Submission process for contributing maps of seabed habitats and coastal wetlands to EMODnet

To contribute data to EMODnet Seabed Habitats, the data must be in a standardised ingestible format to allow it to be combined with datasets from other organisations.

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

The following steps should be undertaken when submitting maps of seabed habitats and coastal wetlands to EMODnet Seabed Habitats.

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.

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

If you are unsure about which GUIs are available, then please contact us and we will supply you with some GUI codes. The GUI will be used in the Shapefile's attribute table and as the 'Alternate Title' field in the metadata record created in Step 6.

2 Prepare the Shapefile

All Shapefiles provided to EMODnet Seabed Habitats **must be supplied in the WGS84 unprojected coordinate system (EPSG:4326)**. 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.

The supplied Shapefile should be named as its GUI. For example, a habitat map with GUI "IT000005" should be named as "IT000005.shp".

2.1 Cleaning the Shapefile

To ensure that your map data can be properly integrated in composite data products, and to allow the data to display properly, the shapefile should be cleaned to remove any geometry or topology errors.

1. **Split multi-part polygons** – this is beneficial so that each individual shape can be given a unique reference number. (ArcGIS tool: *Multipart to Singlepart*)
2. **Check topology** – topology refers to the geometry of polygons in a shapefile; some common topology errors are shown in Figure 1. Shapefiles with these errors are said to have complex topology, or to be unclean or non-simple (ArcGIS tools: *Check Geometry* and *Repair Geometry* tools)

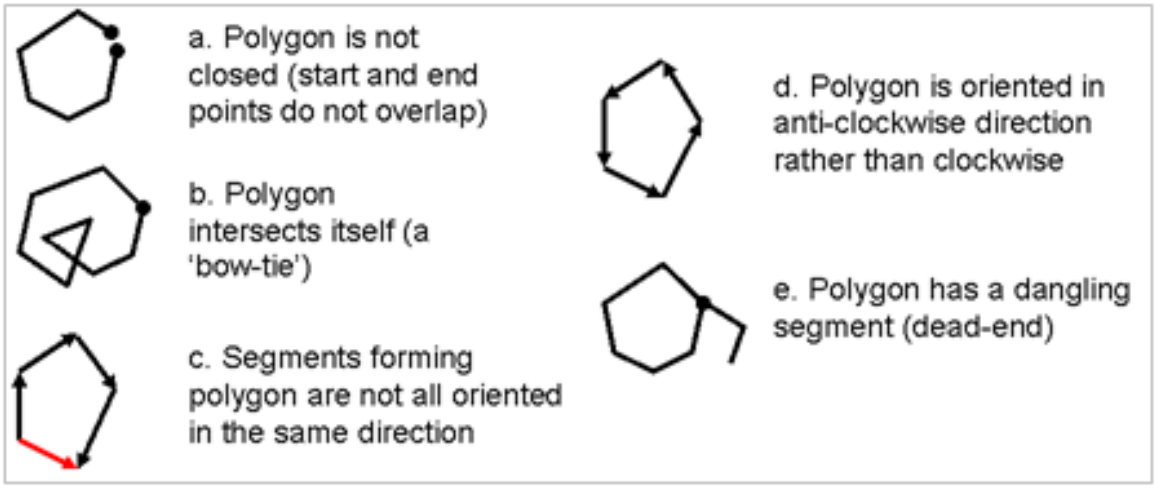


Figure 1: Illustrations of non-simple features according to ESRI™ topological rules for features.

3. **Check for overlaps** between polygons in a shapefile:

   The following procedure is only possible with a ArcEditor or Advanced licence in ArcGIS:

   a) Import map into a file geodatabase dataset

   b) Create a topology within the dataset for the imported map with the rule "no overlaps"

   c) A layer is created showing overlapping areas in red[1].

   A similar alternative solution is possible with QGIS's Topology Checker plugin[2]:

   If any overlaps are identified that are errors and should not exist, then these should be removed.

2.2 Areas With More Than One Habitat Type

In some cases, you may have overlapping polygons that represent the fact that more than one habitat has been assigned for a single area. For example, where more than one habitat is known to exist in a heterogeneous area, but cannot be delineated, or where two habitats may exist as a result of vertical stratification (Figure 2).

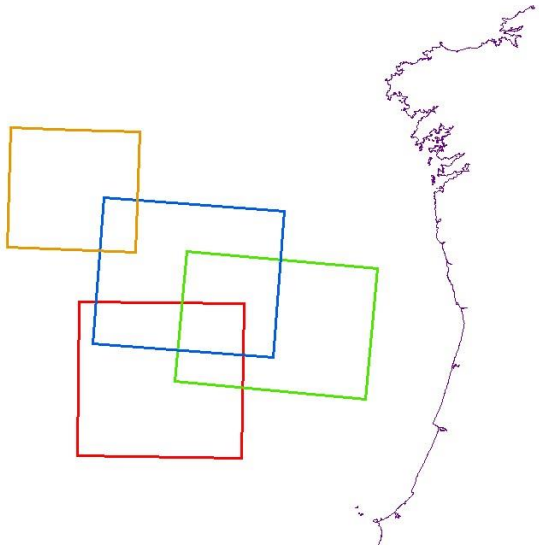


Figure 2: Simplified example of 4 habitat polygons overlapping, with a three-way overlap between red, green and blue.

In order to maintain compliance with the EU INSPIRE data specification for habitats and biotopes, habitat values must refer to only a single habitat type. As an example, 'A1' is valid, but 'A1+A2' is not.

[1] https://desktop.arcgis.com/en/arcmap/latest/manage-data/editing-topology/geodatabase-topology-rules-and-topology-error-fixes.htm

[2] https://docs.qgis.org/3.4/en/docs/user_manual/plugins/core_plugins/plugins_topology_checker.html

The following section provides you with a method to deconstruct overlapping habitat records, and features with multiple habitat values, and where possible, use your expert judgment or additional data sources to provide extra context within overlapping areas.

2.2.1 Conditions that submissions with overlapping habitat data must meet

To enable end-users to handle overlapping habitats within a single map, and to comply with INSPIRE, new submissions will only be accepted if they adhere to the following conditions:

1. Each feature within the dataset must be assigned only one habitat;
2. If two or more features overlap, they must be spatially identical; partial overlaps are not permitted;
3. If two or more features are spatially identical, they must be tagged with identical "POLYGON" identifiers to be identified as a group;
4. Each area containing more than one habitat must define the relative composition between the described habitats within the group in the following way:
   (a) The "COMP" field describes the composition of the habitats within the group. Choose one of the following:
      i. If the proportion of each habitat within the group is known, enter a decimal value in the "COMP" field for each feature within the group, with values ranging from 0-1 determining the proportion of each habitat. Total value for each polygon group must be 1.
      ii. If the proportions are unknown but you know which is the predominant (or primary) habitat, then enter the value 'primary' for the primary habitat polygon, and 'secondary' for all other polygons in the group.
      iii. If the composition is unknown , enter the value 'unknown' in the "COMP" field.
   (b) The "COMP_TYPE" field describes the type of composition for the habitats within the overlapping area. One of the following options should be chosen per group:
      i. 'heterogeneous' - the habitats contained within the group are both dispersed throughout the described area, but the individual patches have not been delineated.
      ii. 'transition' - there is a general trend of change from one habitat to the other(s) across the area. However, the threshold where one habitat changes into another cannot be accurately delineated.
      iii. 'data inconclusive' - based on the available survey data, one or more of described habitats may exist within the area, but cannot be confidently attributed. This may, for example, occur where two habitats are not identifiable from the underlying acoustic data on which a habitat map is based.
      iv. 'single habitat' - there exists only one habitat in the area (no overlaps). *If this option is chosen, the value in "COMP" must be '1.0'*.
      v. 'no information' - no information was provided by the habitat creator as to the composition of the habitats within the area. *If this option is chosen, the value in "COMP" must be 'unknown'*.

