EMODnet Biology: Marine Species Traits Workshop

Suggested citation


Image © Fiona Crouch
Summary

This two-day workshop held at the Hellenic Centre for Marine Research (HCMR), Crete was the third in a series of European Marine Observation and Data Network (EMODnet) Biology meetings to develop a framework for the description, collation and dissemination of traits information relating to marine species.

With contributions from 19 experts with both scientific and technical backgrounds, a wide-range of relevant topics were presented and discussed across the breadth of traits information collated within EMODnet Biology, from those relating to the species morphology to its perceived ‘importance’ to society.

Following presentations highlighting previous work, how traits have been applied to answer science use-cases and international examples, the participants were split into 3 groups to discuss specific aspects of the work and to propose next steps required for the publication of the traits hierarchy, collation of existing traits values from disparate sources and the tools required to search, query and download traits-based information.

The agreed next steps included the imminent publication of the traits hierarchy in wiki format to encourage discussion and development of each of the trait definitions; progress on the completion of 10 the priority traits through the extraction and collation of existing trait material from sources including, but not limited to MarLIN, FishBase, SealifeBase, EoL, PolyTraits; and to develop a timetable and details of responsibilities for use-cases that will feedback on the development of the traits database and associated publications.

http://www.emodnet-biology.eu/
Participating Institutes

- Natural History Museum
- NIOZ
- VLIZ
- UMa
- Universidad de Málaga
- GBIF
- Encyclopedia of Life
- EOL
- HCMR
- MARUM
- NIWA
- Taihoro Nukurangi
- CarMS
  Canadian Register of Marine Species
**Introduction**

The second phase of the biological component of the European Marine Observation and Data Network (EMODnet) began September 2013, following the preparatory work of the previous 3 years. Led by the Flanders Marine Institute (VLIZ) the biological data actions involve a number of key, interlinked work-packages designed to deliver unparalleled access to marine biological datasets and to ensure their utility for effective, evidence-based decision making.

The project consortium is made up of 23 research agencies, marine laboratories and government agencies, with broad experience in the collation and management of biological data.

The project aims to collect and assemble data for all the European sea basins including the Black Sea, Mediterranean Sea, North East Atlantic, North Sea and Baltic Sea, with attention being specifically applied to coastal data sources. In addition marine biological monitoring data from the Norwegian Sea and Barents Sea will be made accessible. The relationship between the six work packages is presented in Figure 1.

![Figure 1. The EMODnet Biology Work Package structure](image)

Work Package 2 (WP2) of EMODnet Biology, the Identification and collection of species, species attributes and species indicator information, has 2 discrete components. The first focuses on the collation of species information relating to legislation and directives, including the Habitats and Birds
Directives, IUCN Red Lists, CITES and those species forming part of the proposed indicators within the Marine Strategy Framework Directive (MSFD).

The second aspect of WP2, relates to harmonising the collation of species attributes and biological traits, including the development of an agreed vocabulary of trait terms and associated definitions. A standard vocabulary will enable a greater degree of interoperability and facilitate the exchange of trait data between groups. A series of workshops dealing with Biological and Ecological Traits have been organised (http://www.emodnet-biology.eu/workshops) over the last 3 years. These workshops resulted in a list of priority traits. The criteria for prioritisation were that (1) the trait could (in theory) be applied across most taxa, (2) that information on the trait existed for most taxa, and (3) it was likely that the availability of the trait would result in new uses of the databases in the short-term. As there are arguments for more and fewer traits depending on user needs, we created a top-10 short-list. The World Register of Marine Species (WoRMS) is used as the taxonomic backbone to store the trait information. A paper dealing with the rationale for the prioritisation of traits for inclusion in WoRMS has been published (Costello et al., 2015).

It was felt timely to organise a follow-up workshop to 1) present the activities that have been achieved since the last workshop in Paris and EMODnet Biology partners meeting, 2) discuss the finalization and publication of the overall vocabulary, 3) discuss the approach to implement the 10 priority traits, and 4) discuss and identify use cases including workflows that combine trait, taxonomic and biogeographic data. The workshop was organised alongside the Steering Committee meeting of WoRMS.

The meeting began with presentations providing background to the EMODnet Biology project, the overall rationale behind the use and importance of biological and ecological traits and highlighting some of the trait data that has already been collated (as highlighted in Costello et al., 2015). This included those traits relating to introduced species and the development of the World Register of Introduced Marine Species (WRIMS) thematic portal. These presentations were followed with examples and demonstrations of how traits have been organised into vocabularies and thesauri within EMODnet Biology and LifeWatch Italy, and introduced how the Lifewatch VRE could support trait-based workflows. The scientific application of functional traits were then explored with presentations on how information from WoRMS, EMODnet and EuroBIS could be combined in the study of European nematodes and how the aggregation of trait information at different taxonomic levels can be applied in the North Sea.

