de Waal, D. B., & Ryabov, A. B. (2018). Biodiversity change is uncoupled from species richness trends: Consequences for conservation and monitoring. *Journal of Applied Ecology*, 55(1), 169–184. <https://doi.org/10.1111/1365-2664.12959>

IUCN (2017). *Guidance for interpretation of CBD categories on introduction pathways*. Technical note prepared by the International Union for Conservation of Nature for the European Commission.

Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et du Jura. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37(142), 547–579. <https://doi.org/10.5169/seals-266450>

Kattge, J., Bönsch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Tautenhahn, S., Werner, G. D. A., Aakala, T., Abedi, M., Acosta, A. T. R., Adamidis, G. C., Adamson, K., Aiba, M., Albert, C. H., Alcántara, J. M., Alcázar, C., Aleixo, I., Ali, H., ... Wirth, C. (2020). TRY plant trait database – Enhanced coverage and open access. *Global Change Biology*, 26(1), 119–188. <https://doi.org/10.1111/gcb.14904>

Langerhaert, W., Faulkner, K., Cartuyvels, E., Groom, Q., & Van Daele, T. (2026). *Report on the criteria for data quality and species characteristics for estimating species status and trends* (B3 project deliverable D4.3). B-Cubed. <https://b-cubed.eu/storage/app/uploads/public/69a/943/ea1/69a943ea17e66799908588.pdf>

Langerhaert, W., & Van Daele, T. (2026a). *dubicube: Calculation and interpretation of data cube indicator uncertainty* (Version 0.12.3) [R package]. <https://doi.org/10.5281/zenodo.14850237>

Langerhaert, W., & Van Daele, T. (2026b). *Report of the indicators of robustness assessments* (B3 project deliverable D5.4). B-Cubed. <https://b-cubed.eu/storage/app/uploads/public/69f/319/fd1/69f319fd13c93876864613.pdf>

Lenth, R. V., & Piaskowski, J. (2026). *emmeans: Estimated marginal means, aka least-squares means* (Version 2.0.4) [R package]. <https://rvlenth.github.io/emmeans/>

MacFadyen, S., & Hui, C. (2026). *dissmapr: A computational workflow for compositional dissimilarity and biodiversity turnover analysis* (Version 0.2.1) [R package]. Zenodo. <https://doi.org/10.5281/zenodo.21363799>

MacFadyen, S., Yahaya, M. M., Trekels, M., Kumschick, S., Landi, P., & Hui, C. (2025). *invasimapr: Workflow to Visualise Trait Dispersion and Assess Invasibility* (Version 0.2.0) [R package]. Zenodo. <https://doi.org/10.5281/zenodo.20842472>



McGillycuddy, M., Popovic, G., Bolker, B. M., & Warton, D. I. (2025). Parsimoniously fitting large multivariate random effects in glmmTMB. *Journal of Statistical Software*, 112(1), 1–19. <https://doi.org/10.18637/jss.v112.i01>

Naczka, A. M., Wilhelm, M., Jakubska-Busse, A., Kalinka, A., Achrem, M., Androsiuk, P., Górniak, M., Zarzycka, M., & Kolanowska, M. (2025). Climate change-driven northward expansion of the Mediterranean orchid *Ophrys apifera* from genetic and ecological perspectives. *Scientific Reports*, 15, Article 29634. <https://doi.org/10.1038/s41598-025-15167-6>

Oldoni, D., Delva, S., Devisscher, S., & Varewyck, M. (2026). trias: Process Data for the Project Tracking Invasive Alien Species (TrIAS) (Version 3.2.4) [Computer software]. <https://github.com/trias-project/trias>

R Core Team. (2026). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://doi.org/10.32614/R.manuals>

Reyserhove, L., Hillaert, J., Adriaens, T., Desmet, P., D'hondt, B., Groom, Q. J., & Oldoni, D. (2022). *manIAS: A community-driven data model and data exchange format for the management of invasive alien species and wildlife*. *Biodiversity Information Science and Standards*, 6, e93453. <https://doi.org/10.3897/biss.6.93453>