In order to achieve this, for every area where two or more habitats co-occur in the dataset:

1. For any partial overlaps, the overlapping area must be intersected to create spatially identical 'partial' polygons for the overlapping area, with one polygon per habitat;
2. Each duplicate *partial polygon* must retain the standard EMODnet DEF attribute information;
3. If a single feature originally describes more than one habitat (e.g. "A5.2+A5.3"), the feature must be split into spatially identical duplicates; one duplicate for each of the habitats originally described. All duplicates retain the same "POLYGON" identifier and must define the relative composition as above.

For futher tips see Appendix 1.

Suggested workflows are presented in Appendix 2 for:

1. How to correct partially overlapping polygons in ArcGIS
2. How to correct multiple habitats within a feature in R and/or GIS

3 Standardise the Attribute Table

Table 1: To save you time, we have created some Shapefile templates and Python scripts to automate the formatting process - explained further below.

Type	Name	Last modified
Template	Data exchange format (.shp)	01/03/2024
Tool	Python scripts for formatting to the Data Exchange Format (on Github)	25/07/2022

To enable 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 a Data Exchange Format (DEF).

2022 update

Before July 2022 there were 3 separate DEFs, depending on the classification system of the habitat map:

- If the map could not be translated into the EUNIS classification system: Original Habitat DEF.
- If the map could be translated into the EUNIS classification system: Translated Habitat DEF.
- If the map described Habitats Directive Annex I habitat types: Habitats Directive DEF.

Over time, it emerged that it would also be useful to be able to translate into other standard classification systems and vocabularies, including regional classification systems (e.g. HELCOM Underwater Biotopes) and priority lists (e.g. OSPAR threatened and/or declining habitats). Furthermore, the Habitats Directive DEF did not include any 'original habitat assignment' field for cases in which the Habitats Directive Annex I habitat type was a result of a translation.

In addition, since 2021 EMODnet has expanded its thematic scope to include coastal wetlands and geographical scope to include European territories in the Caribbean. This further emphasised the need to create a more flexible, future-proof DEF that does not enforce translation to a single classification system (EUNIS).

So, since July 2022, ***there is now a single, consolidated DEF for all classification systems, including coastal wetlands***. Specific changes in this version are summarised in Appendix 3 (summary of changes to the DEF), with a direct comparison between fields in the old and new DEFs.

Table 2 describes the fields that form the DEF for seabed habitat maps and coastal wetland maps in any classification system. A list of the most common classification systems is given in Table 3, with instructions about how to format the text for the attribute table.

Note:

1. Fields are marked as either:
    a. M – mandatory: the field must be completed for every polygon
    b. C – conditional: the field is mandatory in certain circumstances, as described in the table

    All other fields are optional, but we encourage you to complete as many of them as possible.
2. It is no longer necessary to include fields that contain no information.
3. All other non-standard fields (i.e. those not included in Table 2) must be deleted from the attribute table before submission.

Table 2: DEF for seabed habitat and coastal wetland maps. All fields marked "(M)" are mandatory and must be completed for the submission to be accepted. Pay close attention to conditional "(C)" fields, which are mandatory in the circumstances described in the table.

Field name	Data type	Length/ Precision	Description
GUI (M)	Text	8	A globally unique identifier (GUI) for the map. Consists of 2 letter country code (which corresponds to ISO3166-1) plus 6 digits. For example, a dataset from the United Kingdom would be written **GB000005**. This value can be obtained from your country's project partner and should be the same for all features within a habitat map.
POLYGON (M)	Text	10	Permanent identifier unique for each polygon within a dataset (or polygon group in the case of spatially identical polygons reflecting a composite habitat – see Section 2.2). Can be created as ascending integers 1,2,3... etc. Or a string of text characters and/or integers. This label for each polygon is necessary to identify the original polygon because the FID field may change during the processing of datasets.
ORIG_HAB (C)	Text	255	***Mandatory if:*** • *ORIG_HAB is not identical to HAB_TYPE* • *HAB_TYPE is empty, because it has not been possible to convert ORIG_HAB to any standard classification system/ vocabulary.* Original habitat type assigned to the polygon. This can be either a code or a text description of the habitat. It may be in a standard classification or not. If it is in a standard classification system, then refer to **Error! Reference source not found.** for how to format the information, e.g. whether to use a code or a name.
ORIG_CLASS (C)	Text	150	***Mandatory if*** *ORIG_HAB is not empty* The name of the habitat classification system or vocabulary that defines the term used in ORIG_HAB. If a standard classification system/vocabulary has been used, then refer to **Error! Reference source not found.** for how to format the text.
HAB_TYPE (M)	Text	100	Habitat type in one of the standard classification systems/vocabularies included in **Error! Reference source not found.**. For example, if the classification system is the 2022 version of the EUNIS habitat classification, then HAB_TYPE must be of the form "MA1".
HAB_CLASS (M)	Text	150	Name of the habitat classification system/vocabulary that defines the term used in HAB_TYPE – choose from one of those included in **Error! Reference source not found.**. E.g "EUNIS version 2007-11"
HABSUBTYPE	Text	150	This is only relevant for habitat lists where sub-types are defined, such as OSPAR. E.g. where HAB_TYPE = "Intertidal mudflats", HABSUBTYPE may be either "Marine intertidal mudflats" or "Estuarine intertidal mudflats". To find out which classification systems/vocabularies this applies to, see the 'Is HABSUBTYPE relevant for this classification system?' column in **Error! Reference source not found.**.

