Collaborative AI for plant biodiversity monitoring: From Pl@ntNet to GeoPl@ntNet

Pierre Bonnet, Alexis Joly, Antoine Affouard, Rémi Palard, Maxime FromeHoltz,
Benjamin Deneu, Hervé Goëau, J.C. Lombardo, Mathias Chouet, Hugo Gresse,
Cesar Leblanc, Maximilien Servajean, François Munoz

PART I

Pl@ntNet under the hood

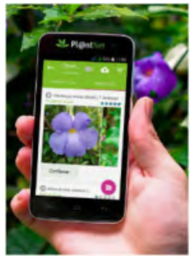
A citizen science platform that uses AI to help people identify plants with their mobile phones

Pl@ntNet app

25 Million users
200+ countries
Up to 2M identifications per day

Personal Usage

Nature, walks



Gardening

Phytotherapy

Professional Usage

Agro-ecology

Natural Areas Management

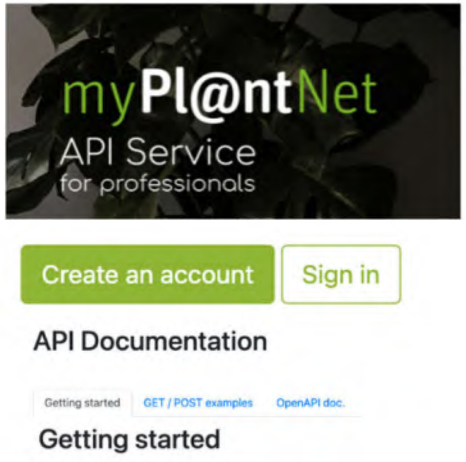
Education, animation

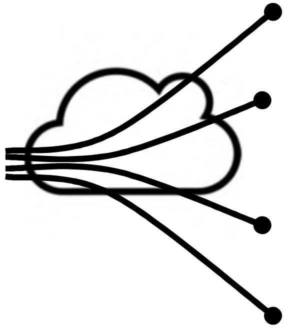
Tourism

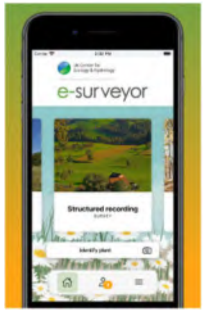
Trade

Pl@ntNet API - https://my.plantnet.org/

- A secured API providing developers programmatic access to Pl@ntNet engine
- **9K developer accounts** (companies, researchers, citizen observatories)
- Integrated in **European Open Science Cloud** (EOSC)



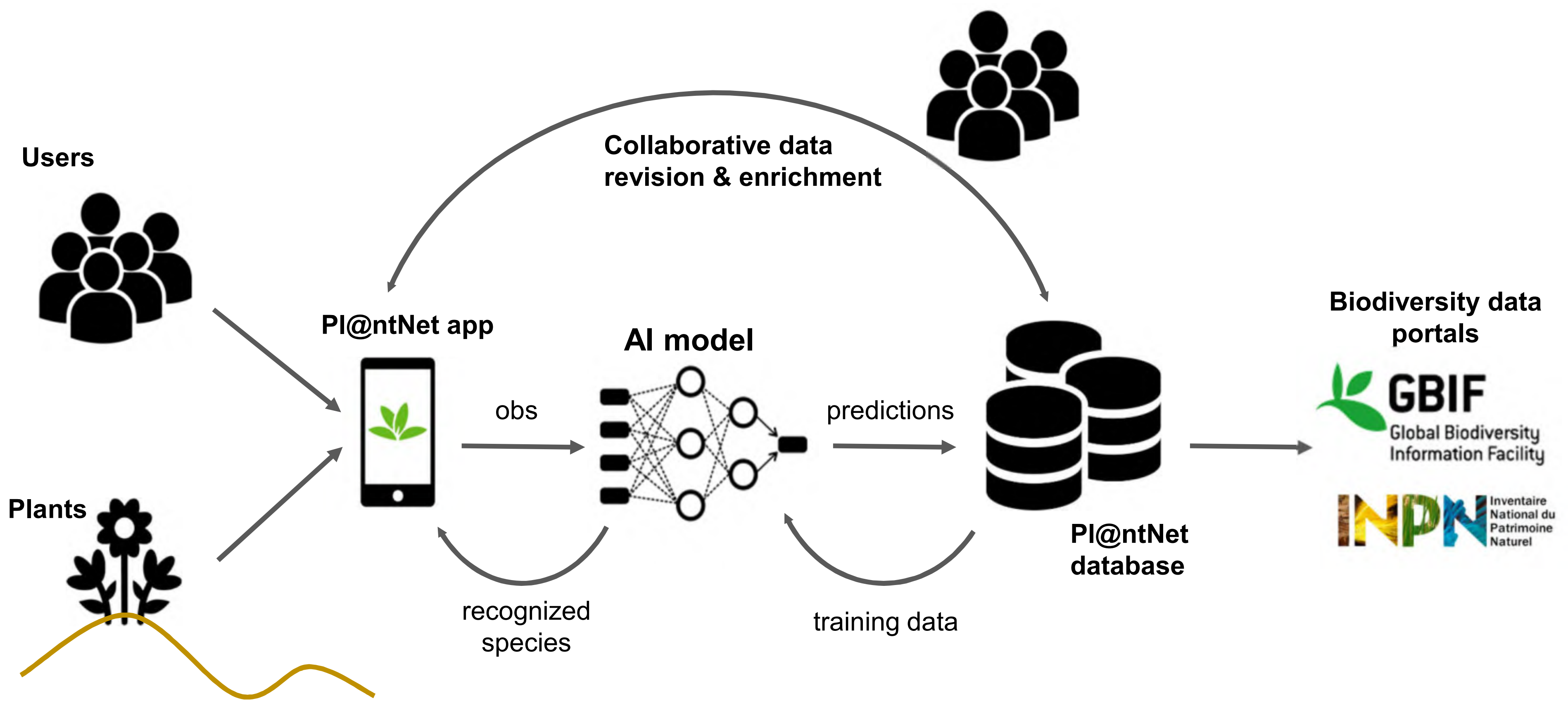




An iNaturalist-Pl@ntNet-workflow to identify plant-pollinator interactions – a case study of *Isodontia mexicana*

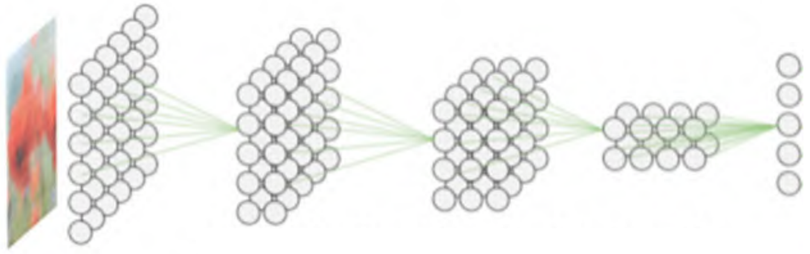
Nadja Pernat[1], Tom August[2], Quentin Groom[3], Daniyar Memedemin[4], and Lien Reyserhove[5]

Key concept of Pl@ntNet: Collaborative AI



Pl@ntNet AI model

Multi-head model trained on **Jean Zay super-computer** on a big dataset of **8M valid observations** (5-6 days of training)



$\sigma_{species}(f(x))$ **= 50K classes**

$\sigma_{organs}(f(x))$

$\sigma_{genus}(f(x))$

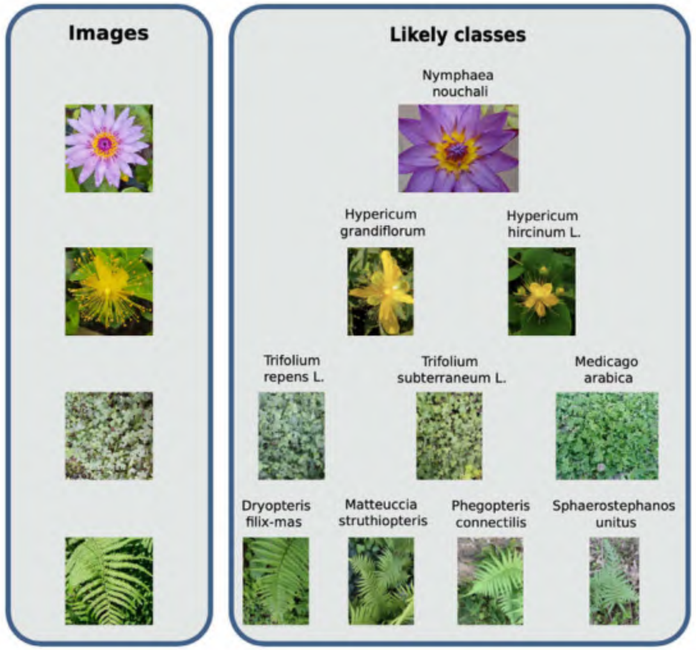
$\sigma_{disease}(f(x))$

Model = Vision transformer DinoV2

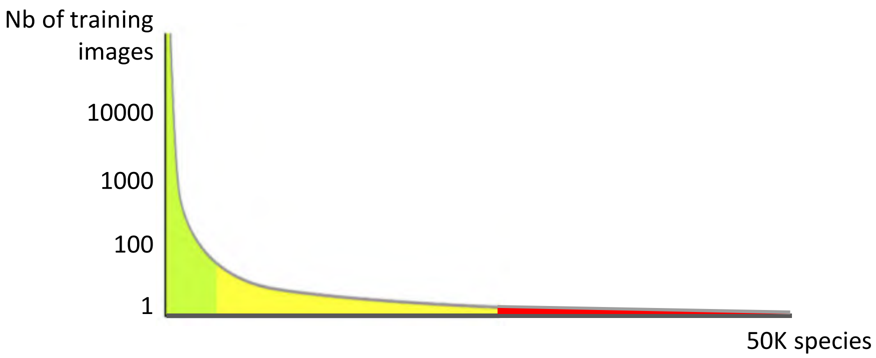
- Backbone pre-trained on 100M images using SSL (by Meta/Inria)
- Final multi-head model fine-tuned on 8M Pl@ntNet images (by Pl@ntNet team)

A difficult problem: uncertainty

Irreducible uncertainty
Species ambiguity



Model uncertainty
Increased by long-tail distribution



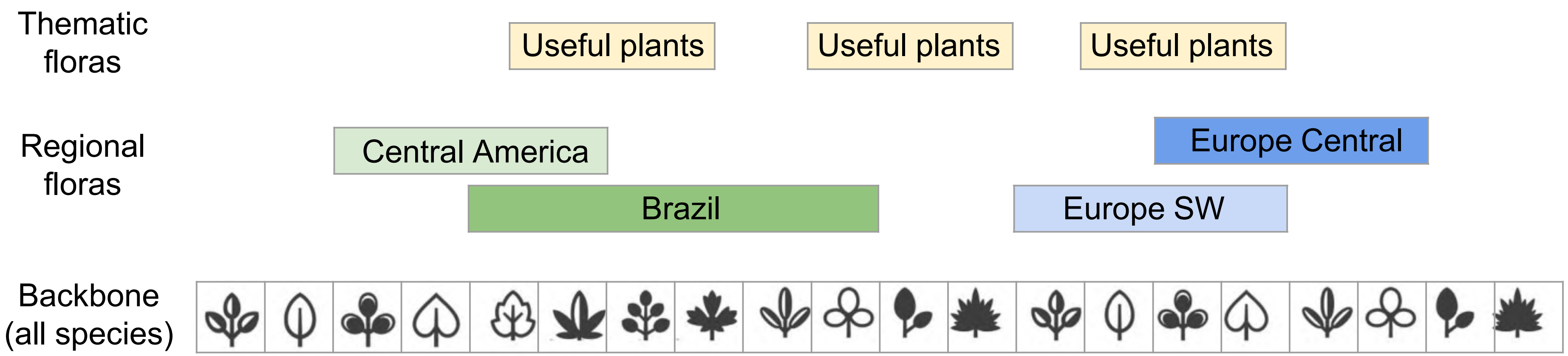
Top1 Identification accuracy:

Common species = ~**90%**
Average species = ~**70%**
Rare species = ~**40%**

Use of regional or thematic floras

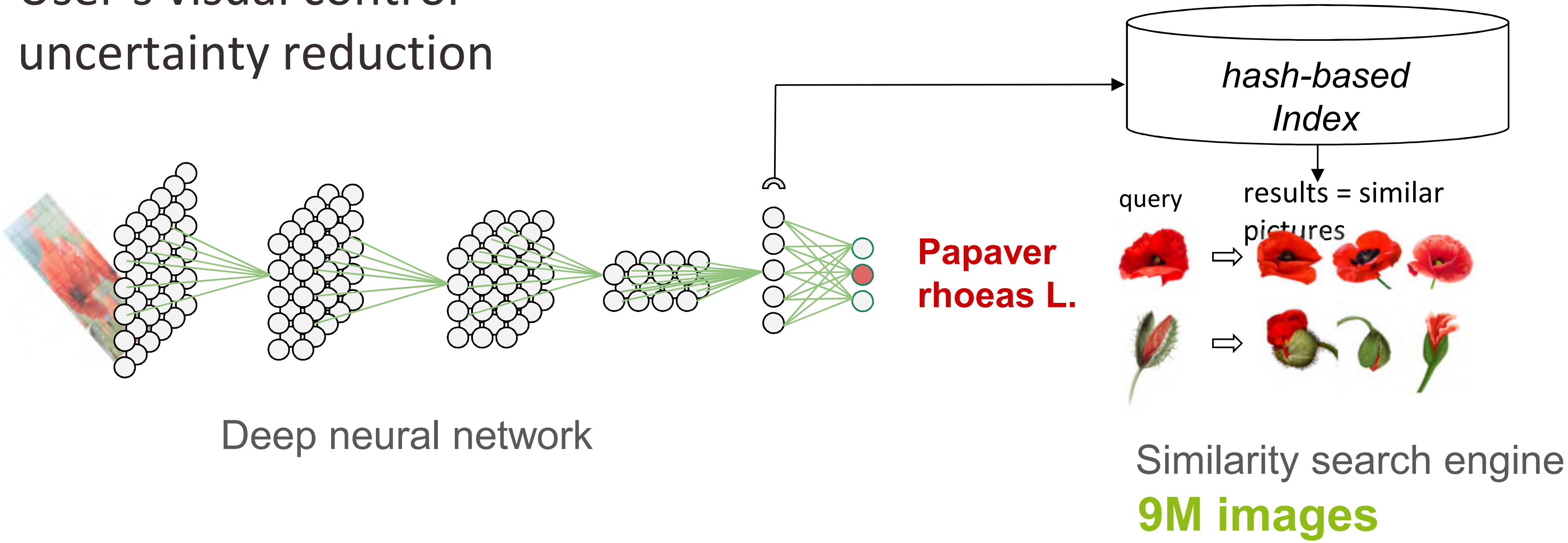
Restricting the hypothesis space to a particular flora allows improving the identification accuracy (based on World Checklist of Vascular Plants (WCVP)).