Robuchon, M., da Silva, J., Dubois, G., Gumbs, R., Hoban, S., Laikre, L., Owen, N. R., & Perino, A. (2023). Conserving species' evolutionary potential and history: Opportunities under the Kunming–Montreal Global Biodiversity Framework. *Conservation Science and Practice*, 5(6), Article e12929. <https://doi.org/10.1111/csp2.12929>

Rowley, J. E. (2007). The wisdom hierarchy: Representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), 163–180. <https://doi.org/10.1177/0165551506070706>

Solow, A. R., & Costello, C. J. (2004). Estimating the rate of species introductions from the discovery records. *Ecology*, 85(7), 1822–1825. <https://doi.org/10.1890/03-3102>

Statistiek Vlaanderen. (2026, July 14). *Bevolking: Omvang en groei*. Vlaanderen.be. <https://www.vlaanderen.be/statistiek-vlaanderen/bevolking/bevolking-omvang-en-groei>

Storkey, J., Meyer, S., Still, K. S., & Leuschner, C. (2012). The impact of agricultural intensification and land-use change on the European arable flora. *Proceedings of the Royal Society B: Biological Sciences*, 279(1732), 1421–1429. <https://doi.org/10.1098/rspb.2011.1686>

Swinnen, K., Bronne, L., Herremans, M., Gielen, K., Adriaens, T., Baartmans, T., Baert, D., Baltus, H., Barendse, R., Boeraeve, M., Boers, K., Bogaert, J., Bosteels, Y., Bouillard, N., Claerebout, S., De Knijf, G., De Beer, D., De Smedt, P., De Cock, R., ... Desmet, P. (2025). *Waarnemingen.be / observations.be—List of species observed in Belgium* (Version 1.3) [Checklist data set]. Natuurpunt. <https://doi.org/10.15468/a7wkuh>

Trekels, M., & Groom, Q. (2025). *b3alien: A Python package to assess the introduction rate of alien species in a FAIR and reproducible way* [Preprint]. bioRxiv. <https://doi.org/10.1101/2025.11.17.688820>



Van Landuyt, W., & Brosens, D. (2026). *Florabank1—A grid-based database on vascular plant distribution in the northern part of Belgium (Flanders and the Brussels Capital Region)* (Version 45.44) [Occurrence data set]. Research Institute for Nature and Forest (INBO). <https://doi.org/10.3897/phytokeys.12.2849>

Van Landuyt, W., Hoste, I., Vanhecke, L., Van den Bremt, P., Vercruysse, W., & De Beer, D. (2006). *Atlas van de flora van Vlaanderen en het Brussels Gewest*. Flo.Wer, Instituut voor Natuur- en Bosonderzoek, & Nationale Plantentuin van België. <http://flora.inbo.be/>

Van Landuyt, W., & Westra, T. (2015). *Een meetnet voor de flora van Vlaanderen: Een verkennend onderzoek naar de mogelijkheden om een meetnet voor de flora van Vlaanderen uit te bouwen* (Rapporten van het Instituut voor Natuur- en Bosonderzoek No. INBO.R.2015.7741571). Instituut voor Natuur- en Bosonderzoek. <https://pureportal.inbo.be/en/publications/een-meetnet-voor-de-flora-van-vlaanderen-een-verkennend-onderzoek/>

Van Moer, K., Brosens, D., Cartuyvels, E., Adriaens, T., Baert, K., Devisscher, S., Neukermans, A., & Huysentruyt, F. (2026). *MICA - Muskrat occurrences collected by RATO in East Flanders, Belgium* (Version 1.17) [Occurrence data set]. RATO vzw. <https://doi.org/10.15468/5fps96>

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, Article 160018. <https://doi.org/10.1038/sdata.2016.18>