

# AN OVERVIEW OF THE GBIF AND LIFEMAPPER DATABASES

## INTRODUCTION

### GBIF

- Pros:
  - Huge amount of data from all over the world
  - Free access to the data
  - Possibility to view most of the data in a map

### Cons:

- Online data as a table is uncomfortable to work with
- There is a fair amount of redundancy in the database
- Some species have no entries
- Some entries are incomplete

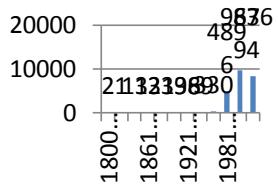
### *Pinus nigra*

- Least concern, spread through Europe
- 23537 occurrences → 14573 (61.91%) complete
  - 798 (3.39%) no coordinates no date
  - 7566 (32.15%) no date
  - 600 (2.55%) no coordinates
- Data goes back to the XIX century
- Redundancy
  - Some big groups (100+ entries)
  - Mostly occurs in entries with only one of the two datasets analyzed
  - The GBIF Spain publisher is the one with the most redundancy, as well as the most entries
- Lifemapper does not have data on this species

### Lifemapper

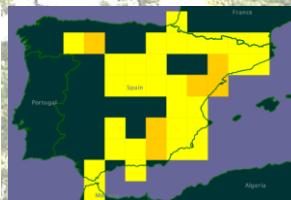
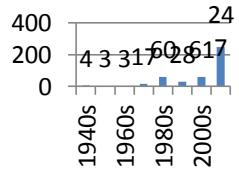
#### Pros:

- Satellite based maps
- Offers the option to run climate models
- Cons:
  - Does not have as much data regarding the species as GBIF
  - Can not run both climate and species distribution maps at once
  - Number of entries differ between what is listed and what is mapped



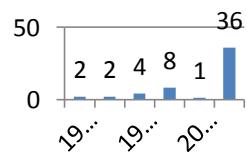
### *Pinus nigra salzmannii*

- Least concern, found in Spain
- Subspecies endemic to the Iberian Peninsula
- 406 entries → 87 (21.43%) complete
  - 49 (12.07%) no coordinates no date
  - 198 (48.77%) only coordinates
  - 72 (17.73%) only date
- Redundancy: many small groups
- Lifemapper did not support this data.



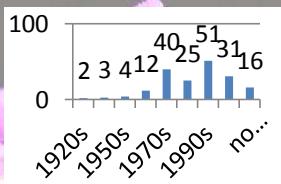
### *Sorbus leptophylla*

- Critically endangered species of England
- 53 occurrences → 14 (26.42%) complete
  - 6 (11.32%) no date no coordinates
  - 3 (5.66%) no coordinates
  - 30 (56.60%) no date
- Redundancy: 3 big groups
  - 23 entries
    - 11 have different dates → Monitoring
    - 12 no date → redundancy
  - 9 entries → only 1 with date
  - 6 entries → 2 dated
  - 50% entries are redundant.
- Lifemapper does not support this species



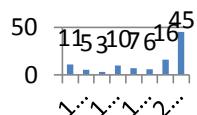
### *Byblis filifolia*

- Least concern, found in Australia
- 184 occurrences → 167 (90.76%) complete
  - 1 (0.54%) no date no coordinates
  - 1 (0.54%) no coordinates
  - 15 (8.16%) no date
- Repeated coordinates through the years → monitoring
- Lifemapper offers 144 hits of the 182 GBIF's map has.



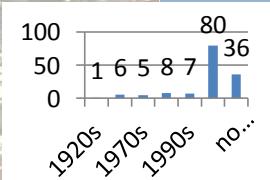
### *Byblis gigantea*

- Critically endangered, found in Australia
- 103 entries → 48 (46.60%) complete
  - 35 (33.98%) no date or coordinates → many come from foreign museums
  - 10 (9.71%) no coordinates
  - 10 (9.71%) no date
- Lifemapper does not use all the entries with coordinates, leaves 12 out
- From the 46 points listed, 13 are displayed in the map



### *Coronopus navasi*

- Critically endangered species of the Iberian Peninsula
- 144 entries → 94 (65.28%) complete
  - 8 (5.56%) no date no coordinates
  - 14 (9.72%) no coordinates
  - 28 (19.44%) no date
- Redundancy:
  - 18 entries
    - 7 dated → monitoring
    - 11 not dated → redundant
  - 16 entries → redundant
  - 28 entries → redundant
- Lifemapper
  - Coronopus navasi PAU → has entries in the sea
  - Later this entry was deleted from Lifemapper
  - Coronopus navasi → correct entries



## DISCUSSION

### GBIF

- The incomplete entries and the redundant ones should be dealt with either by removing them, completing them, or tagging them as such

- There should be an option to directly report incomplete or redundant data so that it could be fixed

- It should have the option to select entries from certain years, as it already has by countries, datasets and publishers.

- Rework the data table provided by the website, 20 entries per page is not enough, and knowing how many pages of entries the species has would be helpful too

### Lifemapper

- Should import more data from the GBIF, as it is quite incomplete

- Being able to use the climate layers together with the species ones would add a lot of interest in this project

- The number of entries displayed on the list of species selection should be the same as the number of entries present in the map