**Data extraction “Module Ulster University” project**

**Extraction performed by Pierre Hélaouët**

**03/02/2025.**

**DOI: 10.17031/67a07cb284efa**

**Web page: https://doi.mba.ac.uk/data/3426/**

**Note:** As always, this dataset has been carefully built and checked accordingly. However, it is the user’s responsibility to perform his own verifications.

**Quick description of the dataset**

**1 – The dataset contains 4 files:**

1. “CPR\_Data\_ModuleUlsterUniversity\_03022025.docx”: This document
2. “CPR\_ModuleUlsterUniversity\_ControlMap\_03022025.png”:

Map representing the selected samples from January 1958 to December 2021 (105644 samples)

1. “CPR\_ModuleUlsterUniversity\_Data\_LargeZooplankton\_03022025.csv”:

Annual averaged abundance data for all selected large zooplankton (104 taxa, see CPR\_ModuleUlsterUniversity\_List\_LargeZooplankton\_03022025.csv) in 13 selected CPR standard areas (A1; B1; B2; B4; C1; C2; C3; C4; D2; D3; D4; E4; F4).

Rows: Annual averaged abundance for each area (13). Total = 832.

Column 1: “StdArea\_Name”. CPR standard area names

Column 2: “Year”

Column 3: “Number\_Samples”: Number of samples used to calculate the averages.

Columns from 4 to 107: Annual averaged abundance for each selected large zooplankton (104 taxa).

Note 1: We may notice very small values (10^-10) or any other number with a very small fraction. Sometimes, our analysts can identify the presence of a specific taxa but are unable to quantify it. In that case, they report the taxa as “present”. This is hard-coded in our database as a very small value (10^-10) for statistical reasons.

Note 2: In a given sample, the abundance value of a specific taxon, is set to NaN (Not A Number) when the corresponding Data of Routine Identification (DRI) is posterior to the date of sample collection.

1. “CPR\_ModuleUlsterUniversity\_List\_LargeZooplankton\_03022025.csv”: List of large zooplankton

Rows: All selected taxa (104 taxa).

Column 1 “accepted\_id”: Unique identifier used by the CPR survey

Column 2 "Aphia\_id”: Identifier used by WoRMS

Column 3 "name-CPR”: Unique name used by the CPR survey.

Column 4 "Name\_worms”: Name used by WoRMS corresponding to the “aphia\_id”.

Column 5 "DRI”: Date of Routine Identification. Before that date, un taxon was not on our routine taxa list. For a given taxon, abundances associated with samples taken before the DRI are set to a NaN (Not A Number).