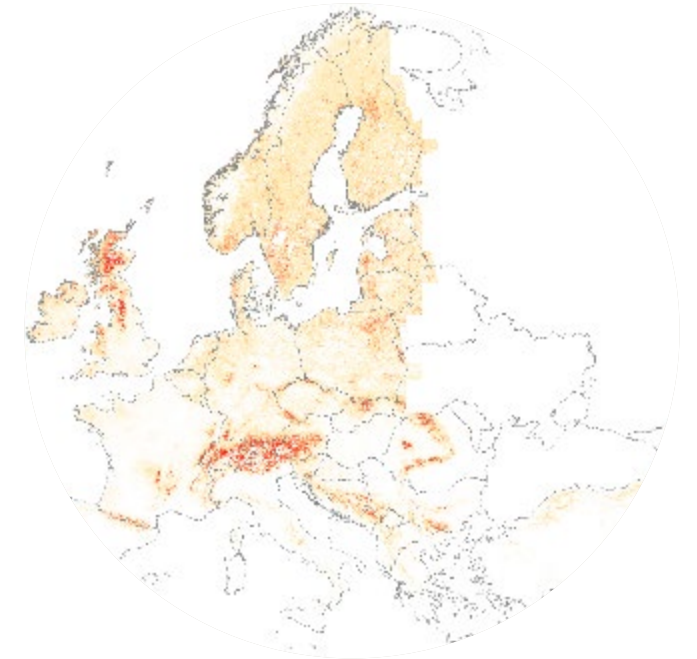


# Availability and use of in situ data for European habitat mapping

**Sander Múcher, Stephan Hennekens & Stan Los**

**Wageningen Environmental Research (WENR)**

Session Ecosystem Extent, 11 February 2025, BIOSPACE25 Conference, Frascati, Italy



# Context



- There is a **strong decline in habitats** and associated **biodiversity** mainly due to **land use & climate changes**.
- **75%** of all European habitats are assessed as **poor or bad** (EEA State of Nature).
- **Therefore mapping & monitoring European habitats** is **key** for the EU's biodiversity strategy for 2030 and the new Nature Restoration law.
- **Improved mapping of European habitats** using machine & deeplearning algorithms is our **major aim**.
- **Key issue is that a huge amount and well distributed EU in-situ data is needed for training of EO classifications ! Needs much effort !**



# Target typology concerns EUNIS habitats at Level 3

EUNIS is a hierachical & comprehensive habitat classification system



Level 1 code	Level 1 Name
MA	Marine habitats (L2:#6, L3:#30)
N	Coastal habitats (L2:#3, L3:#25)
P	Freshwater habitats (L2: #2, L3: #48)
Q	Wetlands (L2: #6, L3: #20)
R	Grasslands and lands dominated by forbs, mosses or lichens (L2:#7, L3: #58)
S	Heathland, scrub and tundra (L2: #9, L3: #43)
T	Forest and other wooded land (L2: #4, L3: #50)
U	Inland habitats with no or little soil and mostly with sparse vegetation(L2:#7, L3: #36)
V	Vegetated man-made habitats (L2: #6, L3: #31)



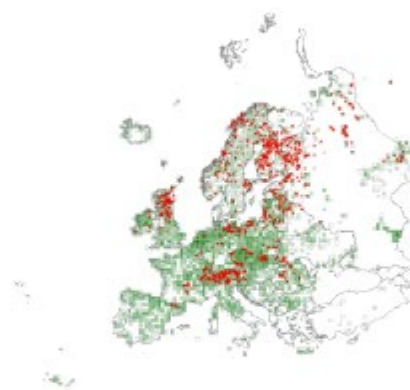
Q11 **Raised bog**

new search

Habitats → Wetlands → Raised and blanket bogs → Raised bog

Overview **Nomenclature** Species Distribution Alliances Pictures

The mire surface and underlying peat of highly oligotrophic, strongly acidic peatlands with a raised centre from which water drains towards the edges. The peat is composed mainly of sphagnum remains. Raised bogs form an nearly flat ground and are ombrotrophic, i.e. derive moisture and nutrients only from rainfall. Raised bog complexes include larger and smaller bog pools, lawns, elevated hummocks and their associated vegetation. Raised bogs form only in cool climates with high rainfall, and they are most widespread in the boreal zone and in the mountains and hills of the temperate zone; they also occur locally in the lowlands of the temperate zone. They are characteristic of lowlands and hills of North-Western and Northern Europe, the adjacent Hercynian ranges, the Jura, the Alps and the Carpathians. Bogs harbour, in addition to sphagna such as *Sphagnum fuscum*, *S. magellanicum* aggr. and *S. majus*, which are often abundant, a small number of dwarf shrubs such as *Andromeda polifolia*, *Rhododendron tomentosum*, *Vaccinium oxycoccos*, and sedges such as *Carex magellanica*, *Carex pauciflora*, *Eriophorum vaginatum* and *Trichophorum cespitosum*, non-sphagnaceous bryophytes and lichens.