$$p(\underbrace{y}_{\text{species}}|\underbrace{x}_{\text{image}}, flora) \geq p(\underbrace{y}_{\text{species}}|\underbrace{x}_{\text{image}})$$



Pl@ntNet Similarity search

User's visual control =
uncertainty reduction

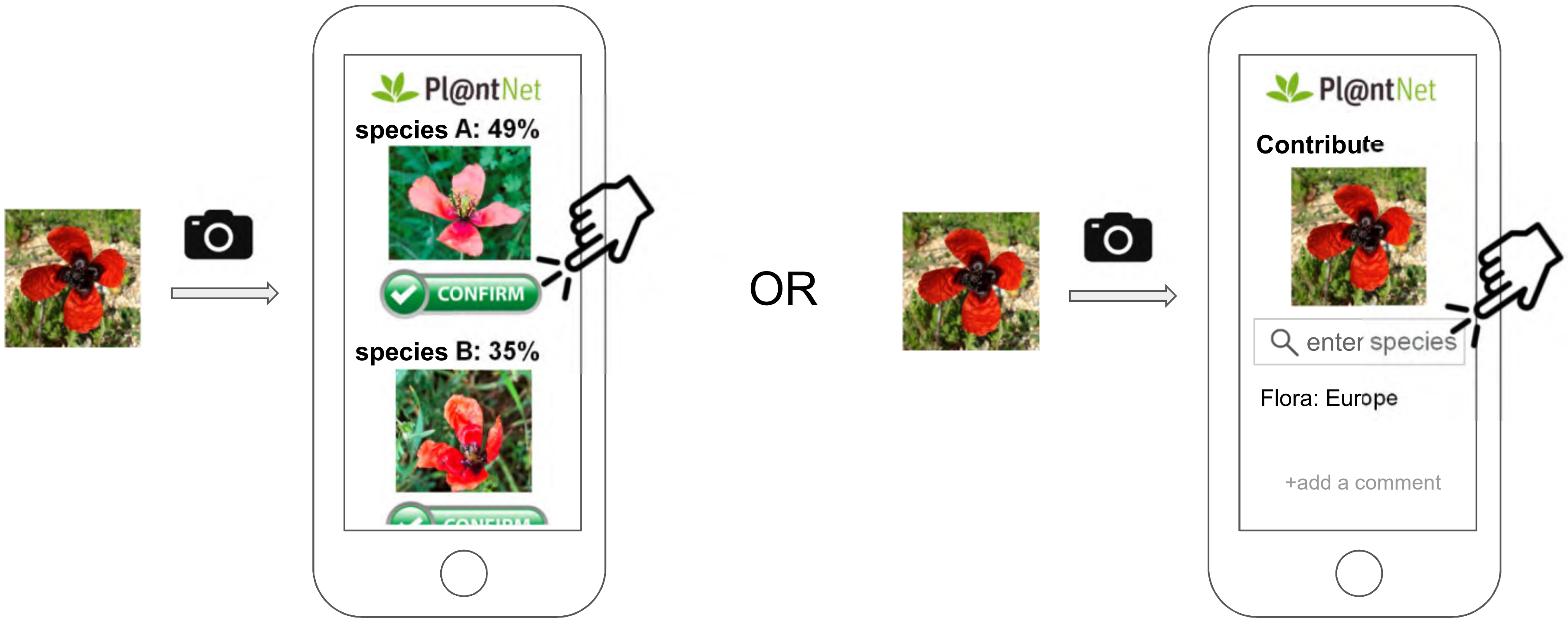


→ Sub-linear algorithm based on locality sensitive hashing

Joly, A., & Buisson, O. (2011, June). Random maximum margin hashing. In CVPR 2011 (pp. 873-880). IEEE.

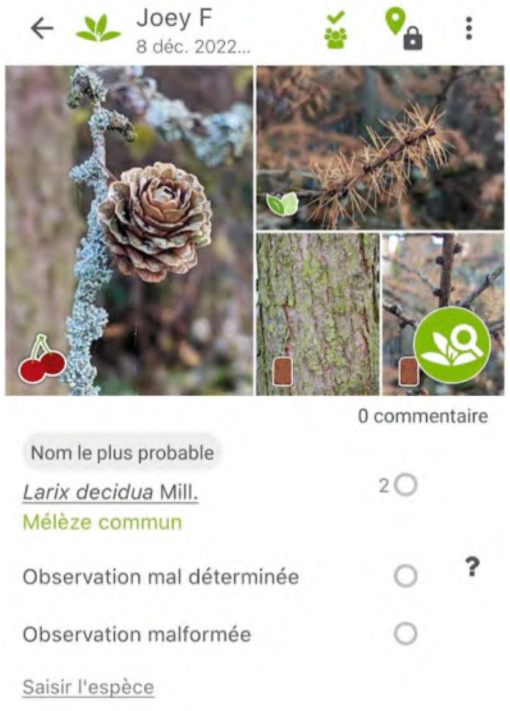
User’s contributions

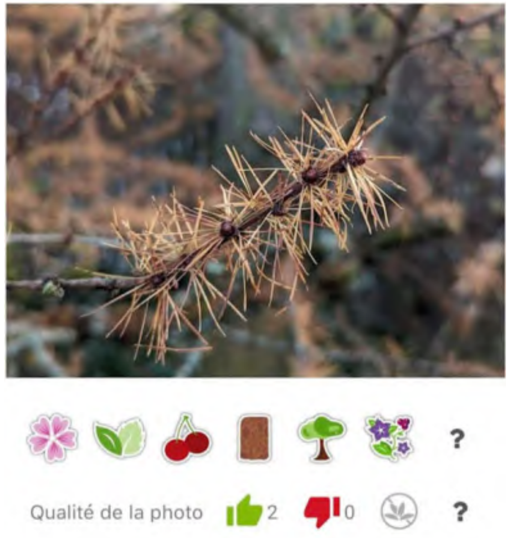
Users can contribute their observations



User’s revisions

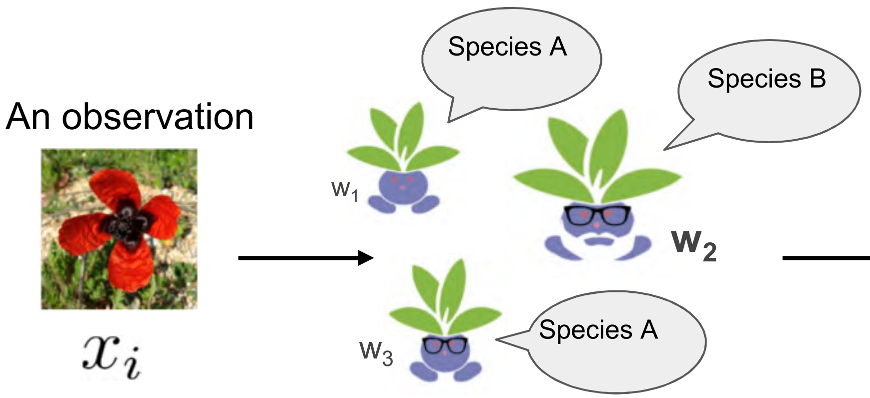
Users can revise observations of other users.





Cooperative Learning algorithm

The most probable label of an observation is determined with a weighted majority voting rule:



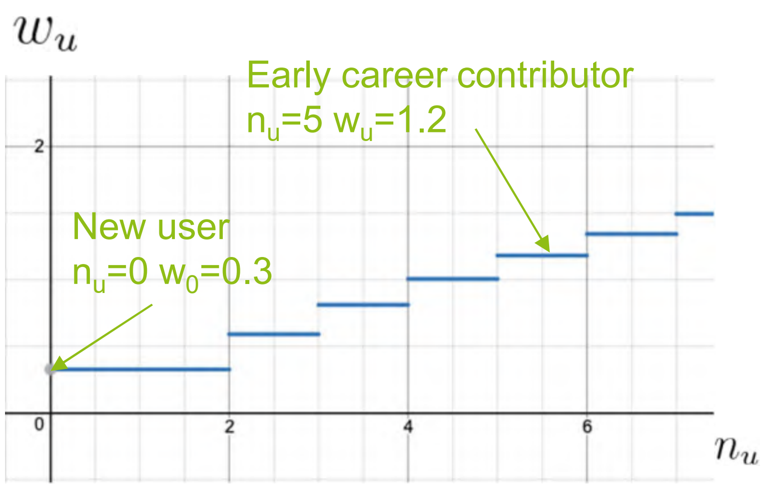
$$y_i = \arg\max_k \sum_{u \in U_i} w_u \mathbf{1}(y_i^u = k)$$

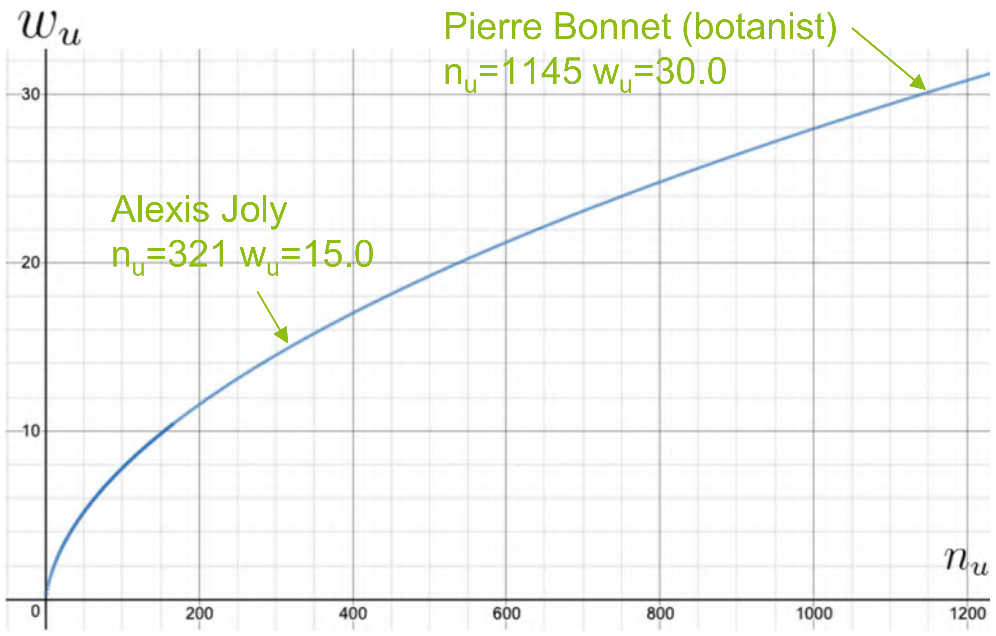
$U_i =$ Set of users who provided a label y_i^u for the observation x_i

Cooperative Learning algorithm

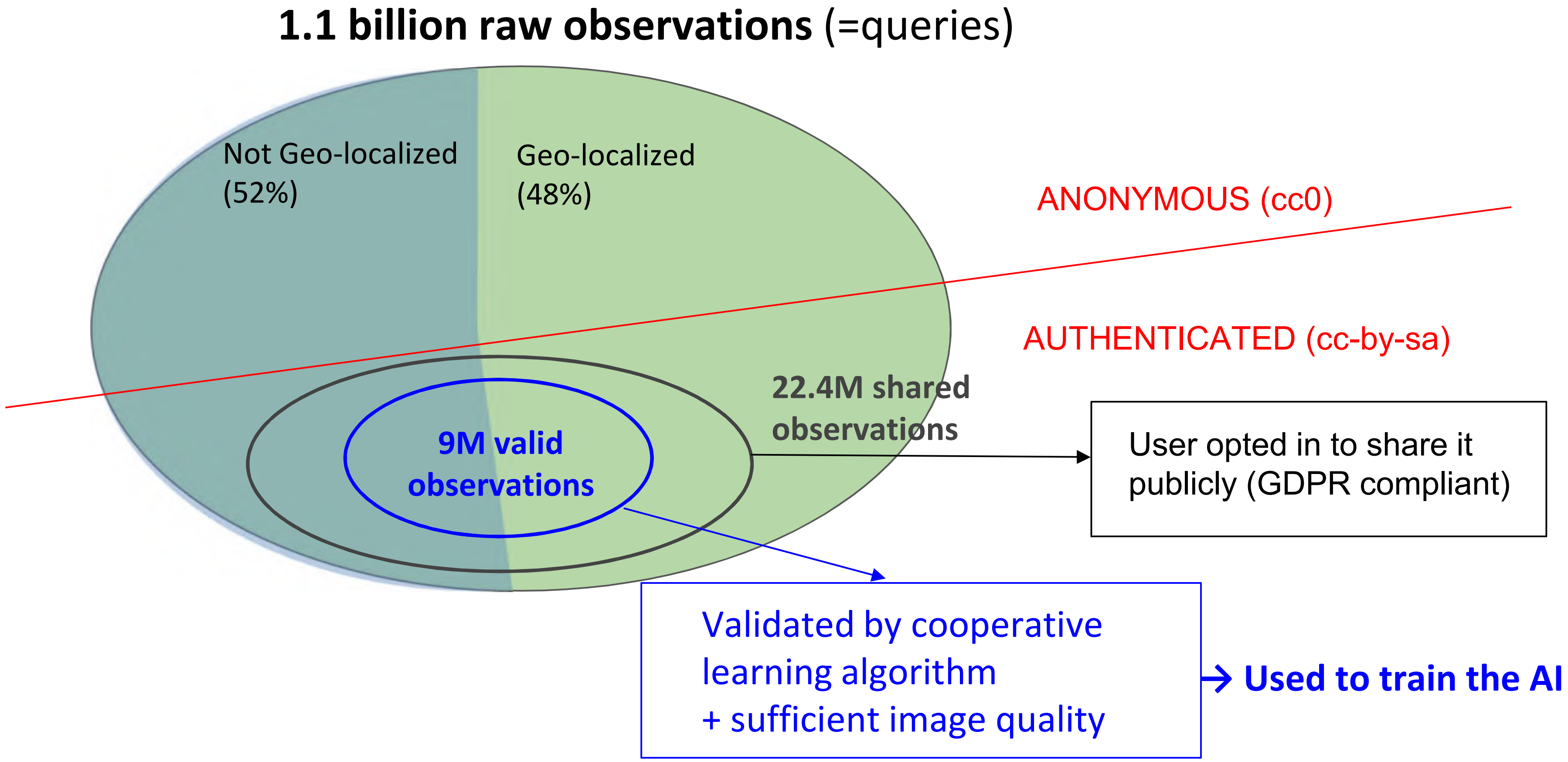
The weight of a user in Pl@ntNet is a function of the **estimated number of species** he is able to identify