DET_MTHD	Text	255	If HAB_TYPE was derived by translation from another classification (ORIG_HAB), enter a short description of the method used to translate the original habitat map. Examples include: • **Manual translation**: Each EUNIS habitat within the map was manually determined and entered by an expert. • **Correlation table**: The translation was performed automatically using known correlations between the original and EUNIS classification systems.
DET_NAME (M)	Text	255	The name of the organisation or institute who determined the HAB_TYPE.
DET_DATE (M)	Date	-	Date that HAB_TYPE was determined.
TRAN_COM	Text	255	This is only relevant if HAB_TYPE was derived by translation from another classification (ORIG_HAB). Enter any comments on the translation from ORIG_HAB to HAB_TYPE. For example, include brief information from survey reports which justifies the translation decision (especially pertinent if the relationship between the data in the ORIG_HAB field and in the HAB_TYPE field is not clear). Also include reasons for assignment of a particular target habitat type, such as the volume and type of additional data used, and use of expert judgement.
T_RELATE (C)	Text	1	***Mandatory if*** *HAB_TYPE was derived by translation from another classification (ORIG_HAB).* Define the relationship between ORIG_HAB and HAB_TYPE. This information is potentially helpful to a user viewing the map, to determine the exact nature of the habitat relationship, especially if the translation is inexact (any symbol other than "=" or "S"). The relationship *must* be described by a one-character symbol chosen from **Error! Reference source not found.**.
VAL_COMM	Text	255	Record any polygon specific comments resulting from the validation of the map using an independent dataset; for example you may judge that there are spatial errors within the map (sublittoral habitat types appearing in the littoral zone and vice versa). Where possible, all translated maps should be validated with independent additional dataset(s). It is not possible to know whether inconsistencies are due to errors in the validation data, the original map, or the translation process. However, any suspected errors or disagreements between different datasets should be highlighted.
COMP (M)	Text	10	A description of the composition of the habitats (as defined in HAB_TYPE) within polygon groups (see Duncan, 2017[3]). **If the polygon is not within a polygon group, the value should be '1.0'.** If the polygon forms part of a polygon group, choose from one of the following options: • If the proportion of each habitat within the group is known, enter a decimal value in the "COMP" field for each feature within the group, with values

[3] https://emodnet.ec.europa.eu/sites/emodnet.ec.europa.eu/files/public/step2_guidance_dealingwithoverlaps_v1_2.pdf

			ranging from 0-1 determining the proportion of each habitat. Total value for each polygon group must be 1. • If the proportions are unknown but you know which is the predominant (or primary) habitat, enter the value '**primary**' for the primary habitat polygon, and '**secondary**' for all other polygons in the group. • If the composition is unknown, enter the value '**unknown'** in the "COMP" field.
COMP_TYPE (M)	Text	20	The type of composition for the habitats within the polygon group (Duncan, 2017[3]). **If the polygon is not within a polygon group, the value should be 'single habitat'.** If the polygon forms part of a polygon group, choose from one of the following options per group: • '**heterogeneous**'- the habitats contained within the group are both dispersed throughout the described area, but the individual patches have not been delineated. • '**transition**' - there is a general trend of change from one habitat to the other(s) across the area. However, the threshold where one habitat changes into another cannot be accurately delineated. • '**data inconclusive**' - based on the available survey data, one or more of described habitats may exist within the area, but cannot be confidently attributed. This may, for example, occur where two habitats are not identifiable from the underlying acoustic data on which a habitat map is based. • '**no information**' - no information was provided by the habitat creator as to the composition of the habitats within the area. If this option is chosen, the value in "COMP" must be 'unknown'.
SUM_CONF (C)	Short integer	-	***Mandatory if*** *a MESH confidence assessment has been carried out for the habitat map.* This is the value of the "Total" MESH confidence score given to the polygon or map during the data submission process.
TEXT_CONF	Text	100	A text description of the confidence in the presence and/or extent of the habitat described in HAB_TYPE. This may be a verbose description, or a simple classification of, e.g.: • High • Potential

Table 3: List of standard classification systems/vocabularies and how to format the information for the ORIG_HAB, ORIG_CLASS, HAB_TYPE and HAB_CLASS fields in the DEF for seabed habitat maps and coastal wetlands (Table 1). Note that ORIG_CLASS describes the classification system of ORIG_HAB and HAB_CLASS describes the classification system of HAB_TYPE.

Standard terms for ORIG_CLASS and HAB_CLASS fields	List of accepted habitat/wetland types for ORIG_HAB and HAB_TYPE fields	Accepted format for terms in ORIG_HAB and HAB_TYPE	Is HABSUBTYPE relevant for this classification system?
EUNIS version 2007-11	https://eunis.eea.europa.eu/habitats-code-browser.jsp We currently accept terms from the **Marine habitats** section (codes starting with **A**) and **Coastal habitats** section (codes starting with **B**)	Code only, not including the name e.g. **A1.1**	No
EUNIS version 2022	https://eunis.eea.europa.eu/habitats-code-browser-revised.jsp We currently accept terms from the **Marine benthic habitats** section (codes starting with **MA-MG**) and **Coastal habitats** section (codes starting with **N**)	Code only, not including the name e.g. **MA1**	No
Habitats Directive Annex I Habitats	https://eunis.eea.europa.eu/habitats-annex1-browser.jsphttp://dd.eionet.europa.eu/vocabulary/biodiversity/n2000habitats/ We currently accept terms from the **Coastal and halophytic habitats** group (codes starting with 1) and **Coastal sand dunes and inland dunes** group (codes starting with 2), plus **8330 Submerged or partially submerged sea caves**.	Code only, not including the name e.g. **1110**	Yes, although there is no EU-wide standard list of sub-types, so free text is allowed.
Marine Strategy Framework Directive Benthic Broad Habitat Types	http://dd.eionet.europa.eu/vocabulary/msfd/broadHabitatTypes/view	Name only e.g. **Littoral rock and biogenic reef**	No
Ramsar Classification System for Wetland Type	https://www.cbd.int/doc/meetings/sbstta/sbstta-08/information/sbstta-08-inf-04-en.doc (page 14)	Code and name e.g. **F Estuarine waters**	No
HELCOM Underwater biotope and habitat classification system (HELCOM HUB)	https://helcom.fi/baltic-sea-trends/biodiversity/helcom-hub/	Code only, not including the name e.g. **AA**	No
Classification of Benthic Marine Habitat Types for the Mediterranean Region	https://www.rac-spa.org/sites/default/files/doc_fsd/habitats_list_en.pdf	Code only, not including the name e.g. **MA1.5**	No

Standard terms for ORIG_CLASS and HAB_CLASS fields	List of accepted habitat/wetland types for ORIG_HAB and HAB_TYPE fields	Accepted format for terms in ORIG_HAB and HAB_TYPE	Is HABSUBTYPE relevant for this classification system?
OSPAR List of Threatened and/or Declining Species and Habitats	https://www.ospar.org/work-areas/bdc/species-habitats/list-of-threatened-declining-species-habitats/habitats	Name only e.g. **Carbonate mounds**	Yes[4]
HELCOM lists of threatened and/or declining species and biotopes/habitats in the Baltic Sea area	http://helcom.fi/Lists/Publications/BSEP113.pdf	Name only e.g. **Seagrass beds**	No
Marine Habitat Classification for Britain and Ireland v22.04	https://mhc.jncc.gov.uk	Code only, not including the name e.g. **LR**	No

[4] Accepted terms for HABSUBTYPE where HAB_CLASS = "OSPAR List of Threatened and/or Declining Species and Habitats":

For HABTYPE = "Intertidal mudflats":

"**Marine intertidal mudflats**" or "**Estuarine intertidal mudflats**"

For HABTYPE = Sabellaria spinulosa reefs:

"**Sabellaria spinulosa reefs on rock**" or "**Sabellaria spinulosa reefs on mixed (sediment) substrata**"

For HABTYPE = "Kelp forests":

"**Kelp forests dominated by Alaria esculanta**" or "**Kelp forests dominated by Laminaria digitata**" or "**Kelp forests dominated by Laminaria hyperborea**" or "**Kelp forests dominated by Laminaria ochroleuca**" or "**Kelp forests dominated by Saccharina latissima**" or "**Kelp forests dominated by Saccorhiza polyschides**" or "**Kelp forests dominated by another species**"

For HABTYPE = "*Zostera* beds":

"**Zostera marina beds**"or "**Zostera noltii beds**"

If sub-type is unknown, or for all other habitats, leave blank.

Table 4: Translation relationship types, and which codes are accepted for the T_RELATE field.

T_RELATE symbol	Translation relationship	Example, where HAB_CLASS = "EUNIS version 2007-11"
=	ORIG_HAB is the same as HAB_TYPE.	**ORIG_HAB:** "Circalittoral rock in a high energy environment" **HAB_TYPE:** A4.1
~	ORIG_HAB is nearly the same as HAB_TYPE.	**ORIG_HAB:** "Circalittoral rock in an energetic environment" **HAB_TYPE:** A4.1
>	HAB_TYPE is contained within ORIG_HAB (i.e. ORIG_HAB has a broader definition).	**ORIG_HAB:** "Saltmarsh" **HAB_TYPE:** A2.5
<	ORIG_HAB is contained within HAB_TYPE (i.e. HAB_TYPE has a broader definition).	**ORIG_HAB:** "High energy circalittoral rock with faunal communities" **HAB_TYPE:** A4.1
#	The definition of the ORIG_HAB partially overlaps with that of the HAB_TYPE.	**ORIG_HAB:** "*Potamogeton pectinatus* community" **HAB_TYPE:** A5.542
S	ORIG_HAB is the source of the HAB_TYPE.	*Similar to "=". Use when the translated EUNIS habitat was created as a result of a successful submission of the original habitat.*

Please see the "Converting maps" section of the archived MESH Mapping Guide for further guidance about how to translate habitat maps, with a focus on translating into the EUNIS classification.

4 Confidence Assessment

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

EMODnet Seabed Habitats uses the ***"MESH confidence assessment"*** method for assigning a confidence value determined by the processes and data that were used to create the habitat map. This was developed during the ***M***apping ***E***uropean ***S***eabed ***H***abitats project which ran from 2004 to 2008.

Important changes from 2020

The MESH confidence assessment is no longer an essential requirement for habitat maps submitted to the project, but it should be seen as **best practice**.

Although the method was designed for assessing habitat maps as a whole, we now accept polygon-level assessments for situations where the confidence varies across a single habitat map.

4.1 Confidence Assessment Guidance

Table 5: This link shows the template used for scoring the confidence assessment

Type	File	Last Modified
Template	Confidence assessment score sheet	24/03/2022

In the Excel template (Table 5), the first sheet ("Score sheet") is the only sheet requiring input, and the scores for the categories should be entered in columns D to R, with the corresponding map or polygon reference in column A.

If a confidence assessment can be applied to each polygon, please supply the map GUI and Polygon number to columns A and B - the PolygonID field will automatically populate. If a confidence assessment can only be applied to the habitat map as a whole, please supply the GUI and leave the Polygon Number column blank.

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.

5 Create a study area polygon

For each habitat map (GUI), a rectangular "Study Area" geographic bounding box should be created, encompassing the surveyed areas. The sides of the bounding box should be precisely horizontal and vertical when viewed unprojected in WGS84. In ArcGIS, this can be performed 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 3.

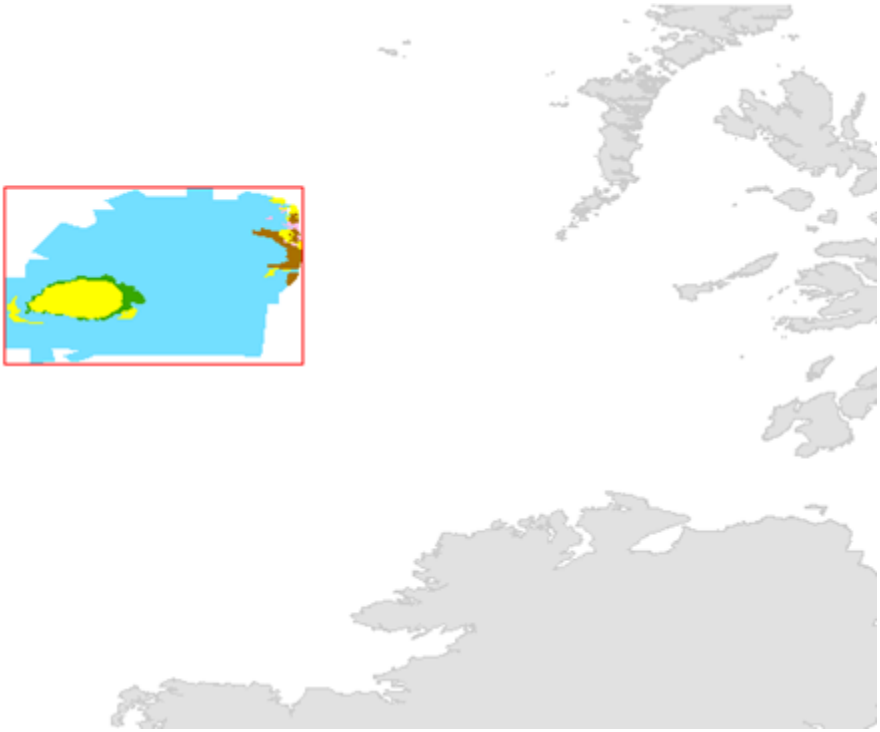


Figure 3: An example habitat map (solid colours representing habitat types) and its respective Study Area polygon (red outline) viewed in unprojected WGS84.

The study area shapefile should contain only one feature and must be supplied in the Study Area Data Exchange Format (Table 6). 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 link to the map metadata.

A Study Area shapefile should be supplied with each habitat/coastal wetlands map. This should be a rectangular bounding box describing the extent covered by the habitat map. The attribute table should be formatted according to the Study Area DEF (Table 6).

*Table 6: Study Area data exchange format (DEF). *If the habitat map does not yet have valid metadata entered on the ICES Geonetwork portal, then your country's partner will upload your metadata and enter the UUID value as required as part of the ingestion process.*

Field name	Data type	Length/ Precision	Description
GUI (M)	Text	8	Unique reference for the study. This must match the GUI of the habitat map to which the study area relates.
UUID (M*)	Text	36	UUID of ICES Geonetwork metadata record in the form of: xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx If the habitat map already has valid metadata on the ICES Geonetwork portal, then the "UUID" value should be entered. This value can be found in two ways: 1. The value after "uuid=" in the URL of the permanent link to the metadata record, for example the UUID of this record is 26f527ac-41f2-4c86-a443-3f655723efdf. 2. The value of "File identifier" within the ICES Geonetwork metadata itself. **Warning:** The UUID must be entered exactly as stated in the metadata and must conform to the pattern xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx (including all dashes) to pass validation.
AVAILABLE (M)	Text	13	Availability of habitat data. The value must be one of the three options: • "View/Download" – fully available to the public • "View only" – only available to view on interactive map or by Web Map Service (WMS) • "Not available" – habitat map not available, only metadata and bounding box.

5.1 Automating using Python

If you have ArcGIS and are comfortable with python, the creation of habitat map study areas can be automated by using the python script "**studyarea_creator_v2020.py**" available in our GitHub repository.

The script should be run on a folder containing only habitat maps that are formatted.

The script will produce a correctly named shapefile in the Study Area DEF. The GUI value will be filled in automatically.

6 Create the metadata

Table 7: This link provides the guidance for the INSPIRE-compliant metadata

Type	File	Last Modified
Template	INSPIRE-compliant metadata for habitat maps	30/07/2022

For EMODnet Seabed Habitats, metadata is entered into the online habitat mapping metadata catalogue maintained by the International Council for the Exploration of the Seas (ICES).

You should supply your habitat map's metadata 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 map viewer can communicate with the metadata catalogue.

There are many tools available to help you write metadata, such as the template and guidance (see Appendix 4).

7 Complete the data provider agreement

Tabel 8: This links to the Data Provider Agreement template which can be used as a guide for this step

Type	File	Last Modified
Template	Data Provider Agreement	10/01/2023

The **Data Provider Agreement template** should be completed for each habitat map to be supplied to EMODnet. 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.

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 "EMODnetSeabedHabitats@jncc.gov.uk", where your data will be validated. Data that does not pass the validation test will be returned to the sender to rectify.

For data passing the validation screening, datasets will periodically be consolidated and:

- Metadata for each habitat mapping dataset will be entered into the ICES GeoNetwork metadata catalogue
- Shapefile datasets will be added to the webGIS – including the habitat map layer and study area layer
- The data provider agreement will be archived

For any further questions, or queries about the submission process - please contact us.

9 Appendix 1: Tips for Translation into a Different Habitat Classification System

If a habitat map contains many (50+) polygons/features comprising of a number of habitats, for example in a highly heterogeneous environment, it is often easier to translate in a semi-automated way to reduce the time needed for the translation.

There are numerous methods for this; two possibilities within ArcGIS are described below.

1.1 Using the Field Calculator

The Field Calculator within ArcGIS can be used to automatically update values within a field.

If a selection is made from the original feature class or shapefile, the field calculator will only change the values of the selected features. This provides a powerful combination for changing multiple features sharing a characteristic, such as habitat type.

The following method can therefore be used if a relatively low number of habitats are dispersed amongst multiple polygons:

1. Open the attribute table for the habitat map.
2. Open the "Select by attributes" dialogue box.
3. Make sure the method is set to "Create a new selection"
4. Double click the field containing the original habitat information in the list below.
5. Click the equals button
6. Click the "Get Unique Values" button
7. In the list that appears, double click the original habitat to be translated and press "Apply"
8. Right click the "HAB_TYPE" field and select "Field Calculator" (press Yes on any message boxes)
9. In the text box at the bottom of the dialogue screen, type in the translated habitat matching the habitat selected in step 7 contained within double quotes (e.g. "A5.3" and not A5.3) and press "Ok"
10. Right click the "T_RELATE" field and select "Field Calculator"
11. In the text box at the bottom of the dialogue screen, type in the relevant symbol for the translation as described in Table 3 contained within double quotes and press "OK"
12. Repeat steps 2 to 11 for each original habitat contained in the map

Translation comments, translation methods, and any other field can be entered in the same way.

1.2 Using an Attribute Join

Alternatively, if the map contains both a large number of polygons/features AND a large variety of habitats, it may be quicker to undertake the following method:

1. Open the attribute table for the habitat map.
2. Export the attribute table as a dbf file.
3. Start ArcGIS's Delete Identical (Data Management) tool*.
4. Select the table created in **step 2** as the Input Dataset
5. Check the field containing the original habitat information and press "OK".
6. In the "HAB_TYPE" field of the table, enter a corresponding EUNIS habitat for each original habitat.
7. In the "T_RELATE" field of the table, enter a corresponding relationship symbol for the translation for each original habitat.
8. In the "TRAN_COM" field of the table, enter any comments relating to each habitat translation.
9. Save the edits and exit edit mode.
10. Right click the habitat map in the table of contents and select "Join" from the "Joins and Relates" sub-menu.

11. Select “Join attributes from a table” from the drop-down list, select the field as “ORIG_HAB”, the table as the table created in **step 2**, and the table field as “ORIG_HAB”.
12. Make sure “Keep all records” is checked, click “Validate Join” and if successful, click “OK”.
13. Once the join has happened, copy the values from the joined columns into the “HAB_TYPE”, “T_RELATE” and “TRAN_COM” columns in the shapefile using field calculator (newcolumn = oldcolumn).
14. Remove the join.

**If you do not have the Arc License to run the “Delete Identical” tool, then you can perform a similar process in Microsoft Excel below starting at step 3.*

3. Load the dbf table created in **step 2** into Excel.
4. Select all of the data and select Excel’s “Remove duplicates” tool.
5. Check the “my data has headers” box, and in the list below, check the box next to “ORIG_HAB” and press OK.
6. In the “HAB_TYPE” column, enter a corresponding EUNIS habitat for each original habitat.
7. In the “T_RELATE” column, enter a corresponding relationship symbol for the translation for each original habitat.
8. In the “TRAN_COM” column, enter any comments relating to each habitat translation.
9. Save the dbf file in Excel and load it into ArcMap. Continue with **step 10** above.

10 Appendix 2: Suggested Workflows in ArcGIS and R

10.1 How to Correct Partially Overlapping Polygons in ArcGIS

1. Run the 'Union' geoprocessing tool.
   - Select the habitat map layer as the only input feature.
   - Set "Join Attribute" to 'All'.
   - Leave "XY Tolerance" blank, the default setting.
   - Enable the "Gaps allowed" checkbox.

This will create a dataset with all overlaps converted to the correct number of intersection polygons – 1 polygon for each overlapping habitat. The polygons will still contain all the original habitat data. See Figu.

2. Run the 'Find identical' geoprocessing tool.
   - Select "Shape" as the only field.
   - Enable the "Output on duplicated records" checkbox.

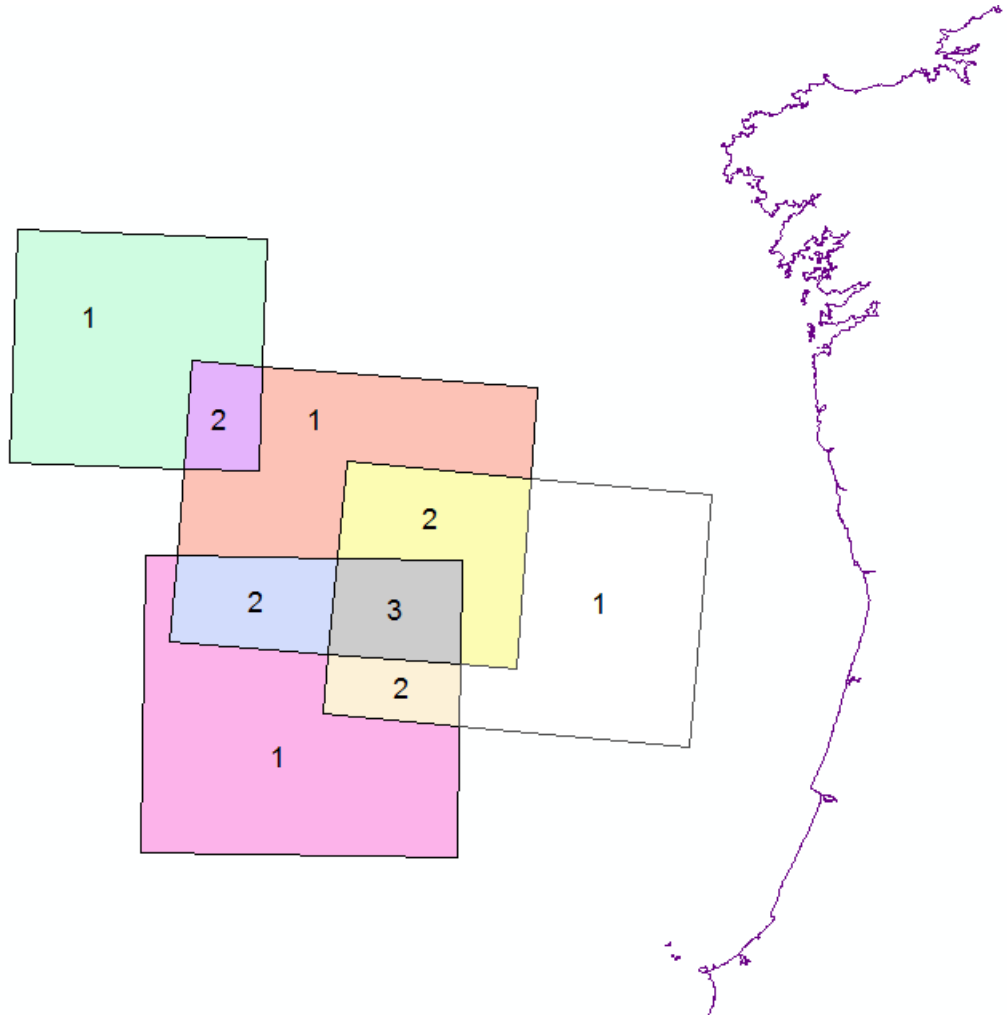


Figure 4: Results of using the 'Union' geoprocessing tool. Overlaps have been separated into their own discrete polygons (15 polygons in total). Numbers show number of spatially identical polygons 'stacked' in each location.

This will create a table showing you "Overlap groups" – the sets of polygons which come from each distinct overlap in the original dataset.

3. Run the 'Add join' geoprocessing tool.
   - Set "Layer Name" to the layer created in step 1.
   - Set "Input Join Field" to 'FID'.

- Set “Join Table” to the table created in step 2.
- Set “Output Join Field” to ‘IN_FID’.

4. Use ‘Field Calculator‘ on the field “POLYGON” with the following (VBScript) code: `[FEAT_SEQ]`

5. For each group of features with the same “POLYGON” value, the “COMP” and “COMP_TYPE” fields should be completed to satisfy the conditions in Section 2.2.1, item 4.

10.2 How to Correct Multiple Habitats Within a Feature in R and/or GIS

As stated above in Section 2.2, each feature (row) of a shapefile must contain only a single habitat type. Multiple habitats cannot be contained within the habitat fields. As an example, ‘A1’ is valid, but ‘A1+A2’ is invalid.

An example of a dataset containing multiple habitats which must be separated can be seen in Table. This workflow separates concatenated habitats into their own features, as spatially identical polygons, and, if possible, provides the user with information on the type of composition of the habitats within the area.

While the entire flow is possible in a GIS package, R’s data frame abilities make it significantly quicker in the separation steps and therefore use of this free and open-source software is highly recommended. EMODnet Seabed Habitats provides a pre-built R function to perform these steps.

Table 9: Example of input file with mulitple habitats per feature, without having undergone separation, note the use of ‘A+B’ type concatenation in the “ORIG_HAB” field. Note: This example is simplified; not all fields within the Data Exchange Format are shown.

GUI	POLYGON	ORIG_HAB	HAB_TYPE	T_RELATE
GB003021	1	CR.HCR.FaT+CR.MCR.EcCr	A4.11+A4.21	=
GB003021	2	CR.HCR.XFa	A4.13	=
GB003021	3	IR.HIR.KSed+IR.MIR.KR	A3.12+A3.21	=
GB003021	4	CR.HCR.FaT+CR.MCR.EcCr	A4.11+A4.21	=
GB003021	5	CR.HCR.FaT+CR.HCR.XFa	A4.11+A4.13	=
GB003021	6	CR.HCR.XFa+CR.HCR.FaT+IR.MIR.KR	A4.11+A4.13+A3.21	=
GB003021	7	CR.HCR.FaT	A4.11	=
GB003021	8	CR.HCR.FaT+CR.HCR.XFa	A4.11+A4.13	=

10.2.1 Initial Steps

The initial steps can be undertaken in the software of your choosing.

1. For each feature containing multiple habitats define the nature of the habitat composition in the “COMP_TYPE” field as stated in section 2 item 4.
2. If a feature contains only a single habitat, enter ‘single habitat’ in the “COMP_TYPE” field and ‘1’ in the “COMP” field.

10.2.2 Separating the Habitats - R

To undertake this section you will need a version of the free and open source software ‘R’. This can either be installed by itself, or as part of a more friendly development environment such as the free RStudio Desktop.

1. Download the “separate_habitats.R” script from EMODnet Seabed Habitats’ GitHub repository.

2. Open the script in edit mode and change the values between the quote “” marks on on the following lines:
   - Line 15: Change to the folder of your input shapefile.
   - Line 17: Change to the file name of your input shapefile.
   - Line 21: Change (if required) to the character that separates individual habitats in your text. The default value is ‘ + ’ , to be used in the case of ‘A+B’ where A and B are two habitats.
   - Line 23: Change (if required) to the field that contains your habitat values requiring separation. The default value is ‘ORIG_HAB’.
   - Line 26: Change to the folder where you would like your output shapefile saved.

