

# Unlocking European marine biodiversity data under EMODnet Biology using the FAIR principles

Christos Arvanitidis, Nicolas Bailly, Klaas Deneudt, Daphnis De Pooter,  
Peter Herman, Dan Lear, Paula Oset, Leen Vandepitte

Simon Claus | Flanders Marine Institute  
Coordinator EMODnet  
Biology

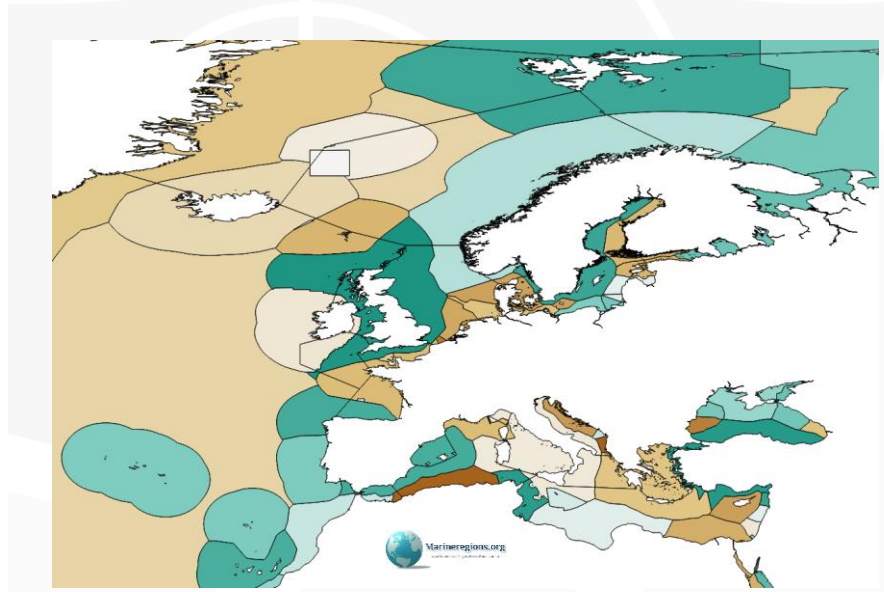
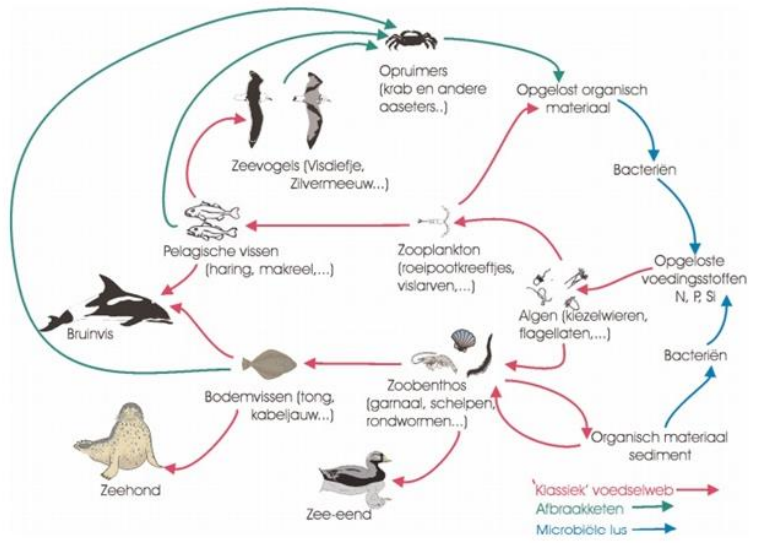


# EMODnet Biology Objectives'

## EMODnet



The objective is to **assemble existing data** from public and private organisations relating to the state of sea basins; processing them into **interoperable formats which includes agreed standards**, assessments of their accuracy and precision, making them **available** and creating **data products**;





**EMODnet**



# The FAIR principle



**Making Marine Biological Data Findable**

**Making Marine Biological Data Accessible**

**Making Marine Biological Data Interoperable**

**Making Marine Biological Data Reusable**



**EMODnet**



European Marine  
Observation and  
Data Network

# Making Marine Biological Data Findable



## Metadata system(s)

- **IMIS**

- **Contains 1315 datasets descriptions (ISO 19115)**

- **As webservice:**

- **show=html** default return offered in current website
- **show=json** JSON format
- **show=xml** XML format
- **show=rss** RSS/ATOM format
- **show=gcmd** (GCMD)