$$w_u = g(n_u) \qquad n_u = \left|\{j : \exists i \; y_i^u = y_i\}\right|$$

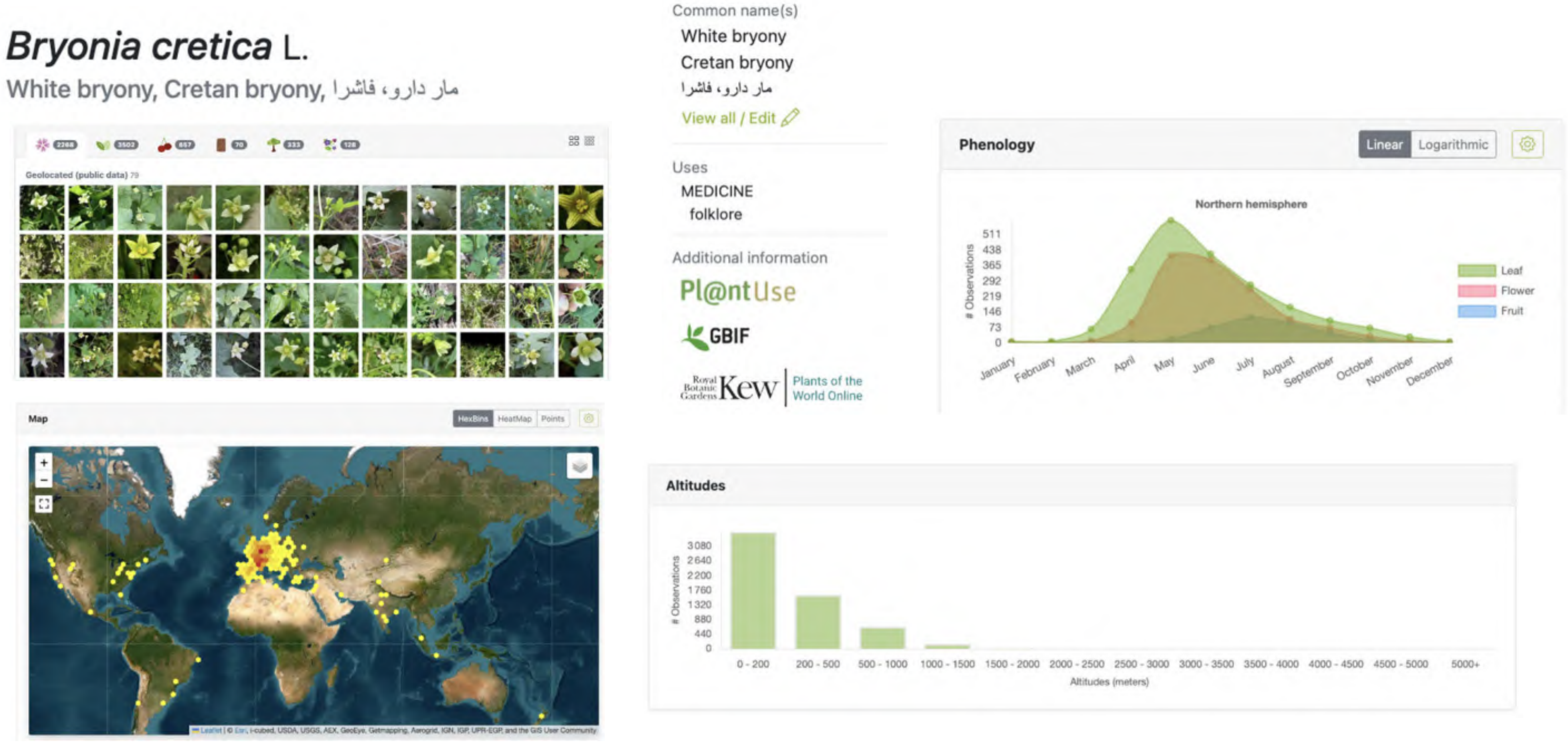




Pl@ntNet Data



Pl@ntNet Data visualisation tools



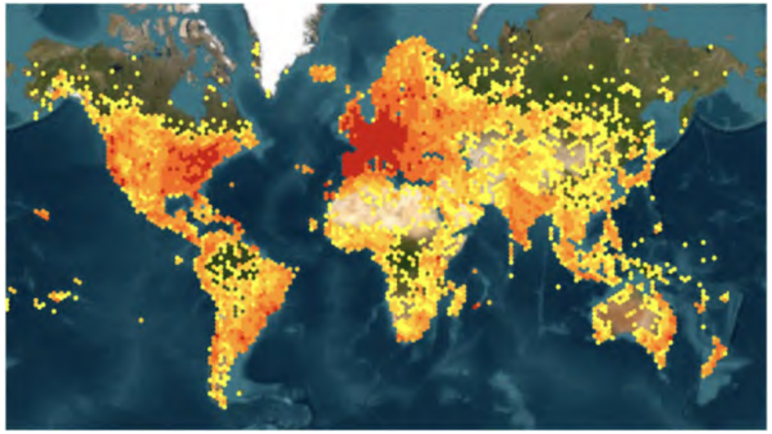
Pl@ntNet Data shared in GBIF

Top-2 plant data provider to GBIF (world's largest infrastructure for biodiversity data)

- Shared data = revised observations + trusted queries identified by the AI (AI score>0.95)
- Quality filters: potted & cultivated plants removal, region-based filtering (Kew POWO)

GBIF 13 856 500 OCCURRENCES

(87% identified by AI, 13% by humans)



https://doi.org/10.15468/mma2ec

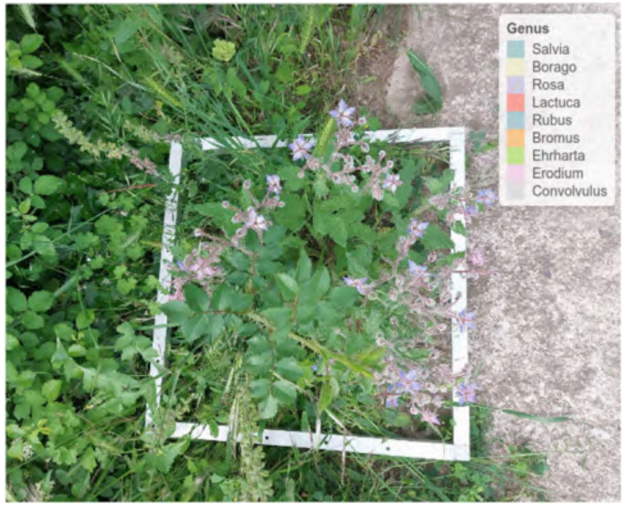
755 citations

RED LIST · nature · PLOS ONE · ANNALS OF BOTANY · WILEY Publishers Since 1807 · ELSEVIER

PART II
From individual plants to plant communities monitoring

Multi-specimen images for community-level monitoring

- Quadrat images for the monitoring of vulnerable habitats or fields biodiversity (e.g. VigieFlore)
- Vegetation cover images (e.g. terrestrial robots, drones, smartphones)
- Landscape views (e.g. car views for the monitoring of invasive species)



