

Reporting data to the MACROFAUNA database v1.5

March 2023

contact : jerome.mathieu@upmc.fr

First of all, thanks a lot for your time and effort. This is very valuable for the project!

The goal of the database is to get estimates of the density and biomass of the major soil macrofauna groups in a variety of situations. Density means here the number of individuals of the group by a unit of surface measured on a short time. Specimens should be alive and found either in the soil or in the litter.

As a general rule, please fill template with as much as possible details. You can describe land cover and land management in your own words, in a verbatim manner, with sentences for instance. Don't worry about space. We will recode afterwards this information to fit it in a controlled vocabulary.

As a general rule, please feel free to add any information that is not in the template, but that you judge important.

If possible, please send any published paper associated to the dataset. This is important for repeatability and data quality checks.

When possible, please send pictures (CC-BY-4 licence) of each transect, monoliths or soil profiles. Please name the file in a way that we can match them to the unit they are representing. For instance:



Reporting abundances

A very important point is that we need to be able to distinguish "abundance = 0" from "no data"

- If a taxa would have been reported if it was found, but it was not present at all, then you should fill the abundance column with 0. Do not leave empty cells. This way we will know that the taxa was absent (instead of "not reported")
- Columns with only empty cells mean "No data": the taxa was not reported, even if it was present.
- You need to do this for all taxa that are mandatory (indicated in the explanation sheet of the template)

Reporting depth

Please indicate the top and bottom depth of the layer, not just the bottom one. Please put a space between each string.

Eg. "0 – 20 cm", not "20cm". If litter is included, please indicate it with "L".

Hierarchical organization of the samples

Layers are parts of a monolith, monoliths belong to a transect. Transects are part of a plot when:

There are several transects in the same plot, at different locations (e.g. : spatial replicates) or at different times (repeated measures).

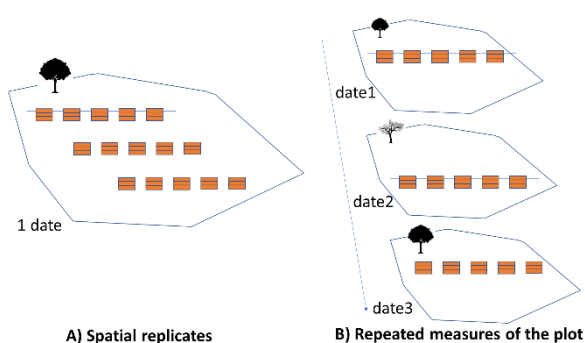
A plot is a physical spatial area, such as a crop, a patch of forest, which is homogenous. A transect is a sampling event in space and time : a group of samples taken near to each others at the same date.

There is one column ID for each level. You need to use the ones that are relevant to your data, according to their resolution. The only ID that is mandatory is the transect ID.

When there is only one monolith in the transect, there should be an ID for the transect anyway.

	Available data	Illustration	Data ID required
Illustration of the	Average density of taxa over the transect		transect_ID
IDs required as a function	Fauna density for each monolith		transect_ID monolith_ID
of the resolution of	Fauna density for each layer of each monolith		transect_ID monolith_ID layer_ID
the available data			

Cases when a plot_ID is necessary



The sheets of the template

- **Sheet "0_data_set_info"**

This sheet summarizes the context of the dataset and the terms of use. Please fill in carefully this sheet.

Please note that by default we plan to release the global data base on open access model (CC-BY-4). If you do not want your data to be released, please advise us. This is not an issue.

Please also check with your eventual co-workers that they agree on this term and on authorship. See our collaboration guidelines here: <http://soil-organisms.org/index.php/SO/article/view/282>

- **1_macrofauna**

Taxonomical coverage

The column "other_groups" in the data sheet refers to the groups that were observed but not reported individually in previous columns.

Please indicate the list of these other taxa at the bottom of the sheet 0.

Taxonomy

Please provide the data at the finest taxonomical resolution as possible. You can add groups by adding columns, named after the taxa and use the suffix _D for density and _fbm for fresh biomass. We will harmonize taxonomy afterwards. We use GBIF backbone classification system by default.

Units of density and biomass

Please provide the density of individuals in ind.m⁻²: For biomass, please use mg.m⁻²

Don't forget to explain the unit used, on the "0_data_set_info sheet".

- **2_lu_sampling**

Please be as much complete and accurate, especially in the lu cover and management columns. We will recode this info afterwards. In case of experiments, you can describe the treatment level in the management column.

Please indicate the date with the best accuracy possible, it will be used to get climatic data at the sampling date. If you have the time of sampling, you can include it.

Geographic coordinates should be in decimal degrees (eg XX.XXX), not hexadecimal (XX°XX'XX"), with WGS84 geodesic system (in doubt, use the google maps coordinates). It is very important to have accurate coordinates because we will use them to extract landscape metrics around the samples. Ideally provide an estimate of the coordinates' precision.

- **3_soil**

Each soil property must be described by a group of 4 variables: value of the property, unit, method, depth. The name of each soil property must start with "soil_". Eg soil_silt, soil_om, soil_cec, and so on.

If you have soil data at several depths, please create group of 4 variables for each depth, and include the depth in the name of the variables. Eg. soil_ph_0_10cm, soil_ph_10-20cm, and so on.

Please explain how you defined the depth: is "0" starting on top of humus or at the transition from humus to mineral matrix.

For soil properties, please indicated the depth of the measurement, this is an important point!

Once we received your dataset, we will assign it an ID, review it and send you a couple questions to do checks.

Thanks again and looking forward hearing from you!

The Soil MACROFAUNA team

Links

Website : <http://www.globalsoilmacrofauna.com/>

Official reporting template : <https://doi.org/10.5281/zenodo.4543852> (double check that you use the latest version)