See distribution...



# Vegetation Science Group & European Vegetation Survey (EVS)

Q11 **Raised bog**

new search

Habitats → Wetlands → Raised and blanket bogs → Raised bog

Overview **Nomenclature** **Species** Distribution Alliances Pictures

**Diagnostic taxa**

*Carex pauciflora*, *Eriophorum vaginatum*, *Drosera rotundifolia*, *Chamaedaphne calyculata*, *Rhododendron tomentosum*, *Rubus chamaemorus*, *Vaccinium microcarpum*, *Vaccinium oxycoccos*, *Andromeda polifolia*

**Bryophytes:** *Mylia anomala*, *Polytrichum strictum*, *Sphagnum capillifolium* aggr., *Sphagnum fuscum*, *Sphagnum magellanicum* aggr., *Sphagnum recurvum* aggr., *Sphagnum rubellum*

**Constant taxa**

*Myrica caerulea* aggr., *Betula pubescens*, *Betula nana*, *Pinus sylvestris*, *Carex nigra*, *Carex pauciflora*, *Trichophorum cespitosum*, *Eriophorum vaginatum*, *Drosera rotundifolia*, *Chamaedaphne calyculata*, *Calluna vulgaris*, *Rhododendron tomentosum*, *Rubus chamaemorus*, *Empetrum nigrum* aggr., *Eriophorum angustifolium*, *Vaccinium microcarpum*, *Vaccinium myrtillus*, *Vaccinium oxycoccos*, *Vaccinium uliginosum*, *Vaccinium vitis-idaea*, *Andromeda polifolia*

**Bryophytes:** *Aulacomnium palustre*, *Mylia anomala*, *Pleurozium schreberi*, *Polytrichum commune*, *Polytrichum strictum*, *Sphagnum capillifolium* aggr., *Sphagnum fuscum*, *Sphagnum magellanicum* aggr., *Sphagnum recurvum* aggr., *Sphagnum rubellum*, *Sphagnum russowii*

**Lichens:** *Cladonia arbuscula* aggr., *Cladonia rangiferina*

**Dominant taxa**

**Bryophytes:** *Sphagnum magellanicum* aggr., *Sphagnum recurvum* aggr.

Chytrý M., Tichý L., Hennekens S.M., Knollová I., Janssen J.A.M., Rodwell J.S. ... Schaminée J.H.J. (2020) EUNIS Habitat combinations and distribution maps of European habitats. *Applied Vegetation Science* 23: 648–675. <https://doi.org/10.1111/avsc.12519>

Version 2021-06-01, <https://doi.org/10.5281/zenodo.4812736>.

For the official presentation of the EUNIS Habitat Classification from the European Environment Agency, please see: EUNIS Terrestrial Habitat Classification 2021. The FloraVeg.EU presentation may show modifications and partial updates to the habitat classification.

Chytrý M., Tichý L., Hennekens S.M., Knollová I., Janssen J.A.M., Rodwell J.S. ... Schaminée J.H.J. (2020) EUNIS Habitat Classification: expert system, characteristic species combinations and distribution maps of European habitats. *Applied Vegetation Science* 23: 648–675. <https://doi.org/10.1111/avsc.12519>

Version 2021-06-01, <https://doi.org/10.5281/zenodo.4812736>.

For the official presentation of the EUNIS Habitat Classification from the European Environment Agency, please see: EUNIS Terrestrial Habitat Classification 2021. The FloraVeg.EU presentation may show modifications and partial updates to the habitat classification.

# ML/AI habitat mapping strategy

Predictor groups

Process

Combinations

ML algorithms

Classification strategy

Evaluation metrics

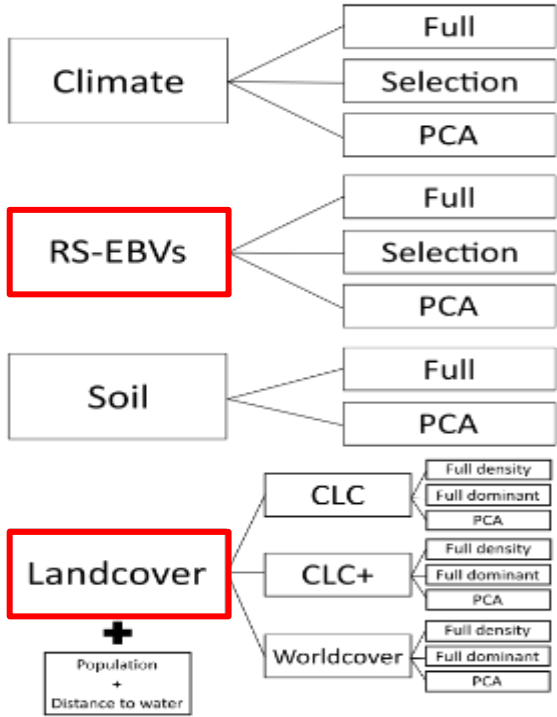
Can the information be summarized? What is the intrinsic information carried by each group? What process combination is the most relevant for the modelling?