Weakly-supervised multi-label classification

Training data

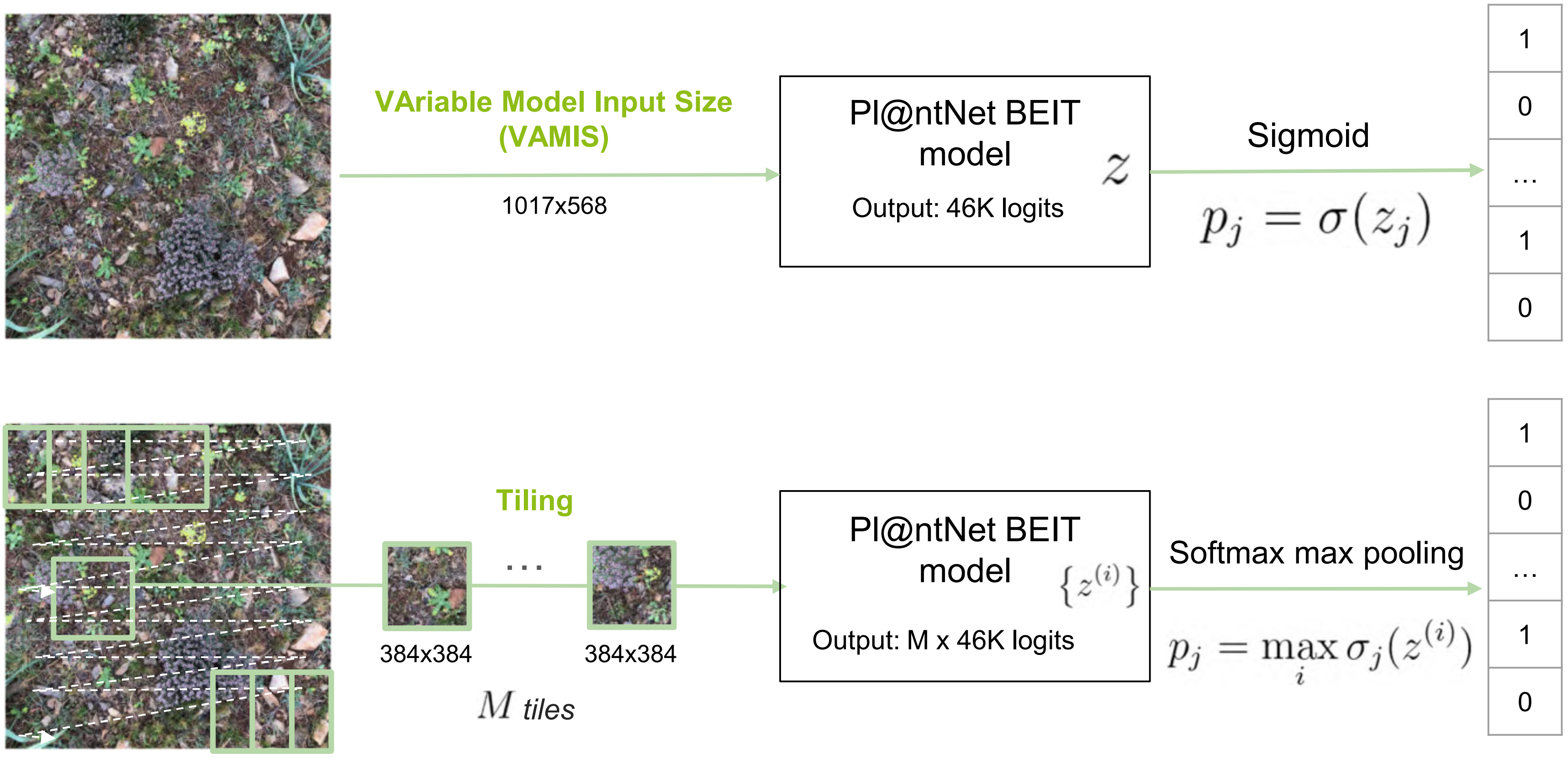
1

Pl@ntNet database

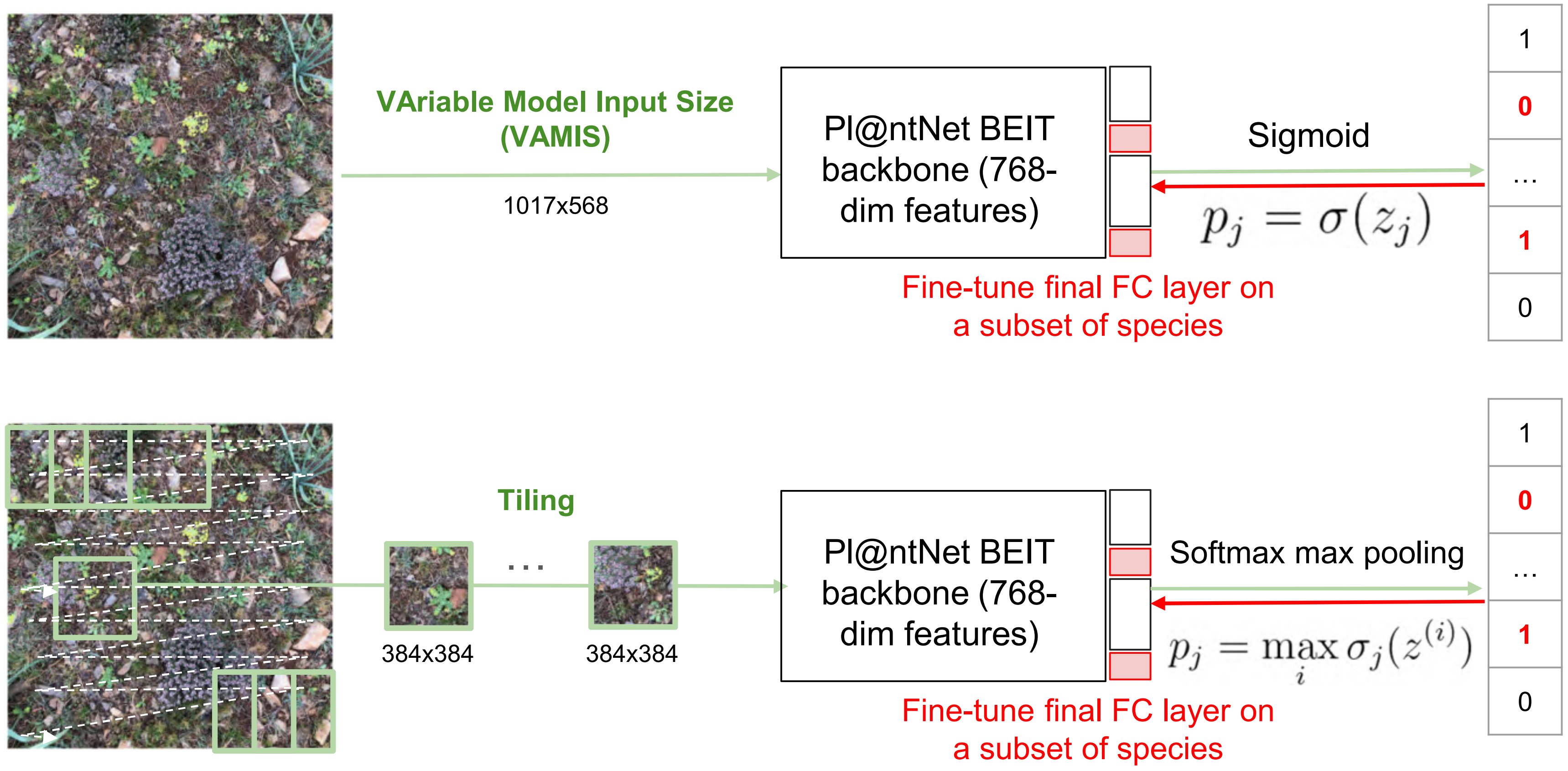
Test data

0	1	0	1	0	1	1	0

Zero-shot multi-label classification (no fine-tuning)



Few-shot multi-label classification (with fine-tuning)



Weakly-supervised multi-label classification

Evaluation on Danish road dataset

- Seven invasive species annotated
- 8.4K images with 1 to 3 invasive species

Dyrmann, M., Mortensen, A. K., Linneberg, L., Høye, T. T., & Bjerge, K. (2021). Camera assisted roadside monitoring for invasive alien plant species using deep learning. *Sensors*, *21*(18), 6126.

(a) (b)

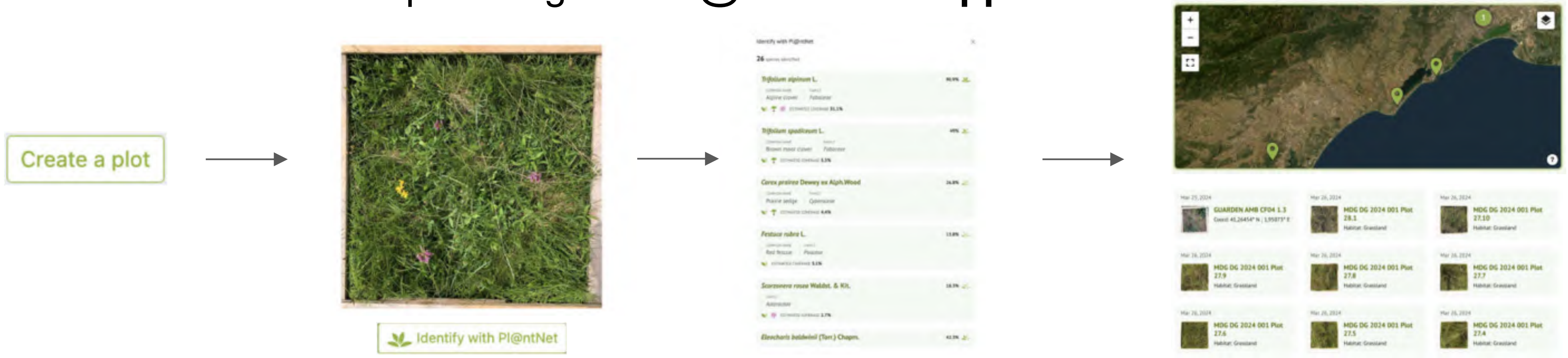
Results

	Zero-shot (no fine-tuning)		With fine-tuning	
	VAMIS	Tiling	VAMIS	Tiling
AUC	75.52	**91.58**	96.49	**96.50**
F1	36.45	**63.39**	74.28	**76.46**

Integration in Pl@ntNet

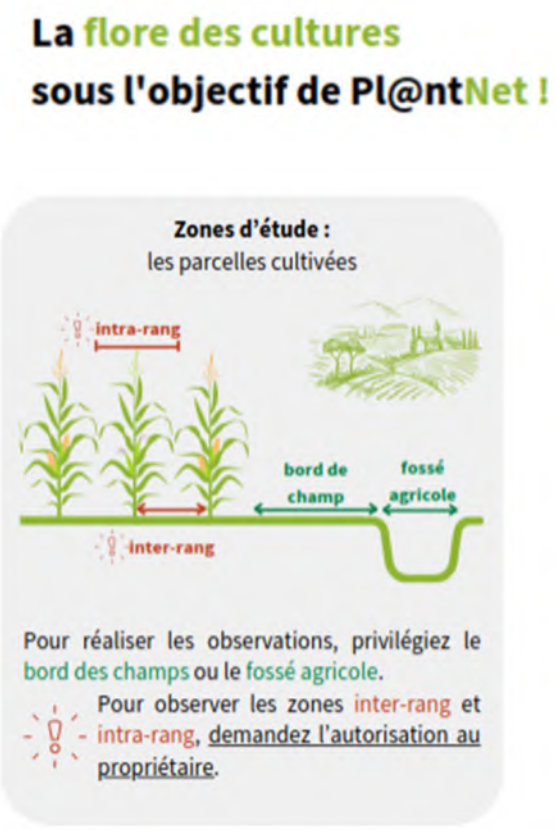
Tiling approach integrated in Pl@ntNet (without fine-tuning)

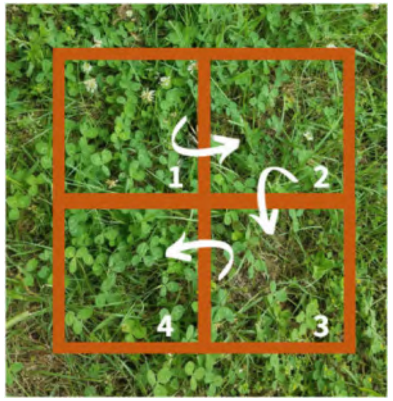
- **API** (my.plantnet.org)
  - used by biodiversa+ project on the monitoring of invasive alien species
  - used for our participation to Xprize rainforest challenge (Brazilian team, finalist)
- **GUI** dedicated to plot images in **Pl@ntNet web app**



Citizen science programme co-organized with

https://www.tela-botanica.org/projets/la-flore-sous-lobjectif-plantnet/






Matériels

Un carré représentant le quadrat de 50cm de côté

N'hésitez pas à utiliser une corde assez épaisse et des piquets bois

Un appareil photo

Veillez à
- *charger la batterie de votre appareil*
- *activer le GPS sur votre smartphone*
- *désactiver le flash s'il a plu ou que la végétation est humide*

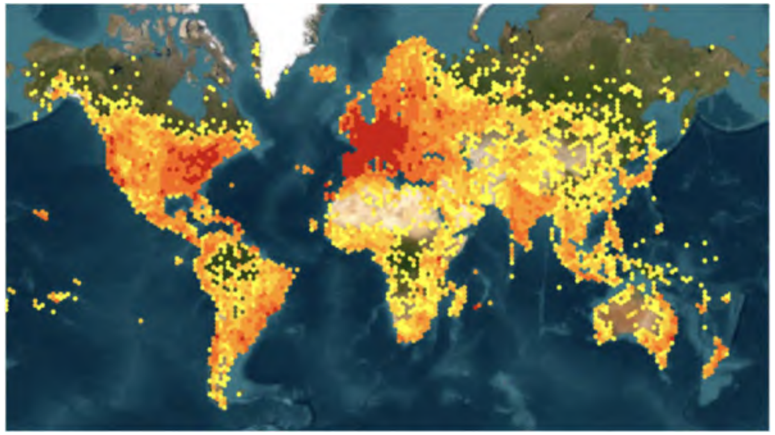
Une fiche de terrain imprimée et/ou carnet de terrain et un stylo

PART III
GeoPl@ntNet: from field observations to Mapping tools and Decision support applications

Objective: high resolution species and biodiversity indicators maps

Raw species occurrence data needs to be interpolated in space and time:

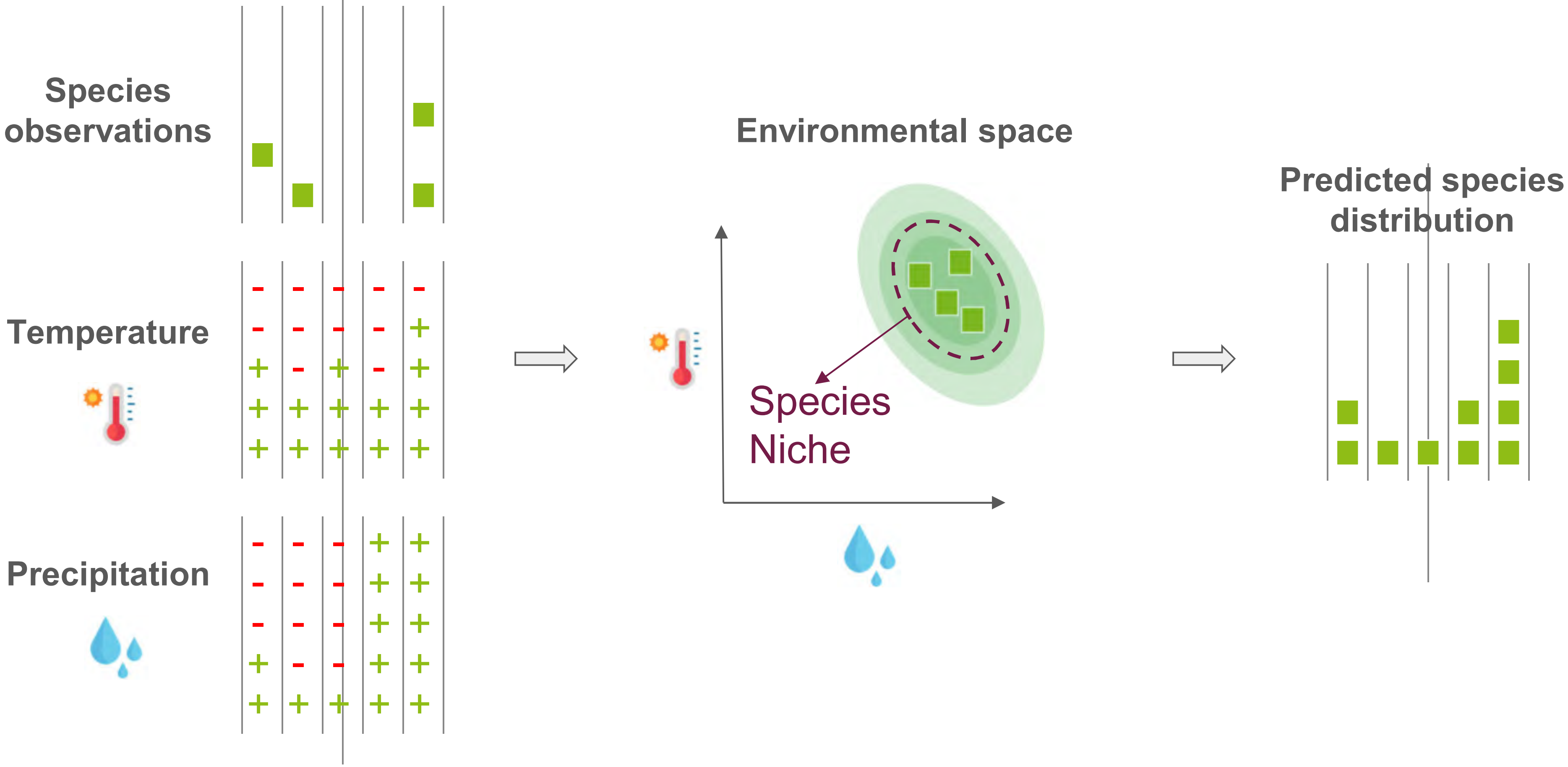
Many plant occurrences at world scale



GBIF

But very few locally (and biased) for most species

Species Distribution Models (SDM)



Remote sensing based SDM (DeepSDMs)

PLOS COMPUTATIONAL BIOLOGY

Convolutional neural networks improve species distribution modelling by capturing the spatial structure of the environment

