



EMODnet Seabed Habitats

EMODnet



European Marine
Observation and
Data Network

Data Submission Technical Guidance

Conforming to EMODnet and INSPIRE feature overlap rules in 'Habitat maps from survey' data submissions

Graeme Duncan, JNCC, UK
Revised: December 14, 2017
V1.2

1 Introduction

Data from EMODnet Seabed Habitats are likely to be used in carrying out processes that require *only one* habitat to be assigned to a particular location.

However, polygon habitat maps can contain geographically overlapping features in certain circumstances, 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. For a simplified representation, see Figure 1.

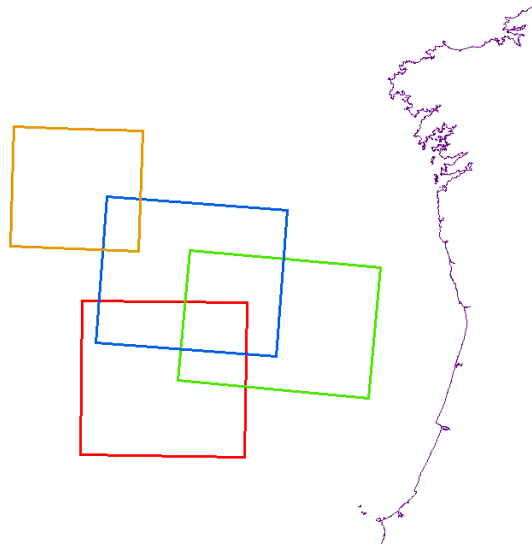


Figure 1: Simplified example of 4 habitat polygons overlapping, with a three-way overlap between red, green and blue.

Furthermore, 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.

The following document 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 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 to be made ready for submission:

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.

3 Suggested workflow for correcting partially overlapping polygons (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 Figure 2.

2. Run the ‘Find identical’ geoprocessing tool.
 - Select “Shape” as the only field.
 - Enable the “Output on duplicated records” checkbox.

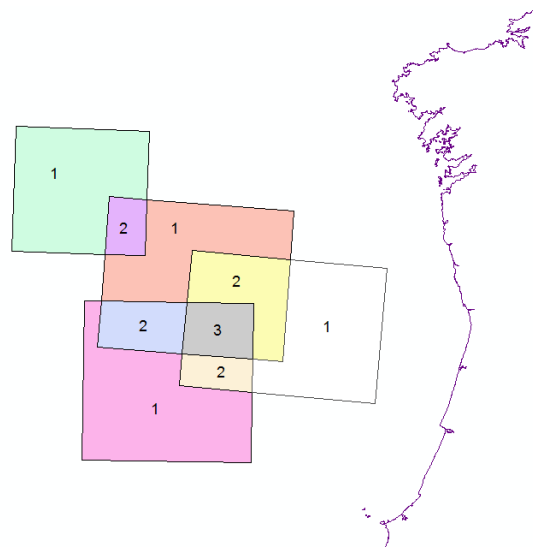


Figure 2: Results of using the ‘Union’ geoprocessing tool on Figure 1. Overlaps have been separated into their own discrete polygons (15 polygons in total). Numbers show number of spatially identical polygons ‘stacked’ in each location.

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:

```
1 [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 item 4.

4 Suggested workflow for correcting multiple habitats within a feature (GIS & R)

As stated above in section 1, 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 1. 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 an pre-built R function to perform these steps.

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	=

Table 1: Example of input file with multiple 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.

4.1 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.

4.2 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 desktop [R-Studio](#).

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 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.

4.3 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 item 4.
2. The attribute table should look similar to that shown in Table 2.

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.FaT	A4.11	=	0.1	heterogeneous
GB003021	6	CR.HCR.XFa	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

Table 2: 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 1. formatting of the “COMP” and “COMP_TYPE” fields. NB: This example is simplified; not all fields within the Data Exchange Format are shown.

5 And finally...

1. Remove all fields not present within the [EMODnet Seabed Habitats Data Exchange Format](#)