Which model gives the better results? What process and combination are the most adapted to what model?

- Recall: Are all of the habitats predicted?
- Precision: How faithful are the predicted habitats?
- F1 score: Is the balance good between the habitat prediction quantity and quality?
- Top k accuracy: Is the real label among the most likely? If the recall is low, but the top k is high, how many similar habitats are co-occurring? (given by k)
- Confusion matrix: What habitats are confused with others? In what proportion?

EO data

EO data

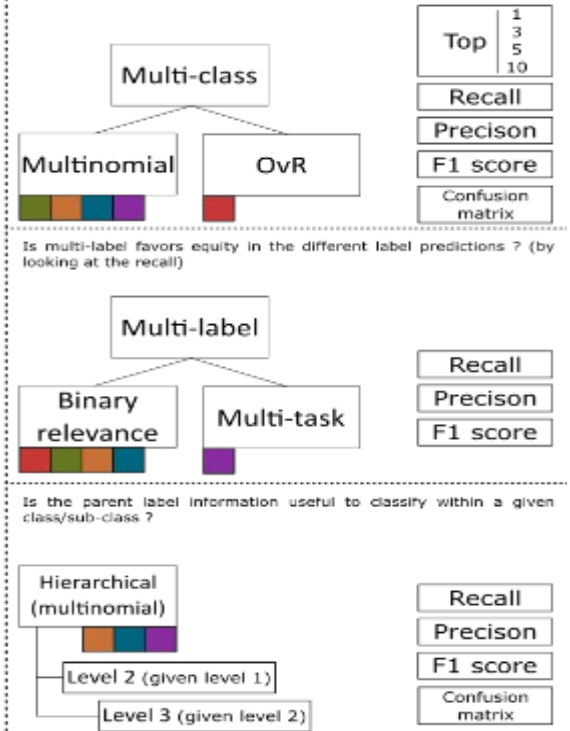


	Climate	RS-EBVs	Soil	Landcover
Combination 3	PCA	PCA	PCA	PCA
Combination 2	Selection	Selection	Full	Full density
Combination 1	Full	Full	Full	Full dominant

For every combination, one predictor group is put aside by turns, to assess its carried information compared to when every groups are present.