Benjamin Deneu, Maximilien Servajean, Pierre Bonnet, Christophe Botella, François Munoz, Alexis Joly

frontiers in Plant Science

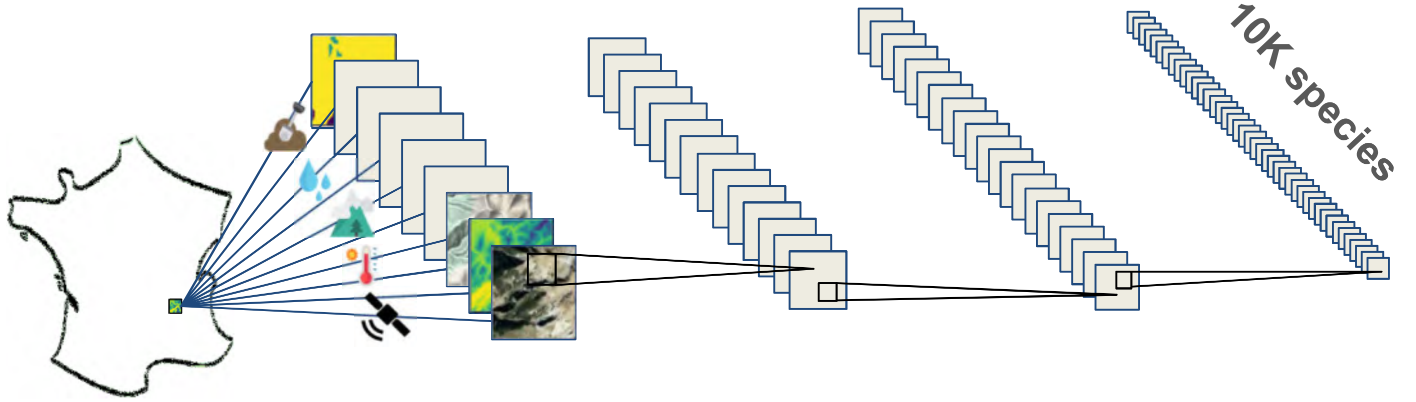
Deep Species Distribution Modeling From Sentinel-2 Image Time-Series: A Global Scale Analysis on the Orchid Family

Joaquim Estopinan, Maximilien Servajean, Pierre Bonnet, François Munoz, Alexis Joly

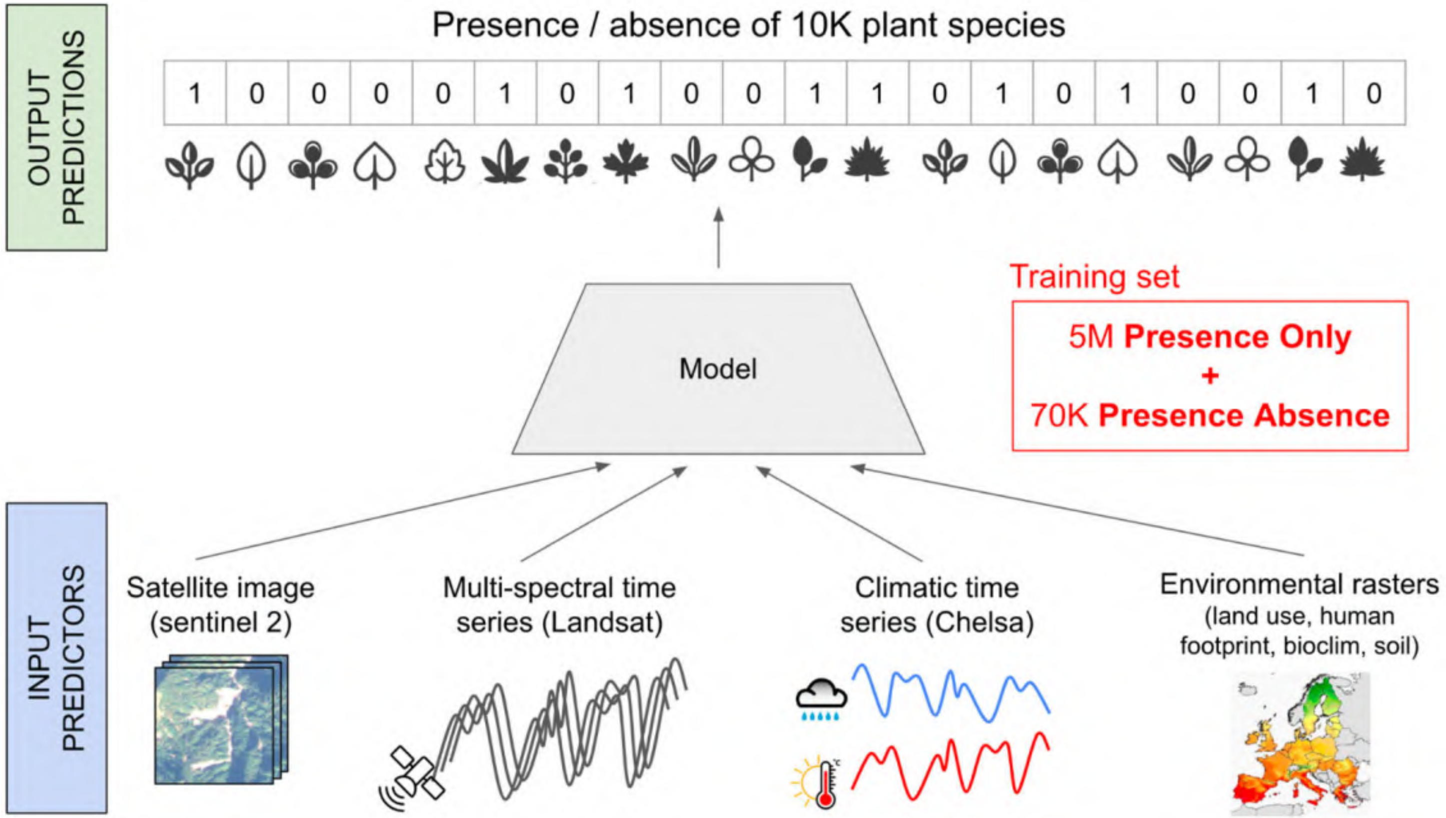
Model Input = data cubes

DeepSDM

Model Output = Suitability score of each species



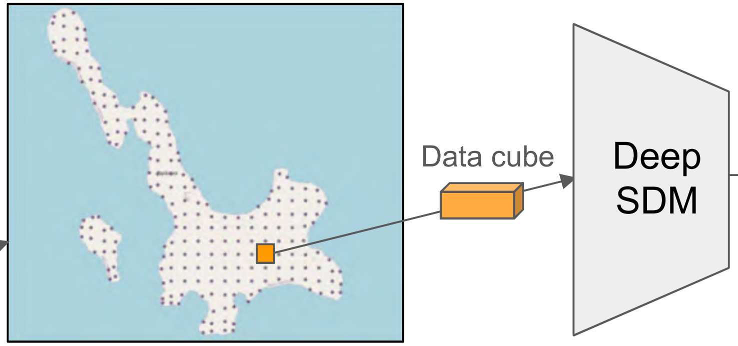
GeoLifeCLEF challenge 2023 & 2024



Integration in GeoPl@ntNet for EU-scale species mapping

25km x 25km grid

= 16K meta-tiles



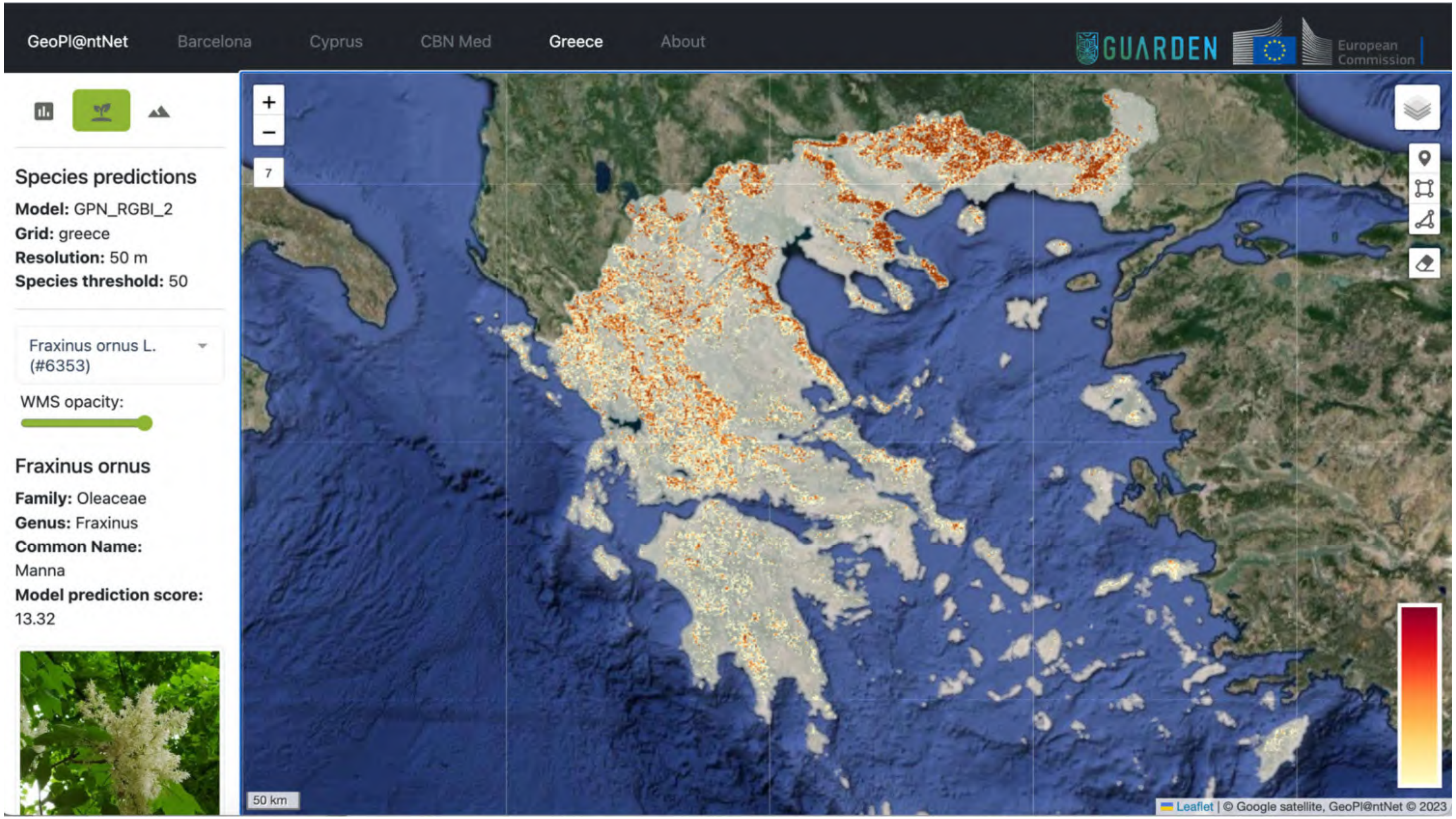
In each tile, points centered per 50 m2 cells

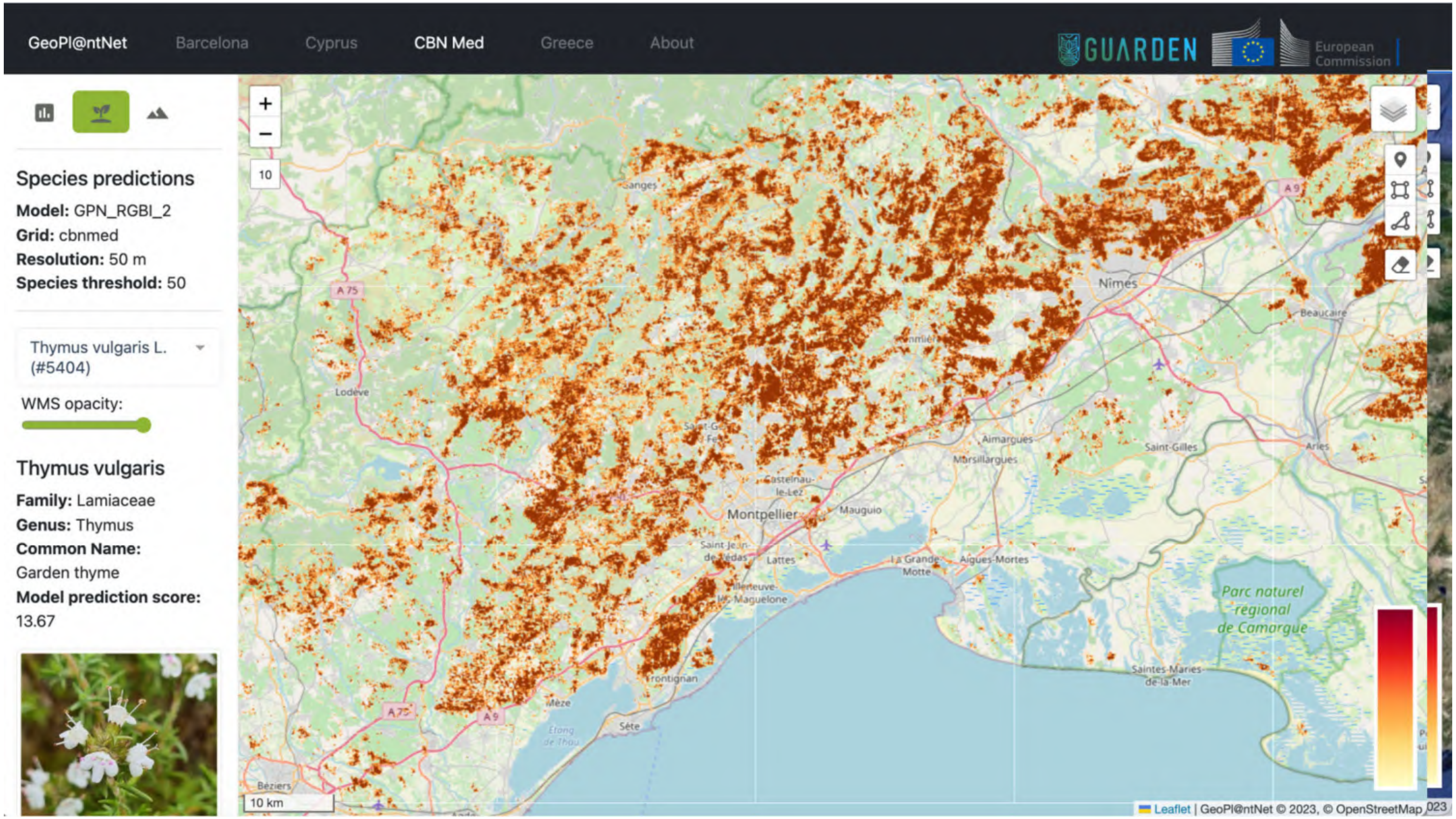
For each point, predict list of species present:

- species A
- species B
- species C

GPKG Species Coverage :
PostGIS Database ~400 Mo / meta-tile

Species Probabilities :
Tiff files ~650 Mo / meta tile


GeoPl@ntNet
Barcelona
Cyprus
CBN Med
Greece
About
GUARDEN
European Commission
Species predictions
Model: GPN_RGBI_2
Grid: greece
Resolution: 50 m
Species threshold: 50
Fraxinus ornus L. (#6353)
WMS opacity:
Fraxinus ornus
Family: Oleaceae
Genus: Fraxinus
Common Name: Manna
Model prediction score: 13.32
7
50 km
Leaflet | © Google satellite, GeoPl@ntNet © 2023


GeoPl@ntNet
Barcelona
Cyprus
CBN Med
Greece
About
GUARDEN
European Commission
Species predictions
Model: GPN_RGBI_2
Grid: cbnmed
Resolution: 50 m
Species threshold: 50
Thymus vulgaris L. (#5404)
WMS opacity:
Thymus vulgaris
Family: Lamiaceae
Genus: Thymus
Common Name: Garden thyme
Model prediction score: 13.67
Nîmes
Montpellier
Lodève
Beaucaire
Saint-Gilles
Arles
Aimargues
Marsillargues
Mauguio
Lattes
Aigues-Mortes
La Grande Motte
Parc naturel régional de Camargue
Saintes-Maries-de-la-Mer
Frontignan
Mèze
Sète
Étang de Thau
Béziers
10 km
Leaflet | GeoPl@ntNet © 2023, © OpenStreetMap

Pl@ntNet Identify Explore Contributions Groups GeoPl@ntNet BETA English alexis joly

GeoPl@ntNet

Anthemis maritima

Family: Asteraceae
Genus: Anthemis
Common Name:
Seaside Chamomile

Mapping biodiversity conservation indicators

From the species assemblage predicted at each point

$$S_\lambda(x) := \{k \in \mathcal{Y} : \hat{\eta}_k(x) > \lambda\}$$


ELSEVIER

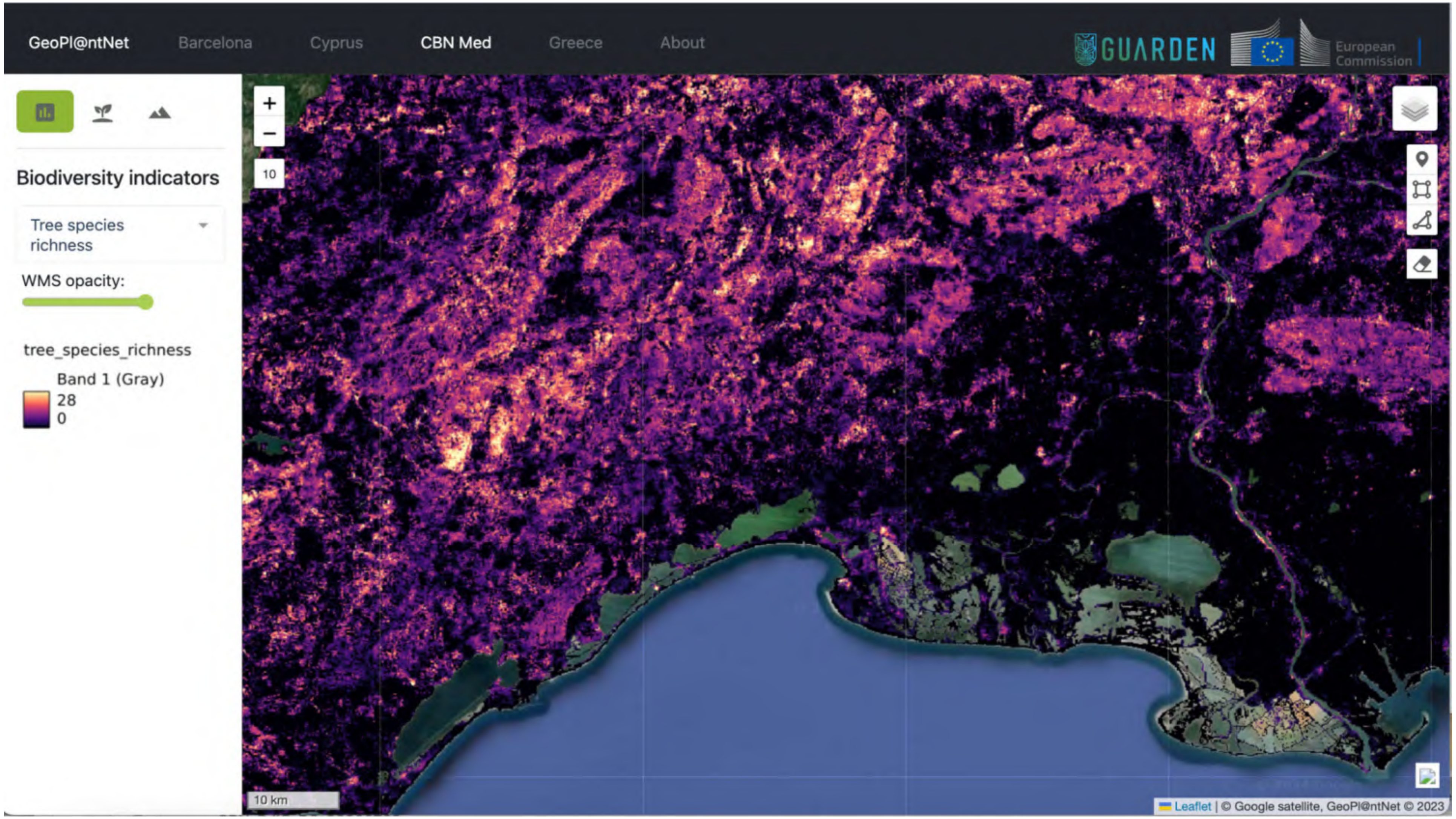
Ecological Informatics
Volume 81, July 2024, 102627

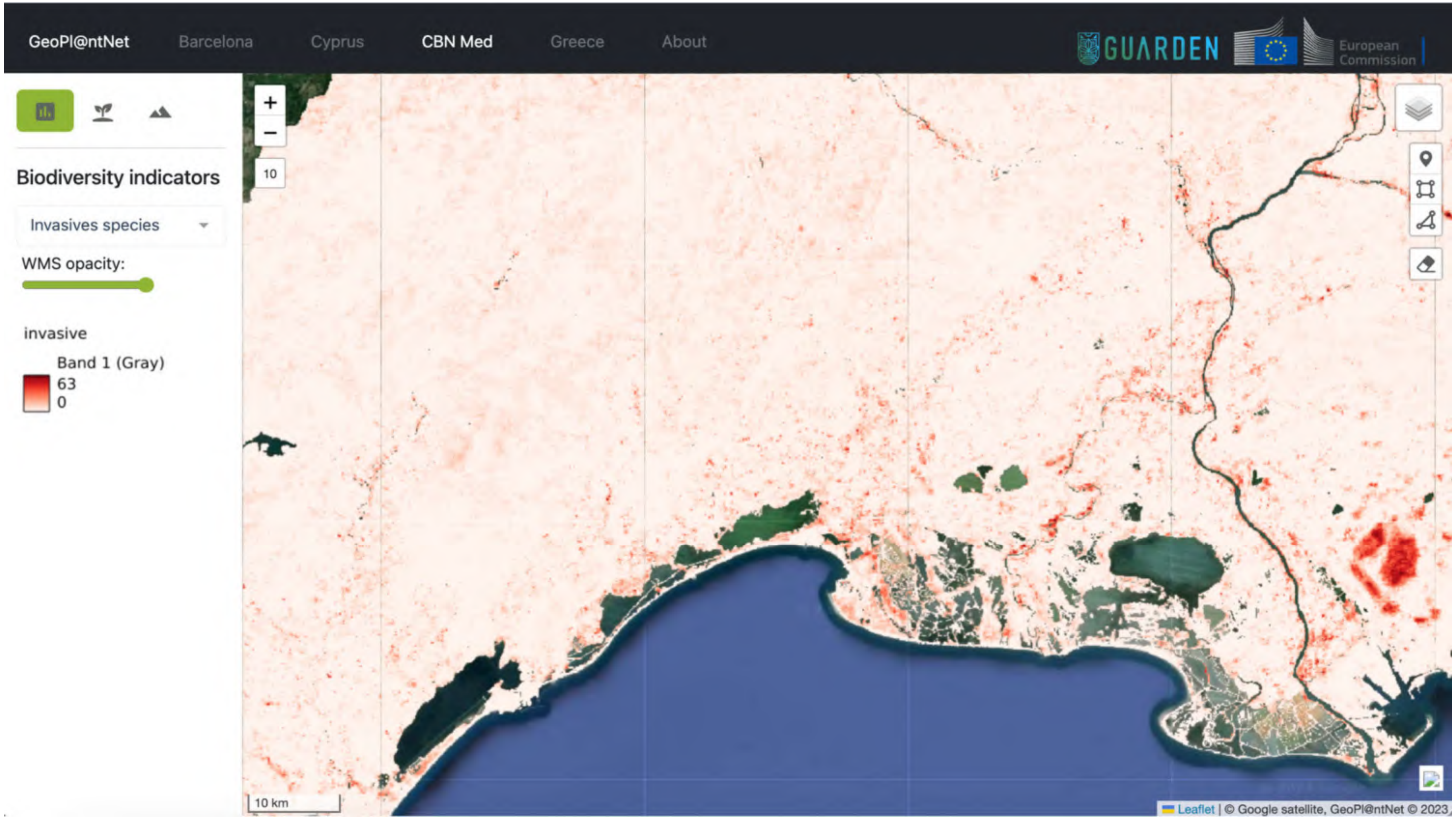
Mapping global orchid assemblages with deep learning provides novel conservation insights

We can compute indicators such as:

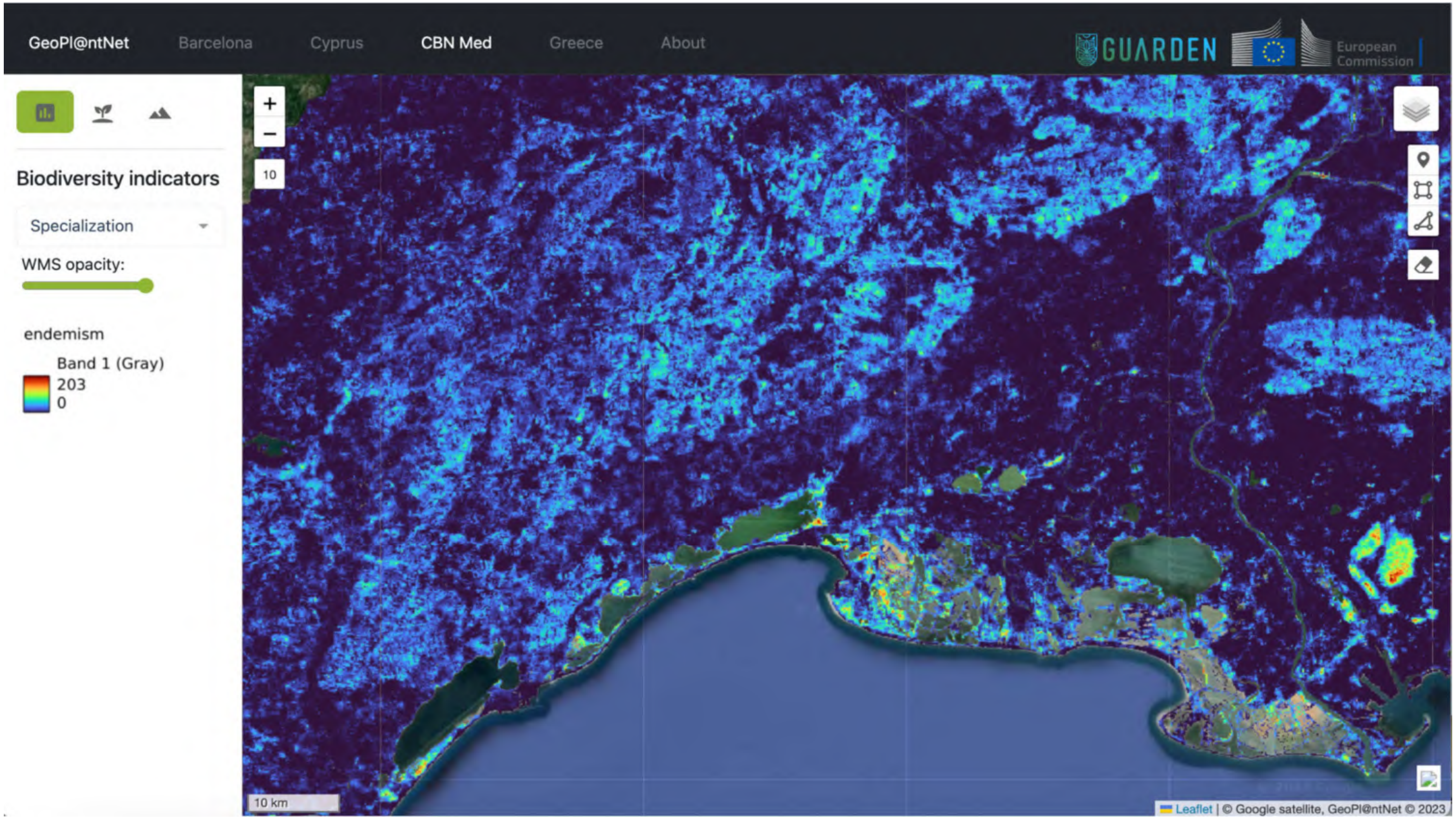
- The number of endangered species (e.g. on IUCN red list)
- The number of tree species (carbon capture)
- The diversity of species (e.g. Shannon index)
- The number of rare species, of species on EU Habitats directives
- The EUNIS habitat (using a species-to-habitat model)