- **MetaGIS**

- **Geographic data layers**
- **INSPIRE compliant**
- **Published via geonetwork (csw)**

The screenshot shows the EMODnet BIOLOGY Data Catalog search interface. The header includes the EMODnet logo and the text "BIOLOGY Dive into data on Europe's marine life". Below the header is a navigation menu with links: Home, Data Catalog, Data Download, Map Viewer, Data Products, Project, and Contribute. The main content area is titled "Data Catalog" and shows "Datasets (1315)". There is a search bar with the placeholder text "Search in all fields: [input] Title or word in abstract". Below the search bar are several filter fields: Theme, Data type, Data origin, Title, Institute, and Person. There is also a "Word in abstract:" section with a search input. At the bottom, there are fields for "Begin date", "End date", "Archived", "Thesaurus term", "Taxonomic term", "Geographical term", and "Availability", each with a search button. A "search" button is located at the bottom left, and a "show full list" link is at the bottom right. The page footer includes the text "simple search".



**EMODnet**



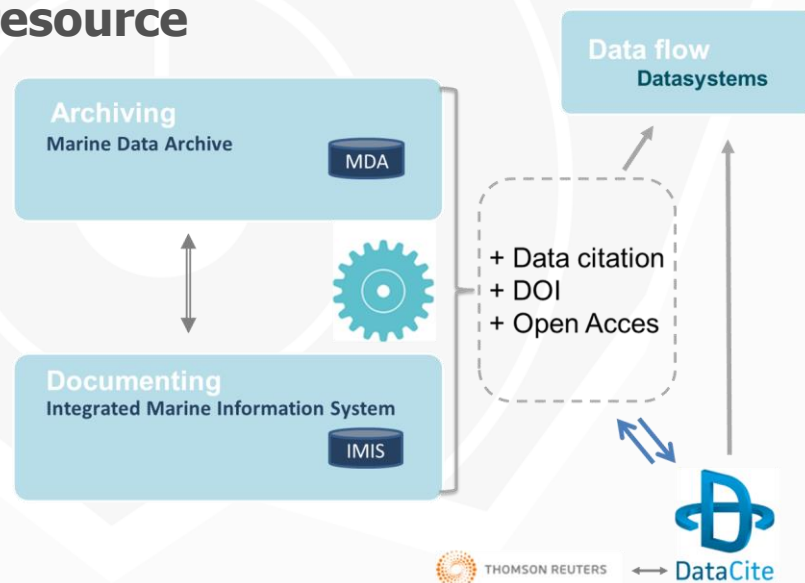
European Marine  
Observation and  
Data Network

# Making Marine Biological Data Accessible



## Create a Digital Object Identifier (DOI) of your dataset

- **Data files stay persistently available**
  - store 'raw' datafiles in Marine Data Archive
  - publish 'mapped' data as IPT resource
- **Data file cannot change anymore**
- **Minimum metadata available**
- **Data has gone through some qc**
- **Citation of dataset**
- **CC-license**