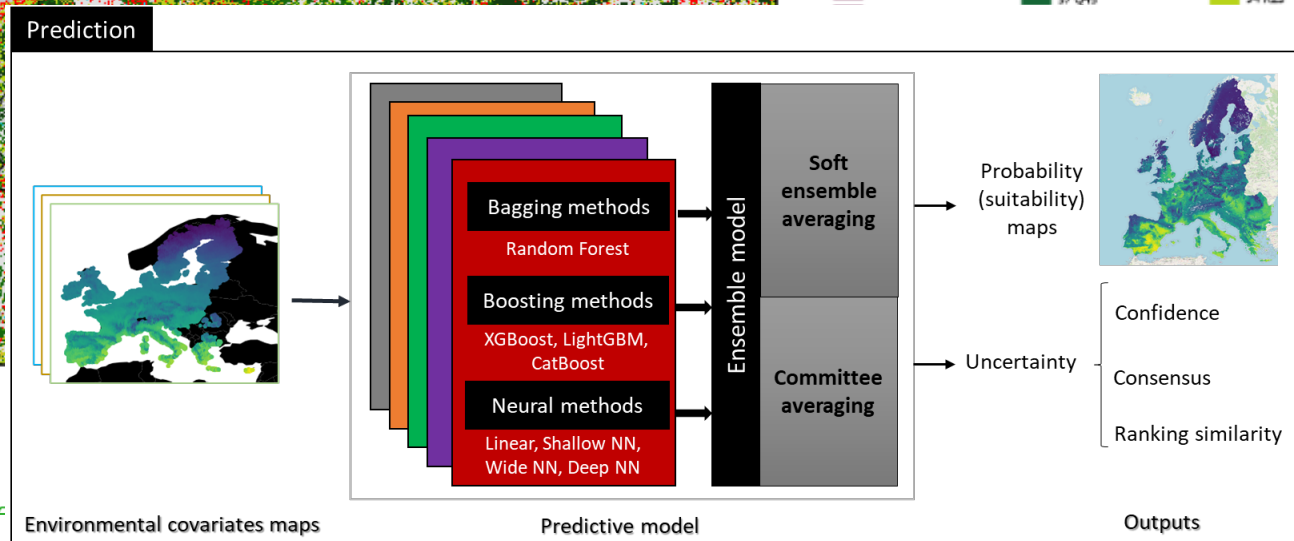
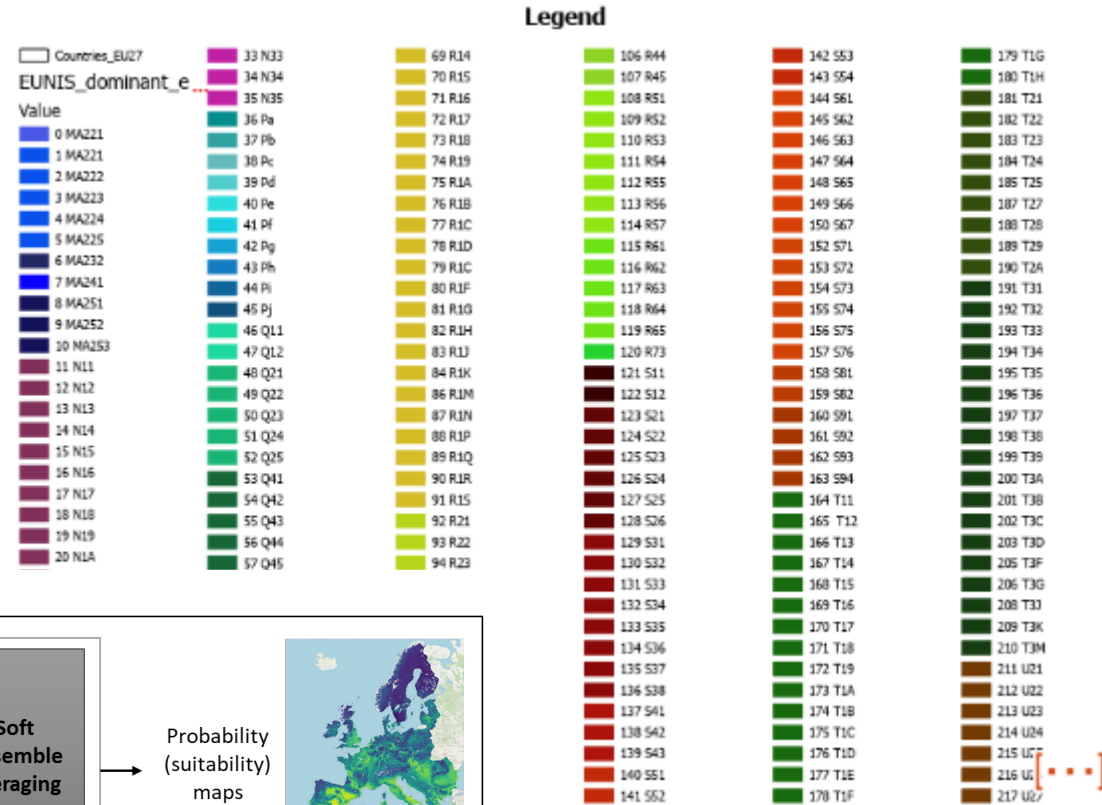
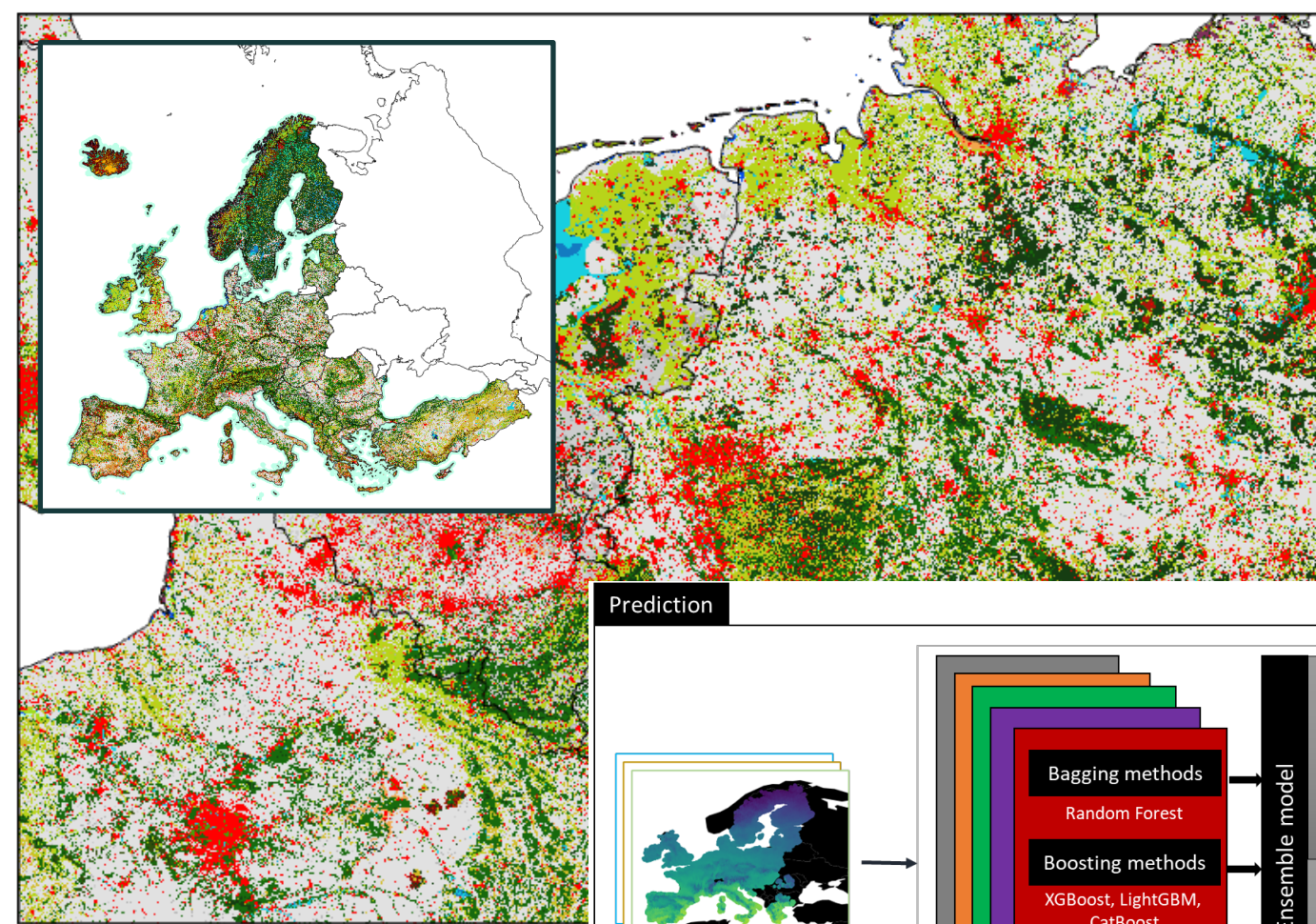
- SVC
- Maxent
- Random Forest
- XGBoost
- CNN