We can construct maps of such indicators at very high resolution by computing $S_\lambda(x)$ for all x_i on a dense spatial grid

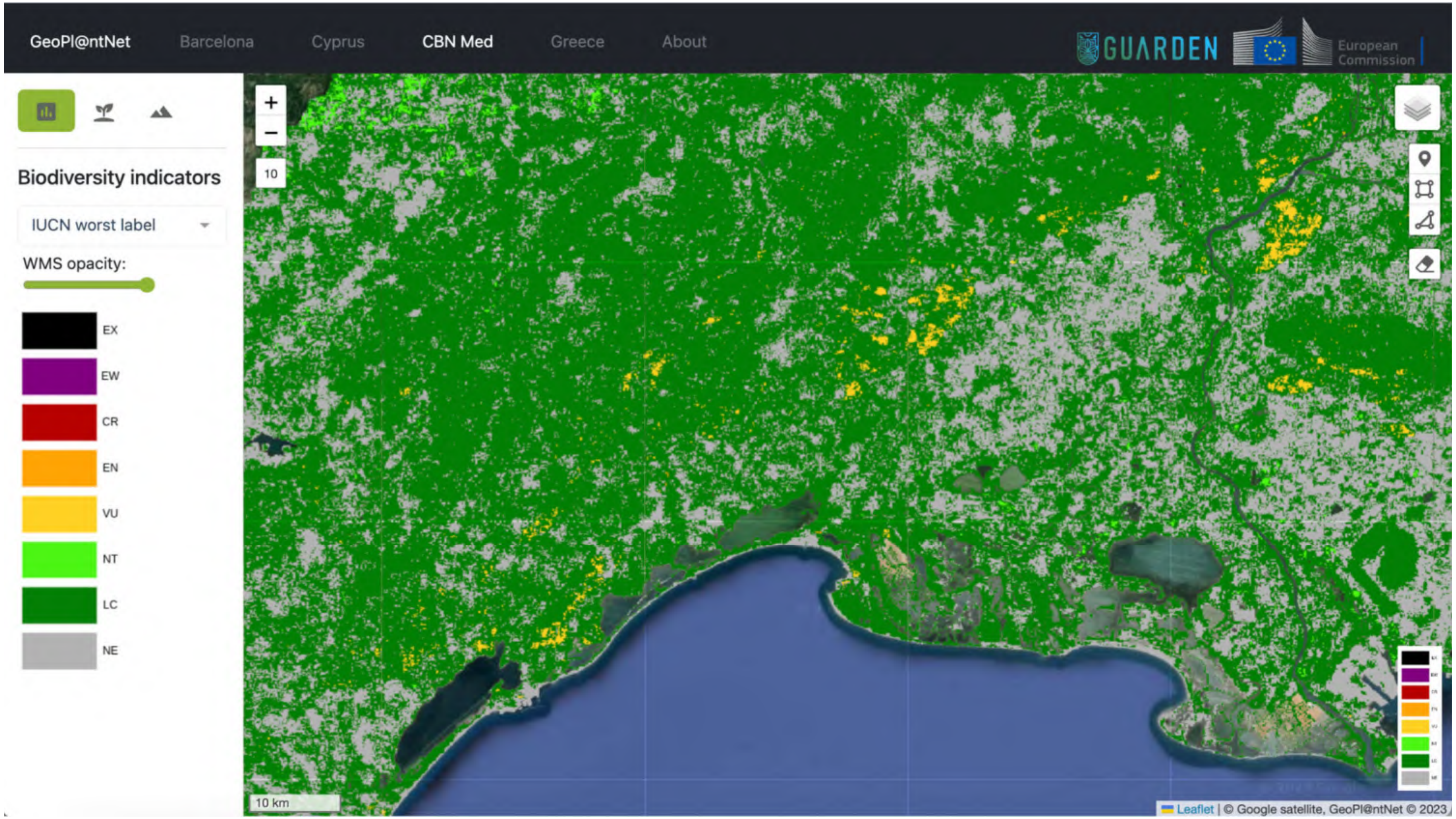

GeoPl@ntNet
Barcelona
Cyprus
CBN Med
Greece
About
GUARDEN
European Commission
Biodiversity indicators
Tree species richness
WMS opacity:
tree_species_richness
Band 1 (Gray)
28
0
10
10 km
Leaflet | © Google satellite, GeoPl@ntNet © 2023

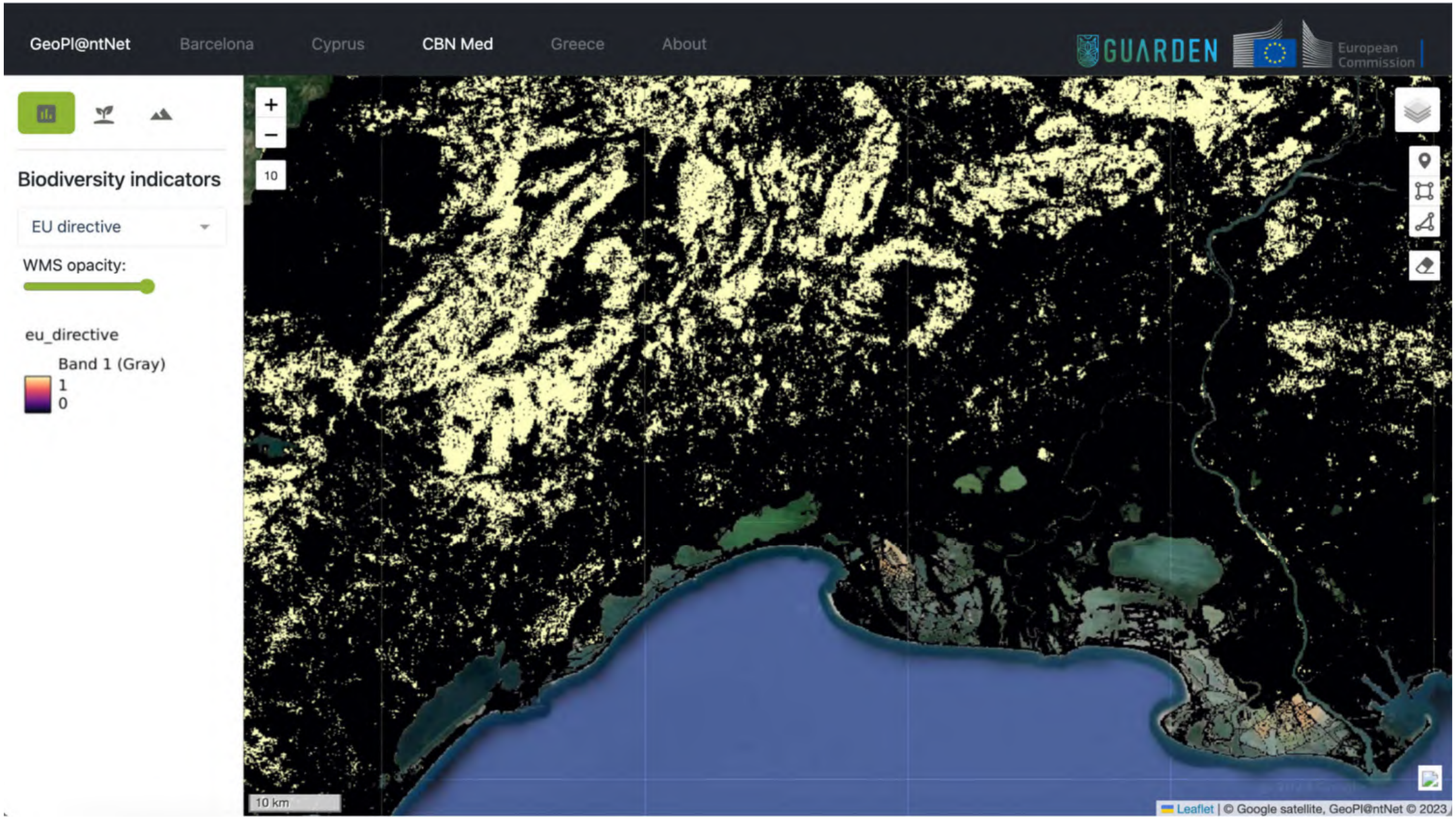


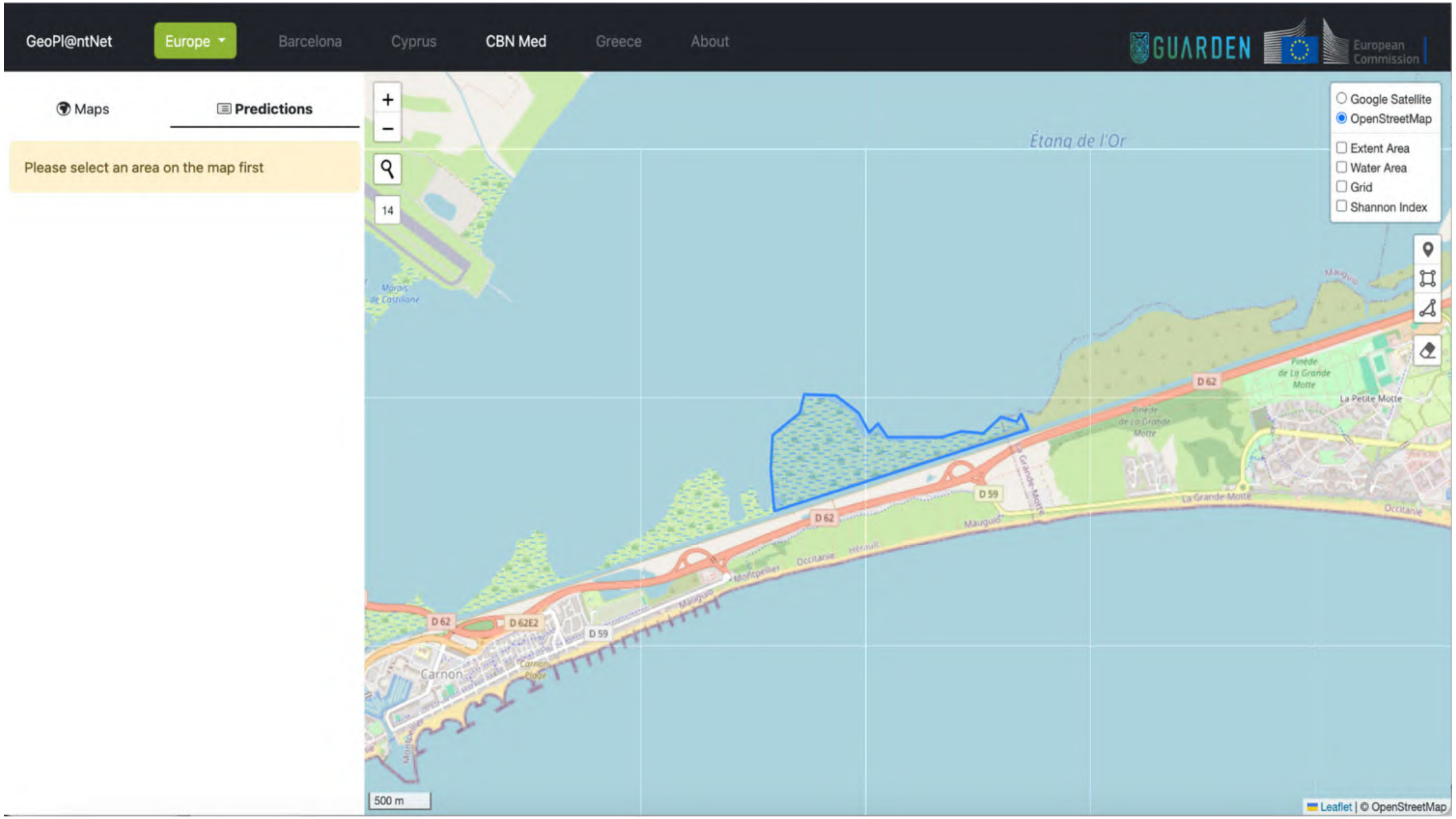
GeoPl@ntNet
Barcelona
Cyprus
CBN Med
Greece
About
GUARDEN
European Commission
Biodiversity indicators
Invasives species
WMS opacity:
invasive
Band 1 (Gray)
63
0
10
10 km
Leaflet | © Google satellite, GeoPl@ntNet © 2023