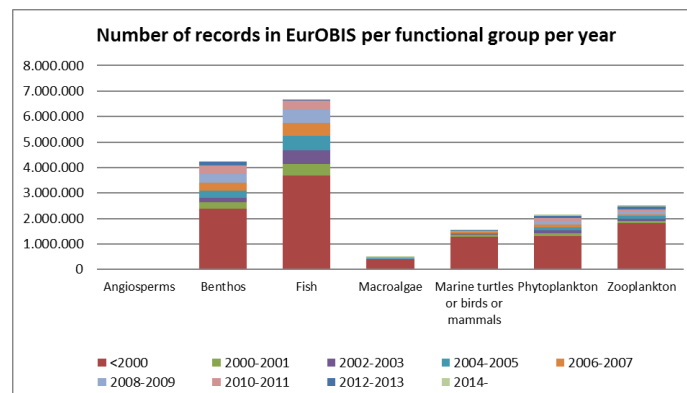


**EMODnet**



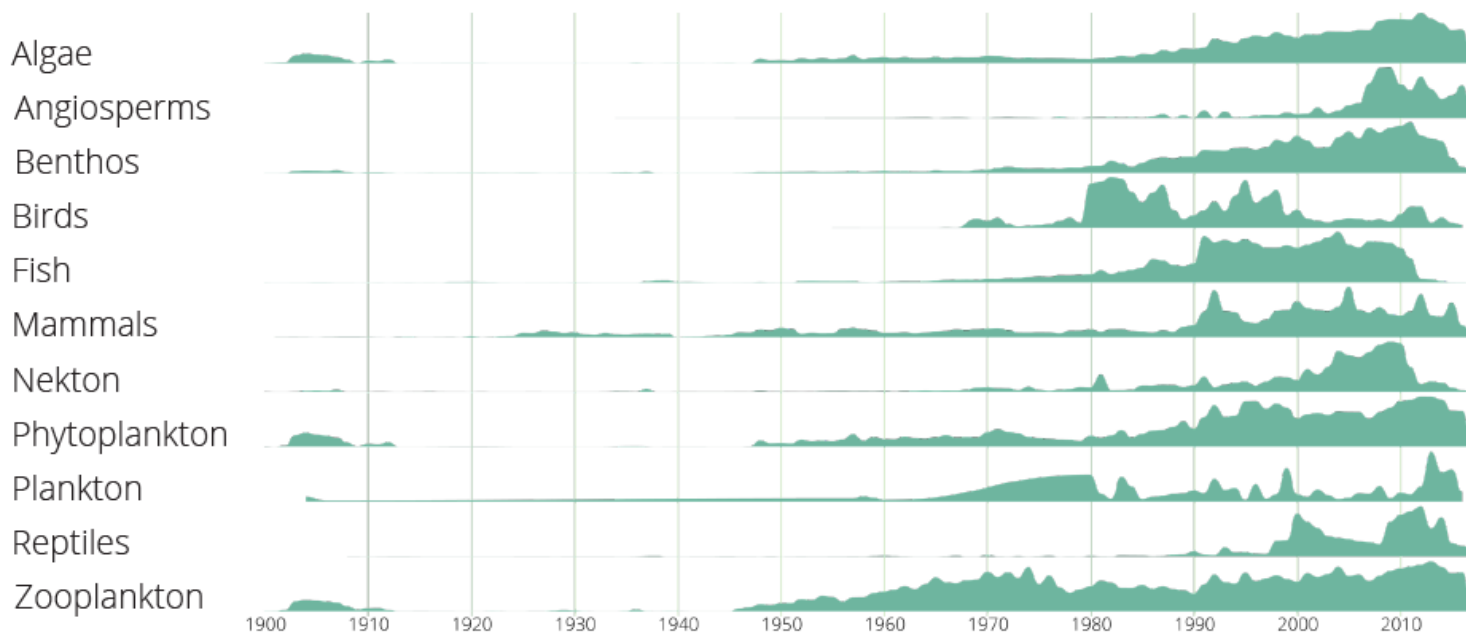
European Marine  
Observation and  
Data Network

## *Taxonomic & temporal scope*



### **Temporal coverage per functional group**

Time series of the relative number of records per functional group from 1900 to present. EMODnet offers historical records of species occurrences that date back to 1526.