TRAINING DATA





# New European habitat map (EUNIS level 3, #251 CLASSES) using 22 predictors & EVA archive & hierarchical ensemble MLC

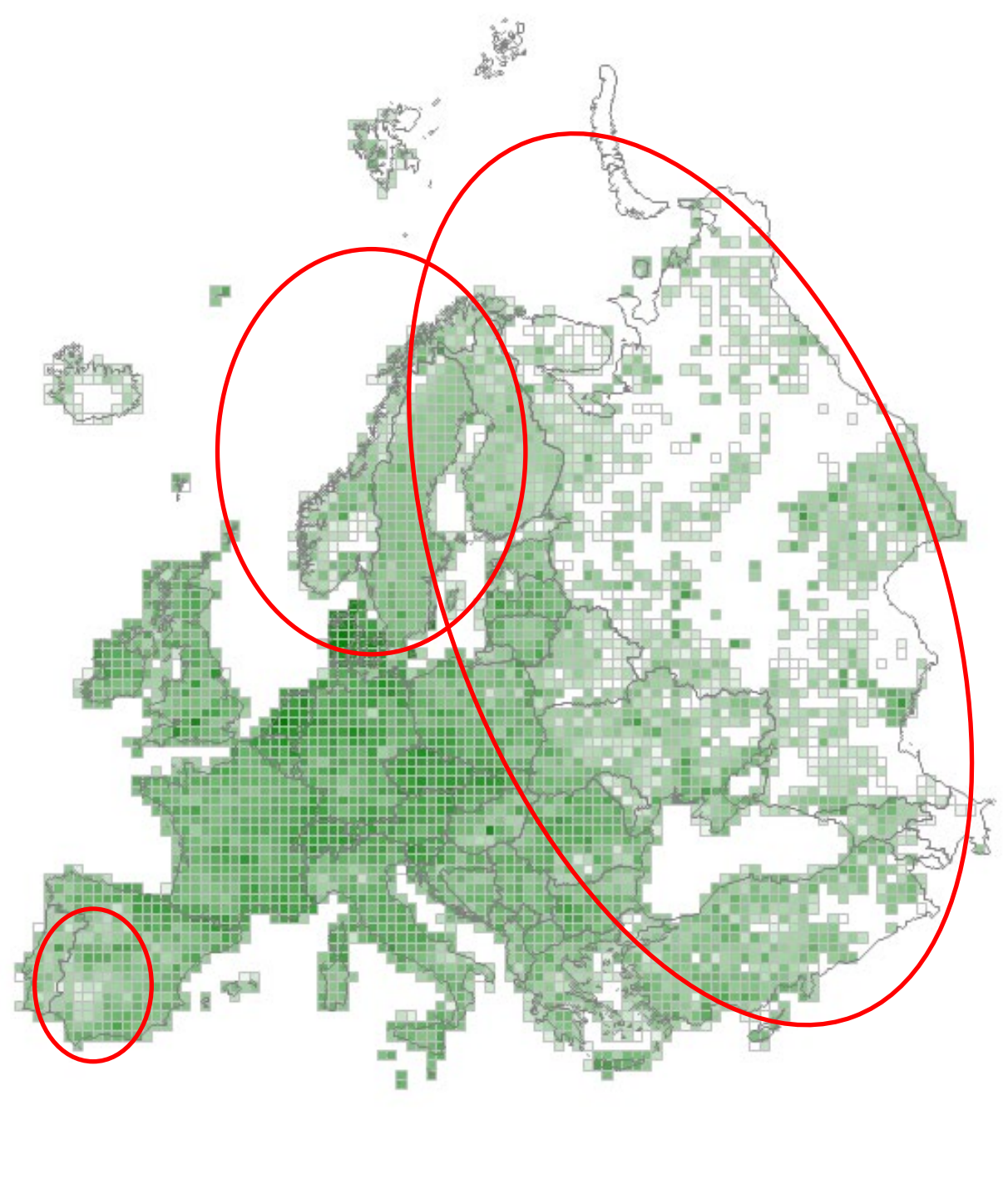
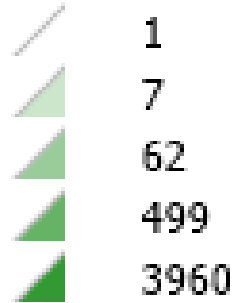