The final set of presentations expanded the geographical scope, investigating the opportunities and challenges associated with developing a global database of planktonic traits and the latest
developments associated with the traits information held within the Encyclopaedia of Life and how EMODnet Biology can best interact with these initiatives.

In the afternoon the participants divided into three breakout groups, each focussing on a different aspect of biological and policy-based traits. The groups were:

1. Review vocabulary and approach and timeline to implement priority traits data
2. Tools, interfaces and workflows for analysis and visualization traits, taxonomic information
3. Identifying scientific use cases using available trait, taxonomy and distributions data

Attendees were split between these groups based on their expertise and interests. The majority of the remainder of the meeting were then spent in these groups, with a final plenary session reporting on the findings, recommendations and next steps.

Through careful co-ordination, the workshop hosts at HCMR were able to run both an EMODnet Work Package 4 (data archaeology) and EUBON meeting alongside the traits workshop. The final set of plenary presentations therefore allowed these three groups to explore the potential for technology to be utilised in the extraction of traits values from electronic resources using the most recent developments in text-mining. Using the agreed lists of traits and focussing on the ‘top 10’ priority traits the opportunities for collaboration between these 3 groups was discussed, and will potentially allow the rapid extraction of existing traits information from published pdf’s and electronic documents.

Summary of Breakout Groups

Breakout Group 1 - Review vocabulary and approach and timeline to implement priority traits data

Simon Claus (Chair), Geoff Boxshall, Jennifer Hammock, Harvey Tyler-Walters, Serge Gofas, Stéphane Pesant, Leen Vandepitte, Mary Kennedy, Nicolas Bailly, +Bart Vanhoorne (10/06), +Dan Lear (10/06)

This group focused discussion on the traits hierarchy and the 10 priority traits as defined at the Paris workshop (ref) and refined through Costello et al. (2015).

Each of the 10 priority traits was discussed in turn, with the current available content assessed and next steps, where required, defined.

1. Taxonomy

Taxonomic traits are already covered through the WoRMS taxonomic entries.

2. Environment

The tagging of all taxonomic entities with the value Marine, Brackish, Freshwater and/or Terrestrial has been identified by the WoRMS Steering Committee as a priority gap to be addressed. Currently around 10% of entries are missing this information, and this field will become mandatory in WoRMS to ensure no gaps occur in the future.
Those species currently tagged as ‘plankton’ or ‘nekton’ will also have the trait ‘pelagic’ assigned to them, as all planktonic taxa are, by definition pelagic. In a similar manner any species featuring in the World Register of Deep Sea Species will be tagged with the ‘deep-sea’ trait.

3. Geography

The geographic range of species can be defined from the distributional data held within OBIS. These distributions can be described using the standard geographic terms and boundaries from within MarineRegions.

4. Depth

Due to different taxonomic groups having varying definitions relating to depth, this trait generated significant discussion. The planned split using intertidal/subtidal/deep sea was judged to be inadequate. In the case of plankton epi-, meso-, bathy- and abbysopelagic were felt to be more useful.

When a quantitative value is given for depth then the most appropriate term can be retrospectively applied, however without clear definitions the reverse cannot be accomplished.

Discussion also centered around whether depth referred solely to water column depth, but also depth within sediment for benthic species. In addition, depth could only be applied subtidally and elevation values were required for seabirds etc.

Later discussion noted that the proposed pelagic depth zones have no empirical basis apart from the euphotic zone and are thus difficult to define except for the subtidal which is equivalent to the epipelagic. Similarly the WoRMS deep-sea group had not distinguished benthic depth zones. Thus for WoRMS the simpler classification was preferred, i.e. intertidal/subtidal(=epipelagic)/deep sea.

Nevertheless, including both a value for depth (below chart datum) and elevation (above chart datum) can be accommodated within the trait catalogue. Depth is clearly a priority, as mentioned above, is the basis for description of any of the zones often used in the literature. However, we can retain the ‘zones’ or ‘vertical zones’ within the catalogue so that we harvest all available information and the information used in the literature and by relevant taxonomic group experts.