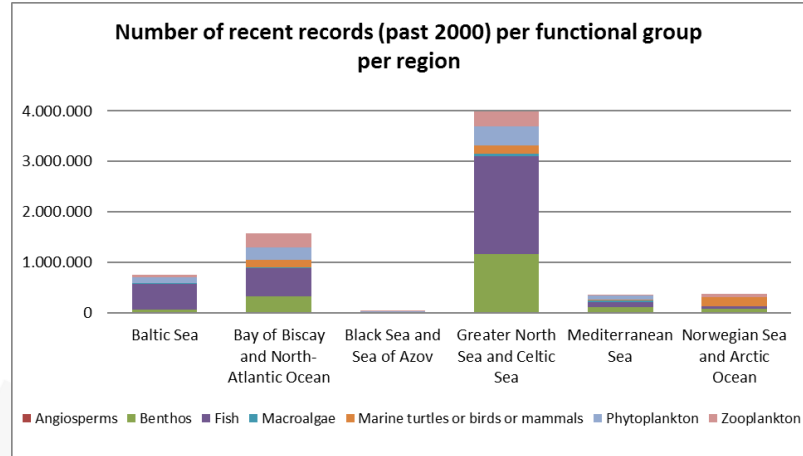




# EMODnet

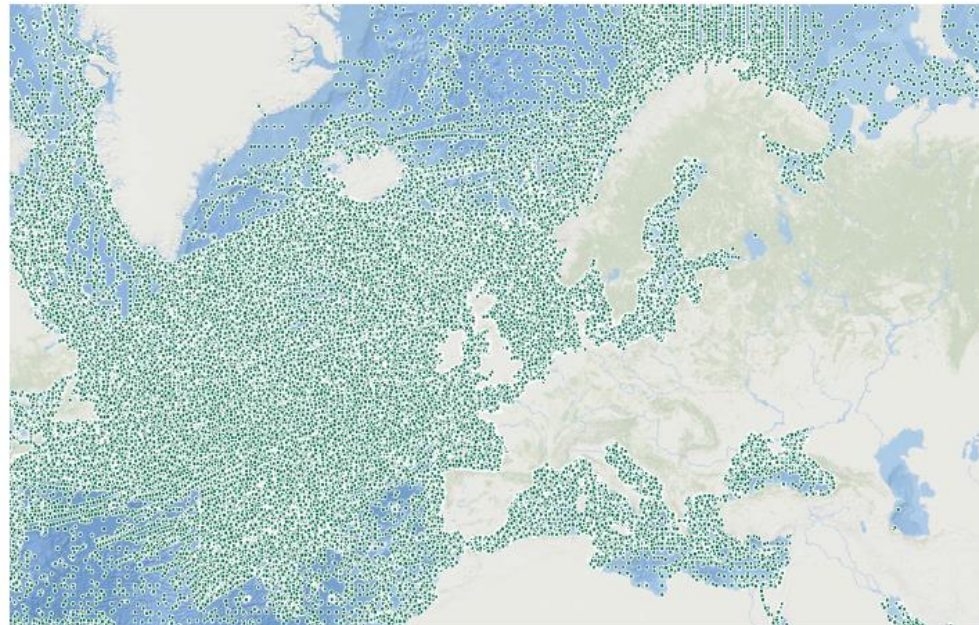
European Marine  
Observation and  
Data Network

## Spatial scope



### Spatial coverage

Map showing the location of the distribution records in EMODnet Biology currently available (25/09/2018): currently 875 datasets representing 23.860.954 occurrence records, from 77.723 species names.

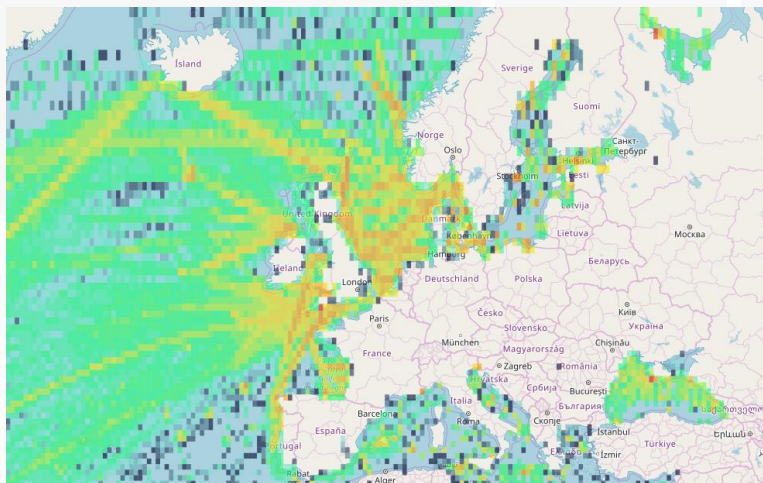
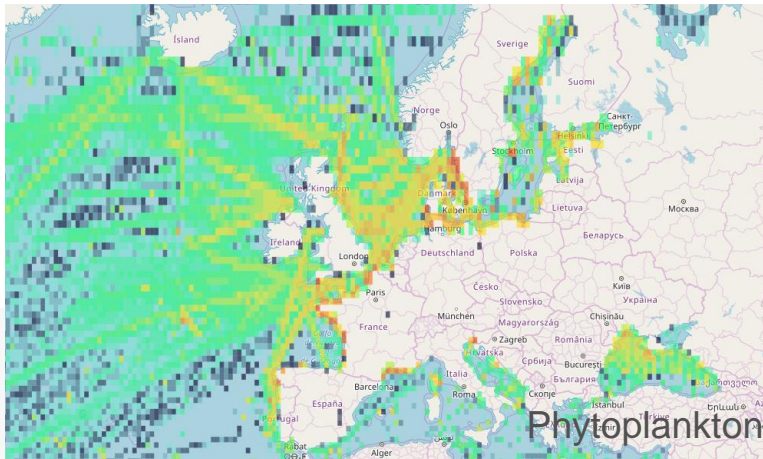




