

Preparing habitat data for submission to EMODnet Seabed Habitats

Contents

Step 1:	Assign a GUI	2
Step 2:	Prepare the shapefile.....	2
Step 3:	Standardise the attribute table	3
Step 4:	Confidence Assessment	5
Step 5:	Create a Study Area polygon.....	5
Step 6:	Create the metadata	7
Step 7:	Complete the data provider agreement.....	7
Step 8:	Submit the data 'package'	8

Introduction

This document provides a summary of the steps that need to be taken by EMODnet Seabed Habitats project data suppliers to:

1. Prepare shapefiles in a suitable format for sharing and displaying on the EMODnet Seabed Habitats webGIS in ESRI ArcGIS software.
2. Produce metadata in a suitable format for sharing and displaying on the ICES metadata catalogue.
3. Conduct a confidence assessment on habitat maps.

Please note: ALL GIS data must be supplied in shapefile (.shp) format.

Guidance and template files are provided in “**EMODnetSBH complete data provider package.zip**”, which can be downloaded by clicking on the “updated data provider files” heading on this page: www.emodnet-seabedhabitats.eu/Default.aspx?page=1636.

The contents of the package are:

- Preparing habitat data for submission.pdf (*this document*)
- Data Exchange Formats.pdf (*includes all sub-documents below*)
 - Original Habitat DEF.pdf
 - Translated Habitat DEF.pdf
 - Habitats Directive DEF.pdf
 - Study Area DEF.pdf
- Confidence assessment scoresheet.xls
- Confidence assessment guidelines.pdf
- EMODnetSBH Data Provider Agreement.doc
- EMODnetSBH Data Provider Agreement supplementary.doc
- Supplying data – GIS tips.pdf

Additionally, python scripts for use with ArcGIS to batch run certain tasks are available to download from the [EMODnet Seabed Habitats GitHub repository](#).

If any further assistance is required, please do not hesitate to contact us at info@emodnet-seabedhabitats.eu

Step 1: Assign a GUI

An important field in the habitat map attribute table is the globally unique identifier (GUI), which is a unique code identifying a dataset (e.g. a single habitat map).

The GUI consists of a 2-letter county code (which corresponds to [ISO3166-1](#)) followed by 6 digits. For example, a dataset from Italy would be written IT000005. The final 6 characters can be used freely although we suggest sequential numbering of datasets to help prevent duplication.

To ensure each GUI is unique, here are some suggested GUI ranges for organisations that have previously submitted data:

lfremer:	FR000001 to FR001999 (excluding previous MESH GUIs)
IMA:	FR002000 to FR002999
DREAL:	FR003000 onwards
Universidade do Algarve:	PT000001 to PT000999
Universidade de Aveiro:	PT001000 to PT009999
IPIMAR:	PT010000 to PT099999
Universidade dos Açores:	PT100000 onwards
AZTI:	ES000001 to ES000999
IEO:	ES001000 onwards
Marine Institute:	IE001000 onwards

Warning: Each GUI must be unique to an individual survey or habitat map and must not be re-used.

If your organisation is not in the list above, please contact info@emodnet-seabedhabitats.eu and we will direct you to your country's project [partner](#) who will supply you with GUI codes, who will supply you with GUI codes.

The GUI in the attribute table will be used as the Alternate Title field in the metadata record created in Step 3.

Step 2: Prepare the shapefile

All shapefiles provided to EMODnet Seabed Habitats must be supplied in the WGS84 unprojected coordinate system ([EPSG:4326](#)¹).

If your habitat map is not in this coordinate system, then it must be reprojected before continuing with the data submission. Within ArcGIS, for example, this can be performed using the "[Project](#)" tool.

To ensure that your map data can be properly integrated in composite datasets, and to allow the data to display properly, the shapefile should be cleaned to remove any geometry or topology errors. For more information on how to clean the shapefile, please read "**Supplying data: GIS tips**".

Finally, the supplied habitat shapefile should be named as its GUI. For example, a habitat map with GUI "IT000005" should be named as "IT000005.shp".

¹Within ArcGIS, this information can be found after "WKID" (well-known identifier) within the Coordinate System dialogue box. WKID numbers are searchable in the Coordinate System dialogue box. Within QGIS, EPSG codes can similarly be entered as a filter term when selecting a Coordinate Reference System.

Step 3: Standardise the attribute table

To enable habitat maps from different sources to be compared and combined into the large datasets used in EMODnet Seabed Habitats and displayed on the interactive map, the map's attribute table must be converted into a standardised format. We call this the Data Exchange Format (DEF).

For the purposes of EMODnet Seabed Habitats, data are required to be submitted in an updated version of the MESH DEFs as they provide a tried and tested method of standardising maps for collation and conflation.