# IN-SITU data: Why GBIF next to EVA

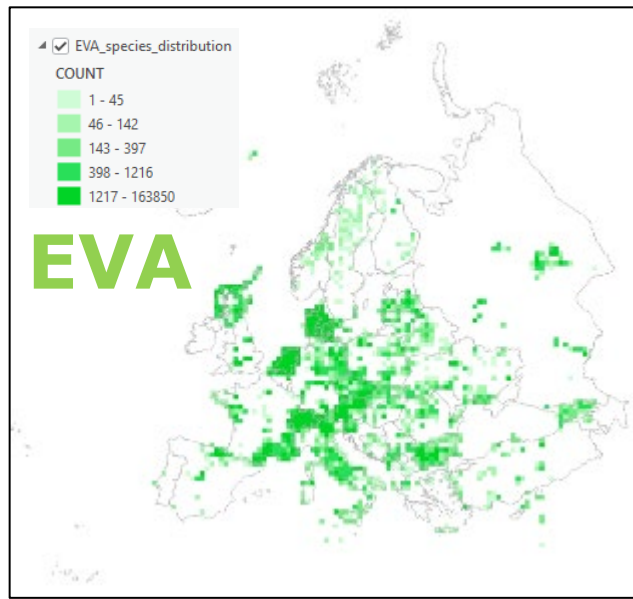
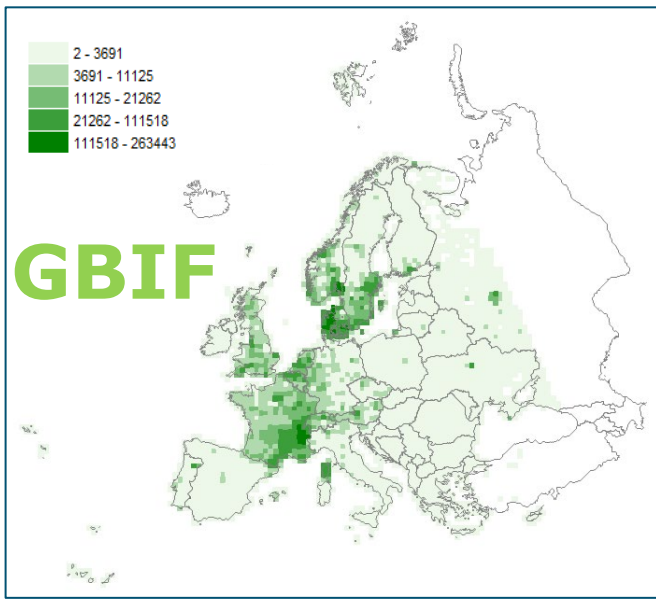
- **European Vegetation Archive (EVA)** with **~2.6 million vegetation plots** is used as main basis for training EUNIS habitat classifications.
- There are three **limitations**:
  - 1. Spatial limitations.** Especially Scandinavia, Eastern Europe and the parts of Spain and Turkey are unrepresented in the database;
  - 2. Temporal limitations.** Only half of the total number of plots is recorded from the year 2000 (1.3 million) and **only 0.72M records from the year 2010 onwards**;
  - 3. Location uncertainty** is a major issue in **EVA**. There are 343,000 vegetations plots NOT georeferenced, and only **183,000 plots** have a **locational accuracy** of  **$\leq 10\text{m}$** .
- If using only EVA vegetation plots from the year 2000 onwards with 10m accuracy -> 115,000 plot observations. Because of underrepresentation certain areas the **GBIF data might provide a solution**