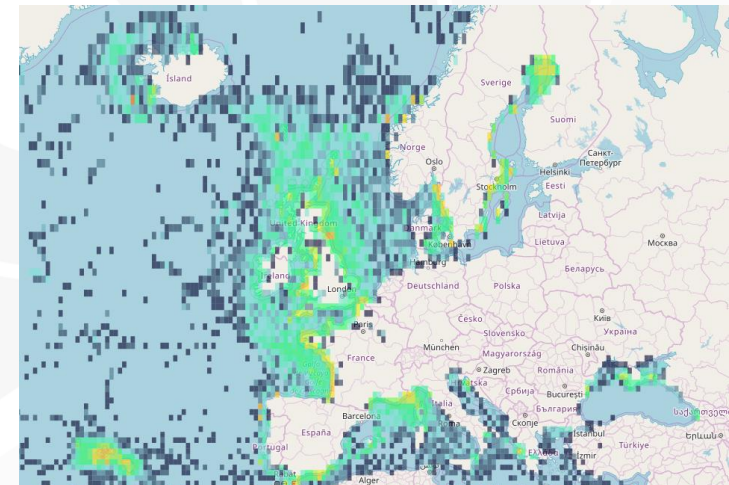
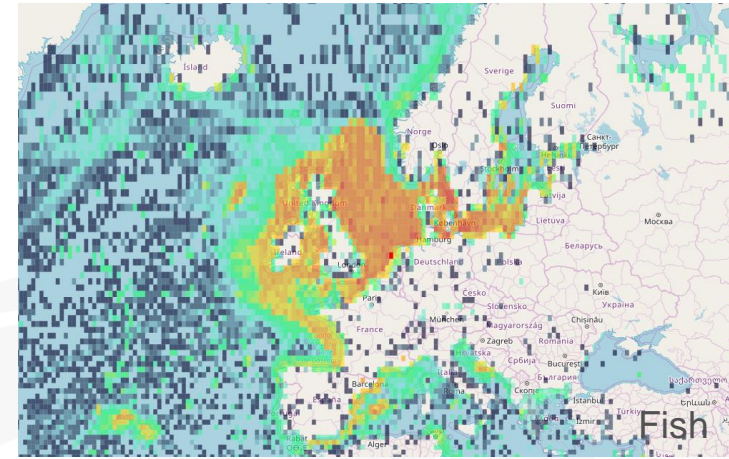
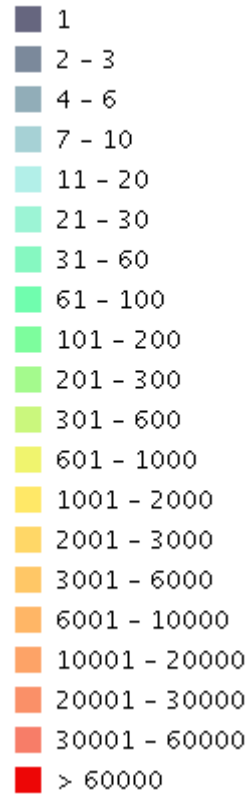
# EMODnet