There are two possible habitat data exchange formats, dependant on the habitat map itself, described in **Data Exchange Formats.pdf**

The map should first be converted into the **Original Habitat DEF**. If the map is detailed enough to be translated into the EUNIS classification system, or is already in EUNIS classification, the attribute table should be extended into the **Translated Habitat DEF** and the habitats translated. If the map represents habitats from Annex I of the EU's [Habitats Directive](#) (for Natura2000 sites), the attribute table should be extended into the Habitats Directive DEF.

The values for the fields in the DEF should be entered as per the Data Exchange Format guidance. Mandatory fields must contain a valid value for each feature/polygon, non-mandatory fields need not contain values but the field itself must still exist in the final submission for the data to be valid.

If the original habitat map does not contain many polygons, it should be viable to enter the DEF field values manually. If the habitat map contains a number of habitats repeated within a large number of polygons, an attribute join can be used to fill the HAB_TYPE and T_RELATE fields with less effort. A possible method of doing this is outlined in **Supplying data: GIS tips**.

Once converted into the required DEF, the final attribute table should contain all of the DEF's fields, and all extraneous fields should be removed.

Python scripting

If you have ArcMap and are comfortable with python, the change to the DEF schema can be automated by using either of two python scripts from the [GitHub repository](#). The script to use depends on your habitat map:

“**add_MESH_ORIG.py**” adds all fields required by the Original Habitat DEF.

“**add_MESH_TRANS.py**” adds all fields required by the Translated Habitat DEF.

After translation, extraneous fields may either be deleted manually, or by using either of two python scripts from the [GitHub repository](#). Again, the script to use depends on your habitat map:

“**trim_MESH_ORIG.py**” removes all fields not required by the Original Habitat DEF.

“**trim_MESH_TRANS.py**” removes all fields not required by the Translated Habitat

DEF.

The scripts provided will act on all shapefiles within a location provided by the user, and so can be used to batch-process addition and removal of fields with multiple maps.

Warning: It should be noted that due to the behaviour described above, maps to be translated must be moved into their own folder before running the scripts to avoid adding fields to or removing fields from other any other shapefiles.

Step 4: Confidence Assessment

Note: This step does not need to be completed if submitting a habitat map in the Habitats Directive DEF.

A confidence assessment should be carried out for each habitat map. **MESH Confidence Scoresheet.xls** contains a template for this.

The first sheet (“**Score sheet**”) is the only sheet requiring input, and the scores for the categories should be entered in columns B to P, with the corresponding map’s GUI in column A.

The second sheet (“**Weightings**”) contains values used in the calculation of the final score and can be left untouched.

Sheets three to five (“**Remote Sensing**”, “**Ground truthing**”, **Interpretation**”) explain the scoring for each categories on sheet one and should be read before undertaking the scoring process. The same information can be found within the accompanying document “**MESH Confidence Assessment Guidelines.pdf**”

Scores for multiple habitat maps can be entered on a single spreadsheet, and it is advised that this is done when it is expected that a package of multiple maps will be delivered. The scoring for each subsequent map should be entered on a new row, with its GUI entered into column A.

You can also use the online [MESH Confidence Tool](#) to generate scores and export a CSV textfile that can be imported into the confidence spreadsheet.

See the [MESH Confidence Assessment Scheme](#) section of the archived MESH Mapping Guide for further guidance about how to conduct a confidence assessment.

Step 5: Create a Study Area polygon

For each habitat map (GUI), a rectangular “Study Area” bounding box should be created, encompassing the surveyed areas. In ArcGIS, this can be performed by using the “[Minimum bounding geometry](#)” tool, selecting “ENVELOPE” as the “Geometry Type”, and grouping by the “GUI” field. The resulting rectangular feature should look similar to Figure 1 (below).

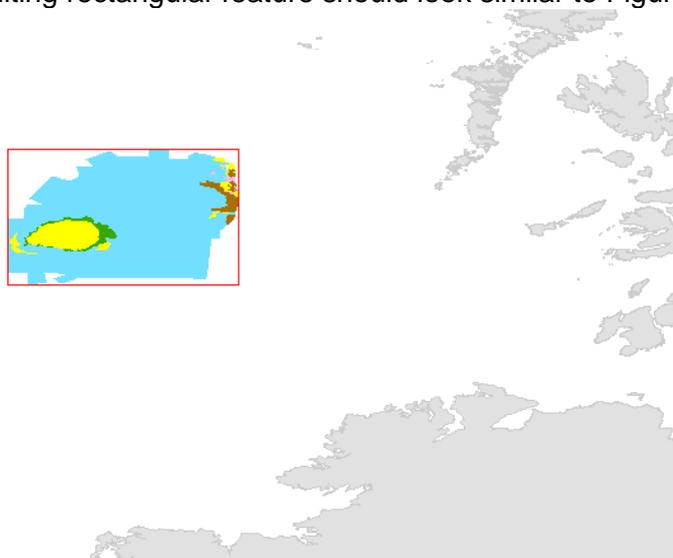


Figure 1: An example habitat map (solid colours representing habitat types) and its respective Study Area polygon (red outline) viewed in unprojected WGS84.

The study area feature class should contain only one feature and must be converted into the **Study Area DEF** (described in **Data Exchange Formats.pdf**).

The feature should be saved as a shapefile with the naming format of “[GUI]_StudyArea.shp”. For example, the Study Area for habitat map GB012345 should be saved as “GB012345_StudyArea.shp”

When these outlines are combined by EMODnet Seabed Habitats, the layer can be used to display map extents and confidence scores.

Python scripting

If you have ArcMap and are comfortable with python, the creation of habitat map study areas can be automated by using the python script `studyarea_creator.py` from the [GitHub repository](#).

The script should be run on a folder containing only habitat maps that are formatted and named in accordance to steps 2 and 3.

The script will produce a correctly named shapefile in the Study Area DEF. The GUI value will be filled in automatically, but the confidence score for the map must be entered into the SUM_CONF field manually.

Step 6: Create the metadata

For EMODnet Seabed Habitats, metadata will be entered into the online habitat mapping metadata catalogue maintained by the International Council for the Exploration of the Seas (ICES) on behalf of its Working Group on Marine Habitat Mapping (WGMHM). The main purpose of this site is for working group members to upload metadata about all habitat mapping activity that had occurred in the previous year.

You should supply your habitat map's metadata to your [country's partner](#) in .xml format. The metadata that you create should conform to the EU's [INSPIRE](#) metadata standard.

When creating your metadata, it is **critical** that the "Alternate title" is equal to the map's 8 digit GUI reference code. This is so that the interactive map can communicate with the metadata catalogue.

There are many tools available to help you write metadata.

Within ArcCatalog:

- Open ArcCatalog and navigate to shapefile requiring metadata
- Open ArcCatalog options, and ensure that "INSPIRE Metadata Directive" is selected as the Metadata Style inside of the Metadata tab
- Click on the *Metadata* or *Description* tab
- Click *Edit* and fill in metadata
- Save the metadata

Other local or online metadata editor tools:

You may be familiar with another tool for filling in metadata and exporting an XML file; if so you may use this. Ideally, it should produce INSPIRE-compliant metadata.

Step 7: Complete the data provider agreement

The supplied data provider agreement template (**EMODnetSBH Data Provider Agreement.doc**) should be completed for each habitat map to be supplied to EMODnet Seabed Habitats. Multiple habitat maps may be entered on one form, with each habitat map as a "Dataset" named by its GUI code.

If the dataset does not originate from your organisation, please enter the name of the owner in the relevant column so that they may be attributed properly.

Any usage limitations should be described in sections 2.3, and section 3 should be filled in if necessary.

Once completed, sign the form and include it within the data package to be sent to your country's EMODnet Seabed Habitats project [partner](#).

Step 8: Submit the data ‘package’

The GIS data (in ESRI shapefile format), confidence assessment spreadsheet, metadata (in .xml format) and signed data provider agreement should then be sent to your country’s project [partner](#). Please contact info@emodnet-seabedhabitats.eu and we will direct you to your country’s partner who will handle your data submission.

The partner will then periodically consolidate data packages and:

1. Enter metadata for each habitat mapping dataset into the ICES metadata catalogue
2. Send shapefile datasets corresponding to the metadata to JNCC for addition to the webGIS – habitat map layer and study area layer.
3. Send a single confidence assessment spreadsheet to JNCC for addition to the webGIS containing information for all maps in the package.
4. Send data provider agreement to JNCC for archiving.

JNCC will then upload the GIS and confidence data to the EMODnet Seabed Habitats webGIS.

The procedure is summarised in **Figure 2**.

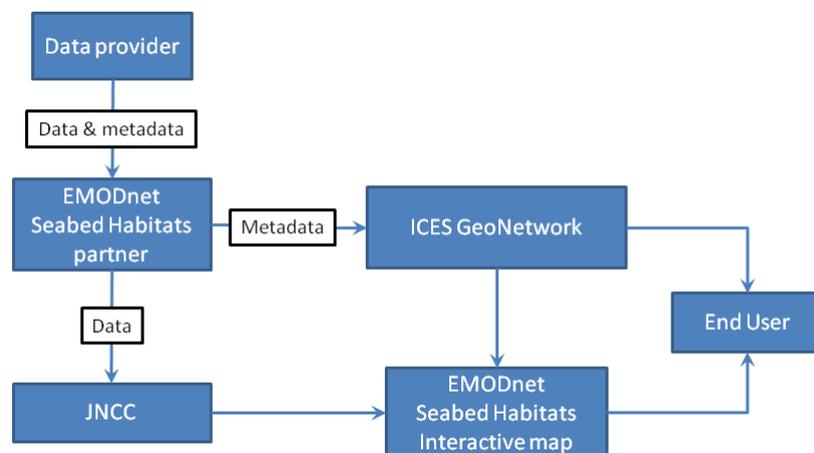


Figure 2: EMODnet Seabed Habitats webGIS data flow.