# Density map of EVA plot observations (2.6 M records)

Plot data (50x50 km)







**Exploiting species combinations at gridcell level 10m or 100m (52.2 million records in total).**

Each EUNIS habitat type is characterized by floristic composition by means of three categories:

1. Diagnostic species
2. Constant species
3. Dominant species



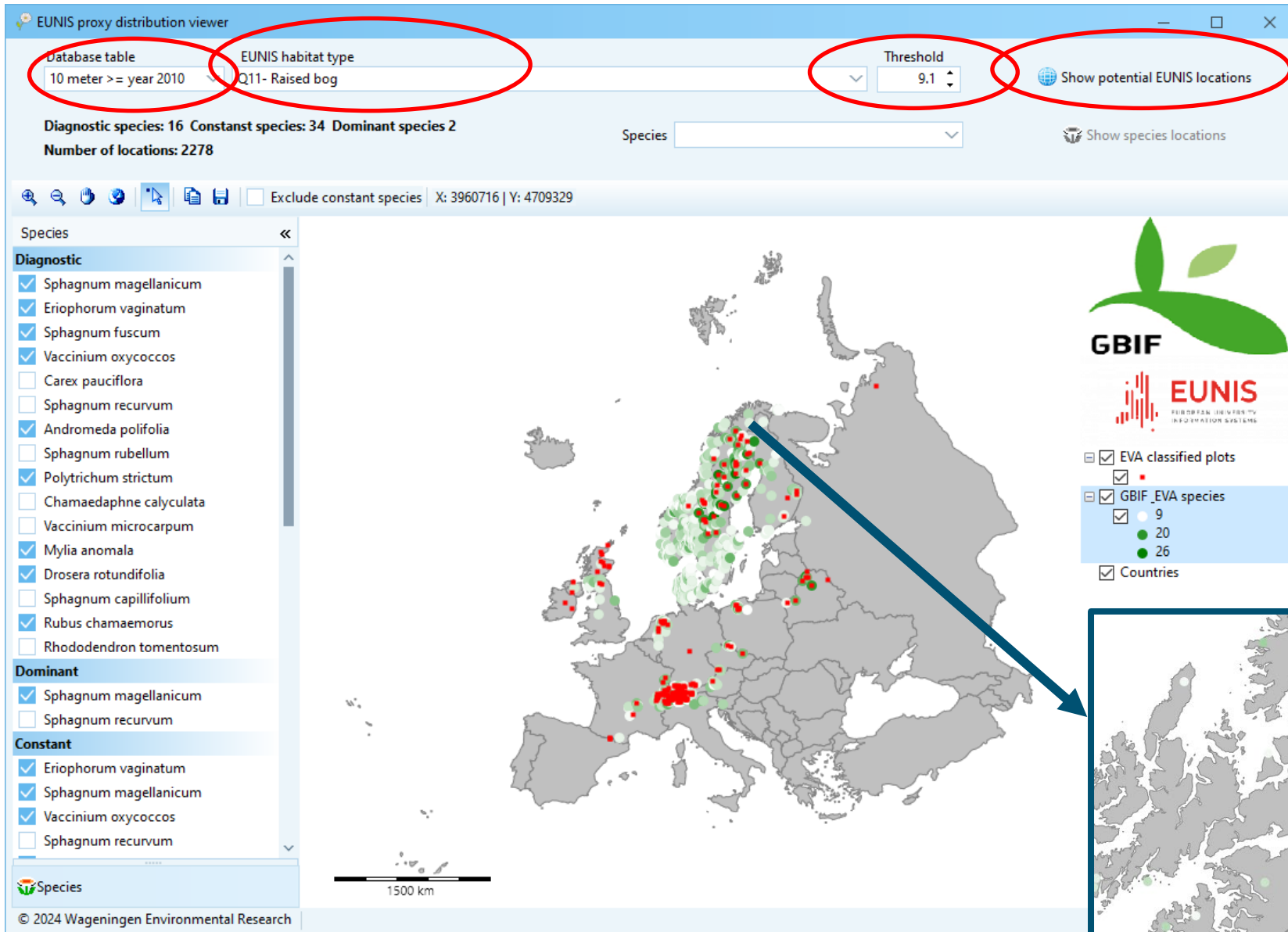
**40.4 million GBIF records extracted + 11.8 million species records from EVA**

Category	Species name	Value	Weight	Value*Weighth*0.01
Diagnostic	<i>Sphagnum magellanicum</i>	44.6	5	2.23
Diagnostic	<i>Eriophorum vaginatum</i>	39	5	1.95
Diagnostic	<i>Vaccinium oxycoccos</i>	33.9	5	1.695
Diagnostic	<i>Polytrichum strictum</i>	27.9	5	1.395
Diagnostic	<i>Drosera rotundifolia</i>	23.4	5	1.17
Dominant	<i>Sphagnum magellanicum</i>	31	2	0.62
Constant	<i>Eriophorum vaginatum</i>	87	1	0.87
Constant	<i>Sphagnum magellanicum</i>	66	1	0.66
Constant	<i>Vaccinium oxycoccos</i>	65	1	0.65
Constant	<i>Calluna vulgaris</i>	56	1	0.56
Constant	<i>Polytrichum strictum</i>	46	1	0.46
Constant	<i>Drosera rotundifolia</i>	46	1	0.46
Constant	<i>Empetrum nigrum</i>	39	1	0.39
Constant	<i>Vaccinium myrtillus</i>	28	1	0.28
Constant	<i>Molinia caerulea</i>	17	1	0.17
Constant	<i>Eriophorum angustifolium</i>	12	1	0.12
<b>Total</b>				<b>13.68</b>

Value

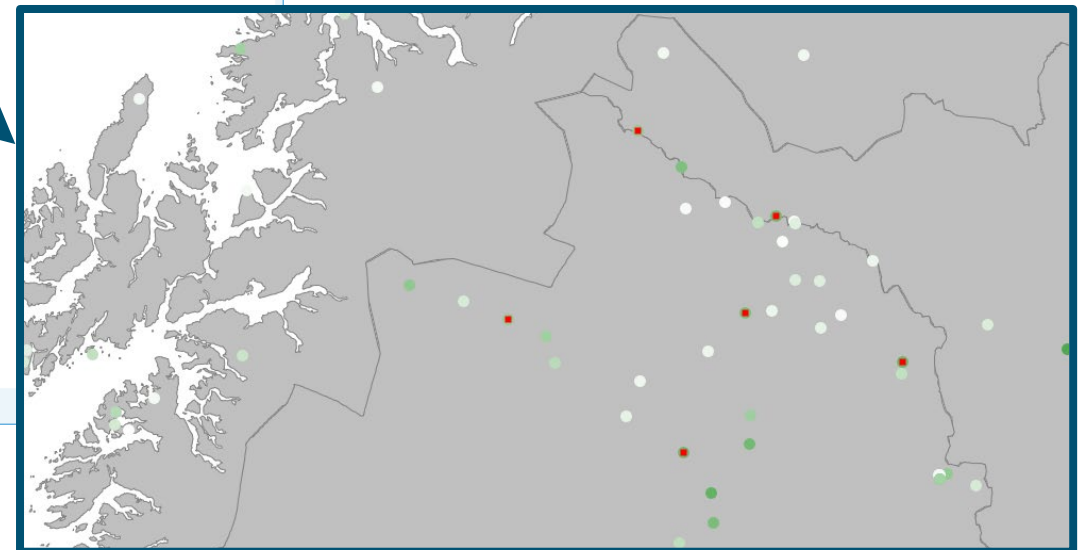
Example of **Q11 Raised bog** with data on **diagnostic, constant and dominant species**

# New software: EUNIS proxy distribution viewer (EPDV)



Option Database Tables:

1. Table species observations with location uncertainty of **10 meter** or less, recorded after **2010** (**15.1 million records**)
2. Table species observations with location uncertainty of **10 meter** or less, recorded **after 2000** (**20.1 million records**)
3. Table with species observations with a location uncertainty of **100 meter** or less and recorded after 2010 (**23.2 million records**)
4. Table with species observations with a location uncertainty of **100 meter** or less and recorded **after 2000** (**29.7 million records**)





# Conclusions

- New potential habitat locations can be found by exploiting GBIF plant species combination at gridcell level with the new EPDV tool
- Preferably a **10m grid** is used with records after **2010 (15.1 million records)**
- **Challenge:** which **threshold value** should be set per EUNIS habitat type ?

# Key message: European Habitat Mapping & in-situ data

- With regard to operational biodiversity monitoring, we cannot monitor all species. But we can monitor the associated habitats with EO & in-situ data.
- For habitat mapping with EO data & machine learning it is key to have a lot of good in-situ data for training purposes.
- While temporal en spatial resolutions of EO data is increasing a lot, in-situ data is lacking behind.
- More effort is needed in the collection in-situ data to support EO mapping & monitoring !





# Thank you for your attention

Contact: Sander Múcher, Wageningen Environmental Research (WENR)  
Tel: +31 (0) 317 481607, email: [sander.mucher@wur.nl](mailto:sander.mucher@wur.nl)