The definition of ‘Deep sea’ requires further discussion within the catalogue. At present the deep-sea group have set a depth cut off of 500m. However, other classification systems, e.g. EUNIS and UK Deep Sea classification use either edge of continental shelf or >200m. The appropriate ‘cut-off’ point needs to be decided with the EMODnet Habitats and Habitat mapping groups to ensure consistency.

5. Body size

The collation of body size values was commenced through the traits pilot projects, and has been identified as a priority for European Register of Marine Species (ERMS) taxonomic entities.

Studies on some species (i.e. Nematodes) use both length and diameter so it is a requirement for several measurements to be permitted, alongside the units, which may also vary. In addition the methodology of the measurement would benefit the reuse value of the information.
Through consultation with the WoRMS taxonomic editors, potential sources of body length information will be reviewed. Currently around one third of the entries in ERMS have an entry for body length. This can be added to with information held within FishBase and SealifeBase and potentially through the analysis of existing images (primarily for Mollusca) held within WoRMS. Where additional body measurements are readily available in digital form they may be included in WoRMS (e.g. biomass, width, wing span). However, the priority with EMODnet will be compile maximum body length for as many species and taxa as possible.

6. Substratum

Whilst the Substratum trait is only applicable to benthic species there are still a wide range of potential values that can be assigned to this attribute. Where sufficient information is available EUNIS classifications should be assigned, however it is anticipated this level of information will not be readily available for most species. A simplified classification of soft, hard or biological can more easily be applied at higher taxonomic levels with these values cascading down using the existing WoRMS inheritance functionality, whilst still allowing for exceptions to be included at lower taxonomic levels.

It was recommended that there is some overlap with the Environment Ontology (ENVO²) and these links should be explored as the EMODnet traits hierarchy evolves.

7. Mobility

It was agreed that the term ‘parasite’ did not sit under ‘mobility’. Therefore the main split would simply be between ‘sessile’ and ‘mobile’, as all species (parasitic or other) have mobility traits.

Further granularity is listed within the hierarchy. The mobility displayed by a species is life stage dependent and so it is necessary to have the capacity to capture at which life-stage that specific trait is applicable.

8. Skeleton

The priority groups identified for this trait were Foraminifera, Crustacea, Mollusca, Echinodermata, Porifera, Cnidaria, primarily due to their potential response to an increase in water pH.

There exists the potential for species to have a combination of materials as such the requirement to select multiple values must be available. The options for this trait are chitinous, calcareous, silicious and hydrostatic. In addition if the information is available calcareous can be sub-divided to include aragonite, calcite or a combination.

9. Diet

It was recognized that it was desirable to disaggregate the feeding method from the food type, and, at the highest level, record food type/diet as one of autotroph, heterotroph or mixotroph.

As mentioned previously the feeding method can vary with life stage and as such the underlying data systems and user interface must allow for this feature.

² http://environmentontology.org/
10. Reproduction

The options discussed and agreed for this trait were; sexual, asexual and alternation of generations. Where possible those species that have been assessed as being important to society (through legislative, economic or environmental means) and benthic species should be prioritized. It was also agreed that the age at which the species reaches reproductive maturity is a key piece of information to attempt to capture.

Breakout Group 2 - Tools, interfaces and workflows for analysis and visualization of traits, taxonomic information

Dan Lear (chair), Francisco Hernandez, Bart Vanhoorne, Dmitry Schigel

This small group looked at the existing technological solutions available for taxonomic and traits information with a view to identifying gaps and assessing how existing solutions can be aligned and integrated to provide a holistic overview of species information. The systems evaluated included:

- Traits wiki – traits vocabulary, glossary of definitions and ontology development and export
- Aphia (WoRMS/ERMS and thematic portals) - taxonomy & traits
- (Eur)OBIS - species/taxon occurrences and distributions
- EMODnet biology & central portal – data products and species information

By improving the linkages within these systems it was agreed that greater utility could be achieved. Future developments would include the need for improved access to traits information through the WoRMS species pages, but with the requirement to ensure taxonomic pages do not become cluttered or confused with large volumes of traits-related information. The requirement for a simple “search by trait” feature was also discussed alongside whether a separate portal for traits information was desirable.

By using the existing EMODnet Biology portal and combining the taxonomy and traits information, Marine Regions gazetteer and occurrence data from EurOBIS/EMODnet it would be possible to develop a powerful, integrated tool to facilitate complex queries to answer the questions of researchers, policy makers and the general public.

Breakout Group 3 - Identifying scientific use cases using available trait, taxonomy and distributions data

Participants: Mark Costello (chair), Andrew Barton, Geoff Read, Stefanie Dekeyzer; Jan Vanaverbeke, Olivier Beauchard.