European Marine  
Observation and  
Data Network



Legend:  
eurobis\_grid\_30m  
records/grid







**EMODnet**



---

## *Data types*

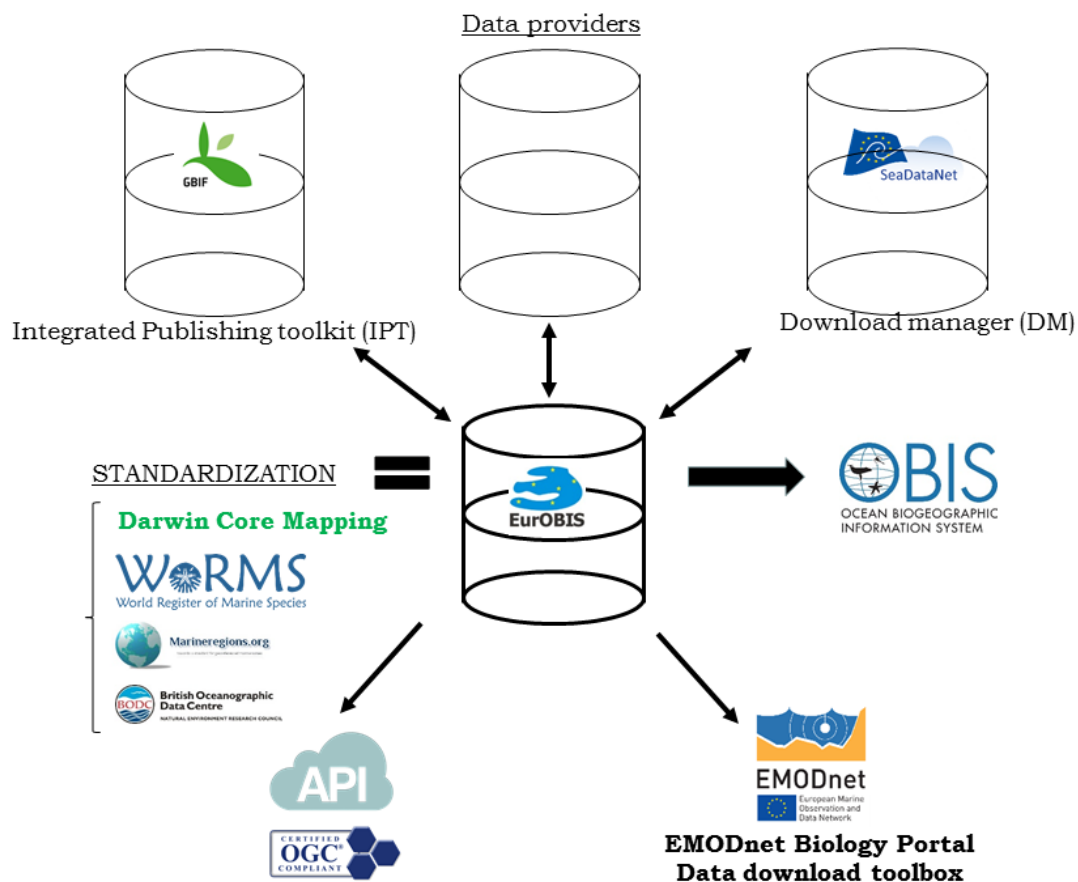
- ① Core: Species name, Accepted species name, Date, Latitude, Longitude
- ① Biota Quantifications : Abundance, biomass
- ① Biota Descriptors: Biota biometrics, Biota chemical properties, Functional traits
- ① Linked environmental parameters: Sediment grain size parameters, Rock and sediment chemistry, Carbon concentrations, Temperature of the water column, Secchi disk depth, Salinity of the water column, Dissolved oxygen parameters in the water column, Dissolved gases, Nitrogen concentrations in sediment, Bathymetry and Elevation
- ① Sampling Descriptors



EMODnet

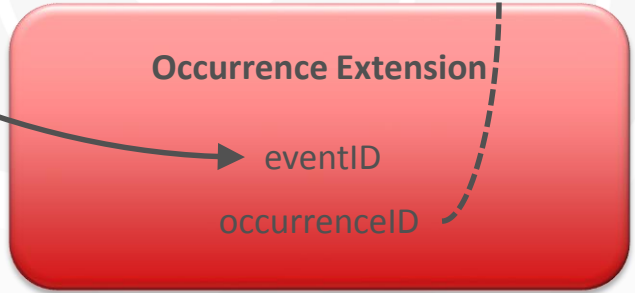
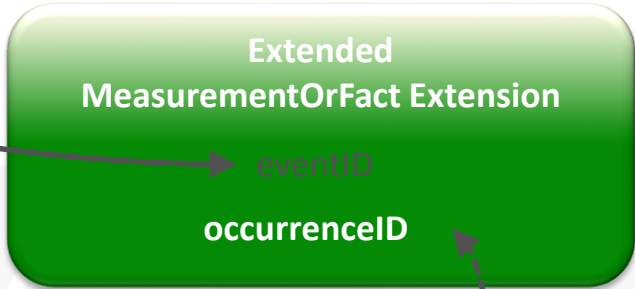


# