

Practical Marine ecology: open data resources

This assignment should be completed during the practical.

The report using this template needs to be submitted before the start of the exam period the latest. The key objective is to utilise the tools and engage in discussion within the group. The marking is either a pass or a fail.

Groups are composed of 2 or 3 persons.

Group Member:

Group Member:

Group Member:

Resources

www.marinespecies.org

<https://dwc.tdwg.org/terms/>

www.obis.org

<https://manual.obis.org/>

<https://mapper.obis.org/>

www.gbif.org

<https://www.gbif.org/data-quality-requirements-occurrences>

<https://www.gbif.org/tool/81281/gbif-data-validator>

<https://www.aquamaps.org/>

Task 1. Species list

<https://www.marinespecies.org>

You may find a species list of the species found during the practical in Ostend [here](#).

1. Quality control the list using online tools. You should retrieve the Full **accepted** scientific name, Authority, LSID
2. How many taxa do not match? Explain why they do not match and make them match.

Task 2. OBIS & GBIF

Select one taxon from the species list per group member (So 2 or 3 in total) and provide an overview of the species using open resources.

You will need to register at GBIF to download information.

Search for occurrence records from OBIS (www.obis.org, <https://mapper.obis.org/>) and GBIF (www.gbif.org).

Compare the results between OBIS and GBIF?

3. What is the largest dataset?
4. How many occurrences globally?
5. What are the most common flagged issues?
6. When were the most occurrences collected?

Environmental conditions

7. What is the temperature range, salinity range and depth range of the species?

8. How many occurrences within the Belgian EEZ?

POLYGON((2.3903 51.2692,2.2383 51.5578,2.2533 51.6131,2.4817 51.805,2.5393 51.8761, 3.36361 51.3707,3.2058 51.3428,3.0851 51.3071,2.91021 51.2308,2.5496 51.0884,2.3903 51.2692))

9. Make a map of all occurrences. Highlighting the occurrences in the belgian EEZ.
10. Are there occurrences that seem wrong? Why?
11. What will the future distribution be?

Task 3. Darwin - Core

Clean the dataset that can be found

Amphipod data set: [Amphipods_example_test_dirty\(xls\)](#), [google spreadsheet](#)

An overview of Darwin Core terms can be found here:

<https://dwc.tdwg.org/terms/>

ObiS manual

https://manual.obis.org/darwin_core.html

12. Match the Column headers with Darwin core terms. Are there important terms missing?

Clean the dataset so that all required terms are there and the values are correct.
Provide a description of the errors you've found and the steps you took to clean them.

13. What terms did you have to add?

14. For each term, describe the errors you found

Use the GBIF data validator on the cleaned file

15. Provide an overview of where the data Improved based on the two data validations