This group developed some of the potential scientific use-cases and questions that could be addressed using the traits information that is intended to be collated through EMODnet Biology and WoRMS.

Building on discussions from the Paris workshop and the expertise of the attending scientists the group initially identified key questions and then discussed what trait information would be required in order for answers to be obtained.
The use-cases focussed on 5 key areas:

1. **IAS**
   - Are they random sample of species in higher taxa?
   - Do any traits correlate with introductions and/or invasiveness?
   - By geographic distribution using place names (cord diagram)
   - Habitat (sediment, fouling/hard substratum, plankton)
   - Have they equal probability of coming from and going anywhere?

2. **Conservation status** (Red List, CITES, OSPAR, and EU Directives)
   - Are there traits associated with conservation status?
   - How does geographic distribution relate to conservation status?
   - Null model is distribution of traits across all species and/or taxa in WoRMS.

3. **Body size by species richness**
   - What is relationship between body size and species richness across taxa?
   - Compare by taxonomic group (phylum-order), pelagic, benthic, introduced, invasive, year of description, number locations (e.g. 5-degree cells), geographic range (from OBIS?)?
   - Use min and max value for range of body size in higher taxon but recognise it may not be normal distribution

4. **Latitude**
   - How do traits map geographically, by latitude and by depth?
   - Species size and richness by latitude with OBIS data
   - Matrix of species richness (as number or %) by log2 body size – is distribution across sizes even or similar between latitudes or geographic areas?
   - Compare groups by taxon, pelagic/benthic, for generality

5. **Combining traits to create functional groups**
   - To develop based on assessment of what traits are available.

Specific actions for group members were identified to enable the development of these use cases, primarily focussing on the collation of trait values from a range of existing sources and the timely aggregation of these values within the WoRMS infrastructure.

To do

1. Look for trait information in Specimen Form, Type Locality, and Notes field e.g. depth [Stefanie]
2. Add traits from FishBase, SeaLifeBase, PolyTraits, Nemys, MarLIN, Olivier Data, EoL [Stefanie]
3. Look at what traits we have and their patterns [Stefanie]
4. Contact WoRMS Editors who wish to collaborate by adding traits and co-author, comment that we are checking Notes and other fields for traits [Mark]
5. Add traits at higher taxon level from Brusca & Brusca or as advised by collaborating editors.
6. Native ranges for IAS into WRIMS [Mark, Shyama]
7. Check length good metric of biomass (e.g. Nikalas 19xx., others)
8. Can we assign trophic levels to all taxa and species? [Mark check FishBase manual]
9. Begin framework for papers [Mark – other volunteers?]
Next Steps

Following each of the breakout groups reporting their discussions, the meeting concluded with a summary of immediate actions and the next steps.

1. The publication of the traits hierarchy is a specific deliverable of the EMODnet Biology work plan and will be made widely available by the deadline of September 2015 (month 24 of the EMODnet Biology Project). Prior to this additional work is required to ensure the wiki instance of the hierarchy is free of typographical errors, associations between terms and their definitions are clear and there is a common writing style across the definitions. In addition a peer-reviewed publication explaining the rationale behind the development of the traits hierarchy will be initiated. Experts will be invited to contribute to the definitions of each of the terms using the ‘Discussion’ features of the wiki application.

2. In parallel to this editorial work, the EMODnet Biology WP2 will progress on the completion of 10 the priority traits through the extraction and collation of existing trait material from sources including, but not limited to MarLIN, FishBase, SealifeBase, EoL, PolyTraits and from individual researchers and published research. The effort required to undertake this activity should, however, not be under-estimated and will require significant staff time to complete. Each of the trait content providers will be asked to enter into an agreement with WoRMS in order to clearly communicate the nature of the relationship, how the traits data will be used and disseminated, and to agree a schedule for future updates.

3. Develop a timetable and details of responsibilities for use cases that will feedback on the development of the traits database and associated publications.
Appendix. Agenda for the Workshop.