3. Run the script within R - the result should be a new shapefile in your output directory. The name of the output file will be ‘split_’ followed by the name of your input file.

4. The output file will contain spatially identical polygons for each habitat in the original text.

5. If you separated your records based on the “ORIG_HAB” field, and are looking to translate your habitats into EUNIS, you can do so now.

10.2.3 Setting “COMP”

This step can be undertaken in the software of your choosing.

1. For each group of features with the same “POLYGON” value, the “COMP” field should be completed to describe the relative composition of habitats as defined in Section 2.2.1 item 4.
2. The attribute table should look similar to that shown in

Table 10: Example final attribute table for a submitted map with overlapping habitats. Note the identical “POLYGON” values for habitats formally contained within a single row in Table. NB: This example is simplified; not all fields within the Data Exchange Format are shown.

GUI	POLYGON	ORIG_HAB	HAB_TYPE	T_RELATE	COMP	COMP_TYPE
GB003021	1	CR.HCR.FaT	A4.11	=	unknown	data inconclusive
GB003021	1	CR.MCR.EcCr	A4.21	=	unknown	data inconclusive
GB003021	2	CR.HCR.XFa	A4.13	=	1.0	single habitat
GB003021	3	IR.HIR.KSed	A3.12	=	0.3	heterogeneous
GB003021	3	IR.MIR.KR	A3.21	=	0.7	heterogeneous
GB003021	4	CR.HCR.FaT	A4.11	=	0.5	heterogeneous
GB003021	4	CR.MCR.EcCr	A4.21	=	0.5	heterogeneous
GB003021	5	CR.HCR.FaT	A4.11	=	primary	transition
GB003021	5	CR.HCR.XFa	A4.13	=	secondary	transition
GB003021	6	CR.HCR.XFa	A4.11	=	0.1	heterogeneous
GB003021	6	CR.HCR.FaT	A4.13	=	0.3	heterogeneous
GB003021	6	IR.MIR.KR	A3.21	=	0.6	heterogeneous

GB003021	7	CR.HCR.FaT	A4.11	=	1.0	single habitat
GB003021	8	CR.HCR.FaT	A4.11	=	unknown	transition
GB003021	8	CR.HCR.XFa	A4.13	=	unknown	transition

11 Appendix 3: Summary of Changes to the DEF Since July 2022

Table 11: Comparison of field name in the new DEF and those in the old Original Habitat, Translated Habitat and Habitats Directive DEFs. Field names are highlighted red where they have changed.

NEW DEF (SINCE JULY 2022)	OLD ORIGINAL HABITAT DEF	OLD TRANSLATED HABITAT DEF	OLD HABITATS DIRECTIVE DEF
GUI (M)	GUI (M)	GUI (M)	GUI (M)
POLYGON (M)	POLYGON (M)	POLYGON (M)	POLYGON (M)
ORIG_HAB (C)	ORIG_HAB (M)	ORIG_HAB (M)	
ORIG_CLASS (C)	ORIG_CLASS	ORIG_CLASS	
HAB_TYPE (M)		HAB_TYPE (M)	ANNEXI (M)
HAB_CLASS (M)		VERSION (M)	
HABSUBTYPE			SUBTYPE
DET_MTHD		DET_MTHD	
DET_NAME (M)		DET_NAME (M)	
DET_DATE (M)		DET_DATE (M)	
TRAN_COM		TRAN_COM	
T_RELATE (C)		T_RELATE (M)	
VAL_COMM		VAL_COMM	
COMP (M)	COMP (M)	COMP (M)	
COMP_TYPE (M)	COMP_TYPE (M)	COMP_TYPE (M)	
SUM_CONF (C)		SUM_CONF (M)	
TEXT_CONF			CONFIDENCE (M)

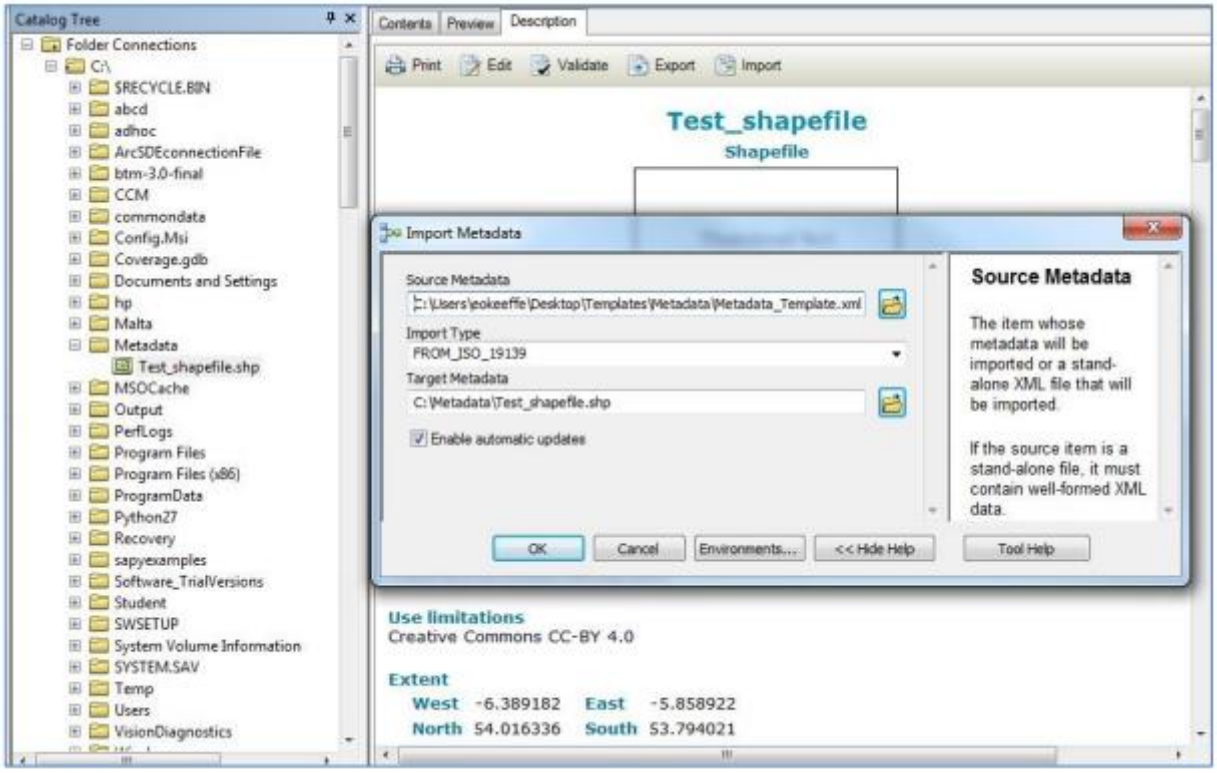
Apart from the change to field names, the other change is that it is no longer required to include all fields in the attribute table, if some of them contain no information.