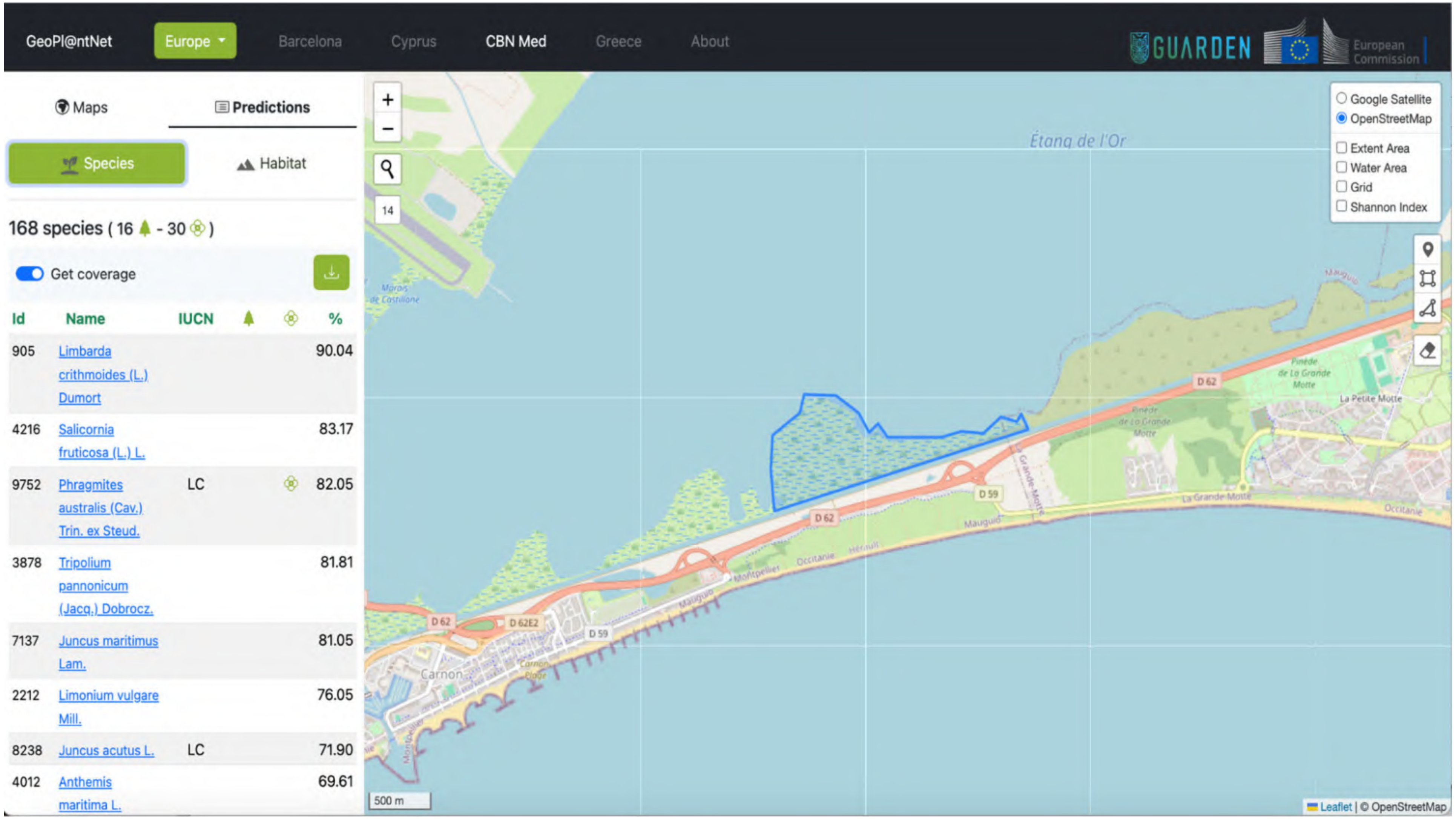


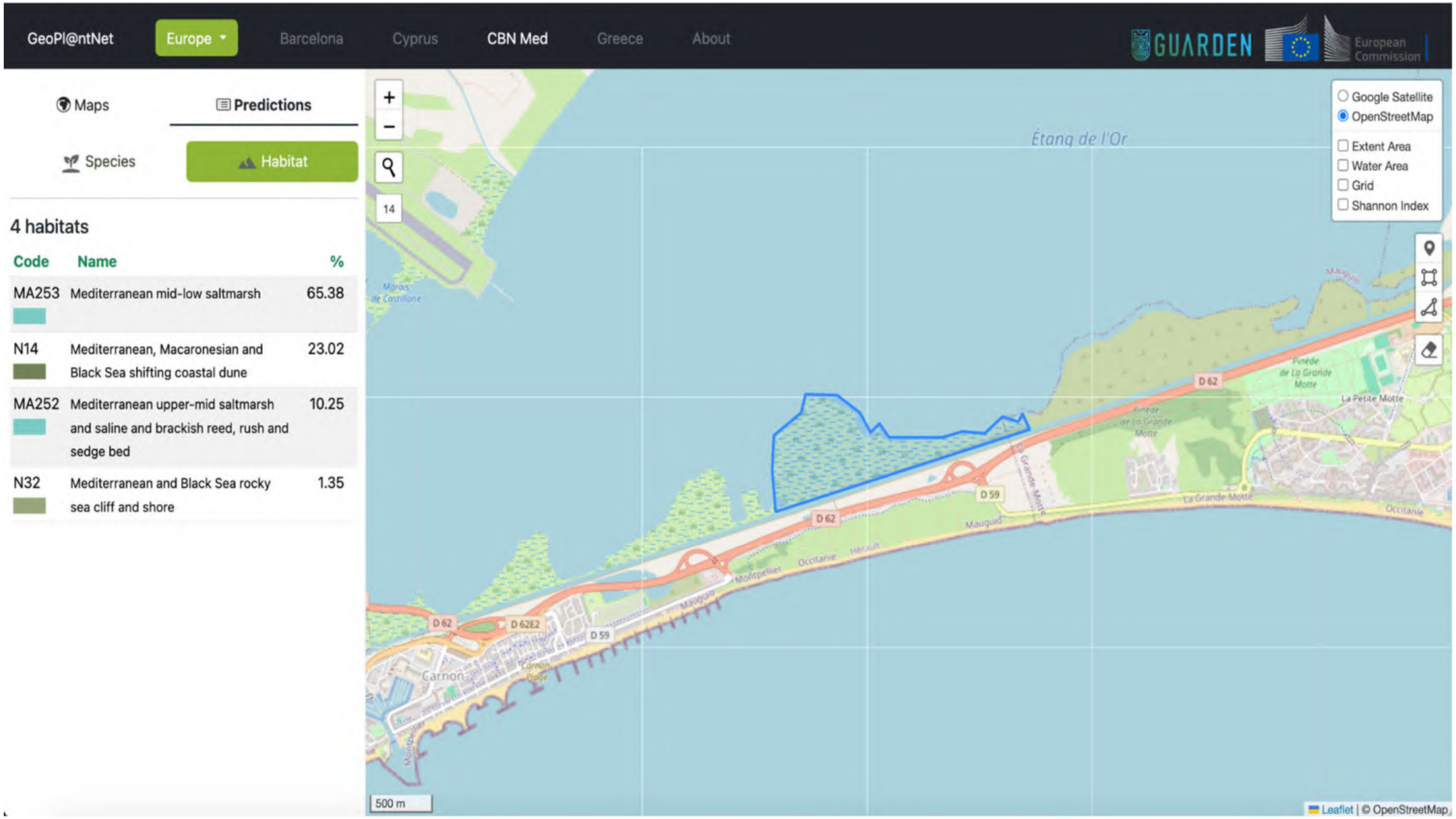
GeoPl@ntNet
Barcelona
Cyprus
CBN Med
Greece
About
GUARDEN
European Commission
Biodiversity indicators
Specialization
WMS opacity:
endemism
Band 1 (Gray)
203
0
10
10 km
Leaflet | © Google satellite, GeoPl@ntNet © 2023


GeoPl@ntNet
Barcelona
Cyprus
CBN Med
Greece
About
GUARDEN
European Commission
Biodiversity indicators
IUCN worst label
WMS opacity:
EX
EW
CR
EN
VU
NT
LC
NE
10
10 km
Leaflet | © Google satellite, GeoPl@ntNet © 2023

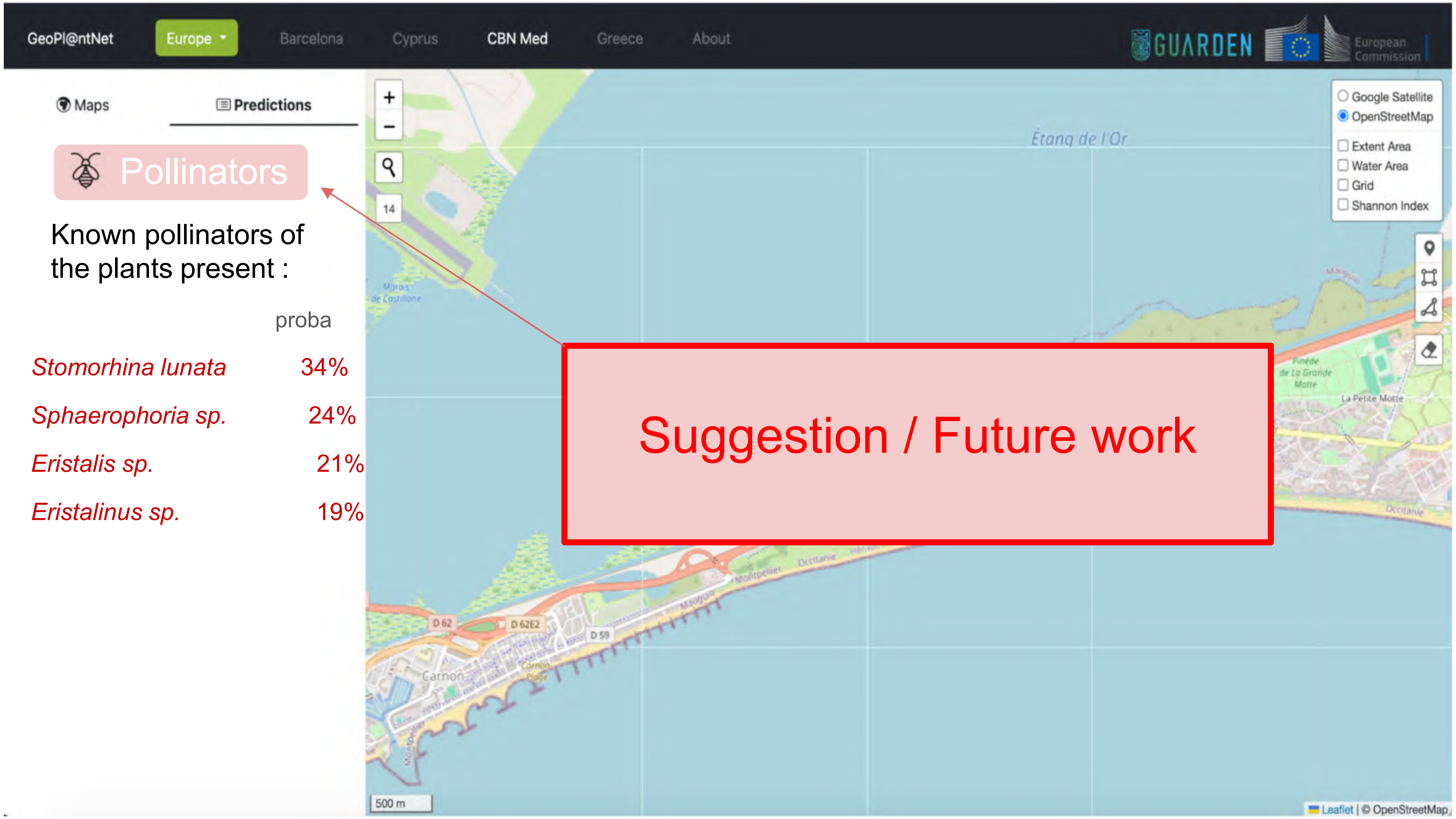

GeoPl@ntNet
Barcelona
Cyprus
CBN Med
Greece
About
GUARDEN
European Commission
Biodiversity indicators
EU directive
WMS opacity:
eu_directive
Band 1 (Gray)
1
0
10
10 km
Leaflet | © Google satellite, GeoPl@ntNet © 2023


GeoPl@ntNet
Europe
Barcelona
Cyprus
CBN Med
Greece
About
GUARDEN
European Commission
Maps
Predictions
Please select an area on the map first
14
Google Satellite
OpenStreetMap
Extent Area
Water Area
Grid
Shannon Index
Étang de l'Or
D 62
D 59
D 62E2
Carnon
La Grande Motte
La Petite Motte
500 m
Leaflet | © OpenStreetMap


GeoPl@ntNet
Europe
Barcelona
Cyprus
CBN Med
Greece
About
GUARDEN
European Commission
Maps
Predictions
Species
Habitat
168 species (16 - 30)
Get coverage
Id Name IUCN %
905 Limbarda crithmoides (L.) Dumort 90.04
4216 Salicornia fruticosa (L.) L. 83.17
9752 Phragmites australis (Cav.) Trin. ex Steud. LC 82.05
3878 Tripolium pannonicum (Jacq.) Dobrocz. 81.81
7137 Juncus maritimus Lam. 81.05
2212 Limonium vulgare Mill. 76.05
8238 Juncus acutus L. LC 71.90
4012 Anthemis maritima L. 69.61
Google Satellite
OpenStreetMap
Extent Area
Water Area
Grid
Shannon Index
Étang de l'Or
Carnon
La Petite Motte
D 62
D 59
500 m
Leaflet | © OpenStreetMap


GeoPl@ntNet
Europe
Barcelona
Cyprus
CBN Med
Greece
About
GUARDEN
European Commission
Maps
Predictions
Species
Habitat
4 habitats
Code
Name
%
MA253
Mediterranean mid-low saltmarsh
65.38
N14
Mediterranean, Macaronesian and Black Sea shifting coastal dune
23.02
MA252
Mediterranean upper-mid saltmarsh and saline and brackish reed, rush and sedge bed
10.25
N32
Mediterranean and Black Sea rocky sea cliff and shore
1.35
Google Satellite
OpenStreetMap
Extent Area
Water Area
Grid
Shannon Index
Étang de l'Or
14
D 62
D 59
D 62E2
Carnon
La Petite Motte
La Grande-Motte
500 m
Leaflet | © OpenStreetMap

Suggestion / Future work



Thank you

inria informatics mathematics

cirad

IRD
Institut de Recherche pour le Développement
FRANCE

INRAE

agropolis fondation

GUARDEN

MAMBO
MODERN APPROACHES TO THE MONITORING OF BIODIVERSITY

Pl@ntAgroEco

GENCI

L'AVENIR DEPUIS 1220
UNIVERSITÉ DE MONTPELLIER