Trait collection work done so far:

- Species’ importance to society - Numbers and statistics (15’ Dan Lear, MBA)
- The World Register of Introduced Marine Species (Leen Vandepitte, VLIZ 15’)
- Inclusion data from pilot projects, including bodysize data (Stefanie Dekeyzer, VLIZ 15’)
- Interfaces, vocabularies webservices and workflows
  - Interface attribute selection and trait based workflows, developed under the Lifewatch VRE (Simon Claus, VLIZ 15’)
  - Interface vocabulary published in semantic Wiki (Harvey Tyler Walters, MBA 15’)
- The application of traits thesauri in LifeWatch Italy (Nicola Fiori, University of Salento 15’’)

11.00-11.30: Coffee

11.30-12.15
- Applications of trait data
  - "Individual traits or life-history groups: does it matter? - An example from the North Sea benthos" (Olivier Bauchard, NIOZ 20’)
  - Where EurOBIS, WoRMS & EMODnet meet: progress and possible user cases for European nematode biomass data (Jan Vanaverbeke, UGhent 20’)

12.15 - 13.00
- Global trait related initiatives and discussion
  - "A global marine plankton trait database: user community, challenges, and possibilities (Andrew Barton, Princeton University 20’)
  - Marine trait data within EOL: structure and statistics (Jennifer. Hammock, EOL 15’)

Lunch: 13.00-14.00

Afternoon: 14.00-17.30
- Discuss objectives break out groups (plenary)
- Breakout groups (Suggested content and participants, open to changes)
  - Group 1: Review vocabulary and approach and timeline to implement priority traits data
    - Chair: Simon Claus - Suggested participants: Harvey Tyler Walters, Leen Vandepitte, David Johns, Geoff Boxshall, Jennifer Hammock, Serge Gofas, Mary Kennedy
  - Group 2: Tools, interfaces and workflows for analysis and visualization traits, taxonomic information and and distributions
    - Chair: Dan Lear. Suggested participants: Bart Vanhoorne, Dmitry Schigel, Lea Roselli, Nicola Fiore, Stephane Pesant, Tjess Hernandez
  - Group 3: Identifying scientific use cases using available trait, taxonomy and distributions data
    - Chair: Mark Costello - Suggested participants: Jan Vanaverbeke, Andrew Barton, Jan Mees, Olivier Bauchard, Geoffrey Read, Stefanie Dekeyzer

Evening: Joint dinner

Wednesday 10/06/2015

09.00-10.30
- Morning: Plenary session: trait data mining: tools and possibilities
  - Automatic extraction of habitat information from literature / web pages and visualization and processing of that info (by Evangelos Pafilis, HCMR)
  - Traits mining from legacy literature in the framework of the EUBON project (Donat Agosti, PLAZI)
- Discussion on best possible approach for EMODnet traits

10.30-10.45: Coffee

10.45-12.30
- Continuation work in break out groups and wrap up
Lunch
14.00-16.30
- Afternoon: Outcome breakout groups (presentation and discussion: 30 minutes/group)
- Planned actions related to traits within Lifewatch/EMODnet/WoRMS for 2015-2016 (30 minutes)
- Closing.
## Participants

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nicolas Bailly</td>
<td>HCMR</td>
</tr>
<tr>
<td>Olivier Beauchard</td>
<td>NIOZ</td>
</tr>
<tr>
<td>Geoff Boxshall</td>
<td>NHM</td>
</tr>
<tr>
<td>Simon Claus</td>
<td>VLIZ</td>
</tr>
<tr>
<td>Mark Costello</td>
<td>University of Auckland</td>
</tr>
<tr>
<td>Stefanie Dekeyzer</td>
<td>VLIZ</td>
</tr>
<tr>
<td>Sarah Faulwetter</td>
<td>HCMR</td>
</tr>
<tr>
<td>Nicola Fiore (teleconf)</td>
<td>University of Salento</td>
</tr>
<tr>
<td>Serge Gofas</td>
<td>University of Malaga</td>
</tr>
<tr>
<td>Jen Hammock</td>
<td>EoL</td>
</tr>
<tr>
<td>Francisco Hernandez</td>
<td>VLIZ</td>
</tr>
<tr>
<td>Mary Kennedy</td>
<td>Canadian Register of Marine Species</td>
</tr>
<tr>
<td>Dan Lear</td>
<td>MBA</td>
</tr>
<tr>
<td>Geoffrey Read</td>
<td>NIWA</td>
</tr>
<tr>
<td>Dmitry Schigel</td>
<td>GBIF</td>
</tr>
<tr>
<td>Harvey Tyler-Walters</td>
<td>MBA</td>
</tr>
<tr>
<td>Jan Vanaverbeke</td>
<td>UGhent</td>
</tr>
<tr>
<td>Leen Vandepitte</td>
<td>VLIZ</td>
</tr>
<tr>
<td>Bart Vanhoorne</td>
<td>VLIZ</td>
</tr>
</tbody>
</table>