12 Appendix 4: Creating INSPIRE-Compliant Metadata using ArcGIS

These steps outline how to create an INSPIRE-compliant metadata record using an xml file as a template in ArcCatalog[4].

12.1 Importing Metadata Template

Import "Metadata_Template.xml" to shapefile that you wish to write metadata for in ArcCatalog. To do this, highlight the shapefile in the ArcCatalog tree. Then, click on the Description tab on the main panel and select the Import option. Select "Metadata_Template.xml" as the input file and click OK.



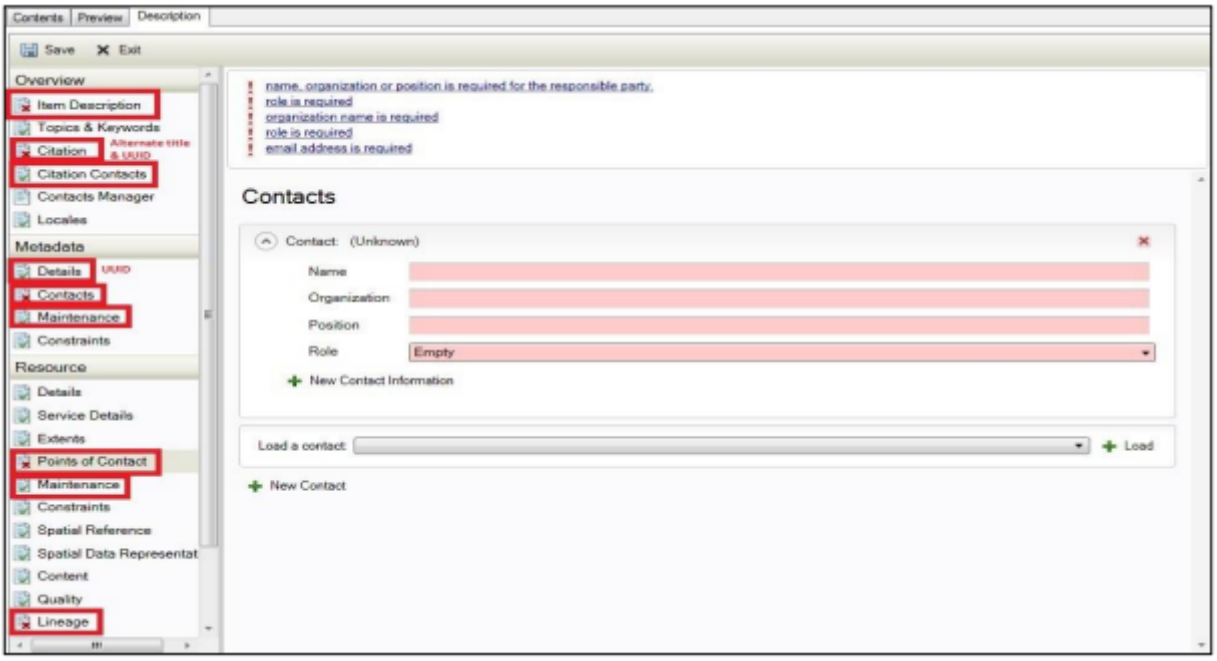
After a minute the import should be complete and a dialog box opens to confirm the import was successful.

12.2 Editing Unique Metadata Fields

Now that a metadata template has been added to the shapefile, additional metadata fields unique to each shapefile need to be input. Common fields including information on distribution licence, keywords, etc. are included in the template and automatically fill the required sections of the metadata file after import. Information on owner/point of contact, UUID and alternate titles are unique to every dataset and therefore have to be filled in manually after the import.

In order to fill in the additional metadata fields, click on the shapefile of interest (the one you just imported the metadata template to) and again click on the Description tab in the main panel. Then, select the Edit button. The fields highlighted all need data to be input.

[4]Eimear O'Keefe, Marine Institute (2019)
https://emodnet.ec.europa.eu/sites/emodnet.ec.europa.eu/files/public/step6_guidance_creatinginspirecompliantmetadata.pdf



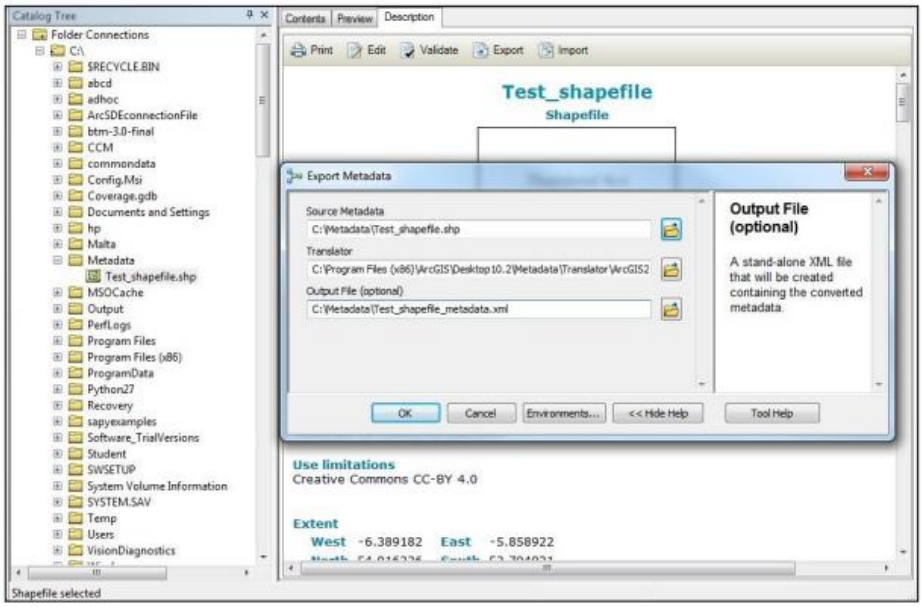
To summarise, the following data needs to be included for each section:

- Item description – title and abstract
- Citation – title, alternate title and UUID
- Citation contacts – add new contact, include name, organisation and role.
- Details – copy and [paste the UUID created in previous step.
- Contacts - add new contact, include name, organisation and role.
- Maintenance – select frequency of update from dropdown menu.
- Points of contact - add new contact, include name, organisation and role.
- Lineage – brief history on the origins of the data.

When complete, click Save.

12.3 Exporting Metadata

Finally, export the metadata from the shapefile into an .xml file. Highlight the shapefile in ArcCatalog, click on the Description tab on the main panel and then select Export. A dialogue box opens prompting you to provide a name for the output .xml file.



Select an output name for the .xml and click OK.

This .xml file is a fully INSPIRE-compliant metadata file and needs to be submitted along with all of the formatted shapefiles (DEFs) to EMODnet Seabed Habitats.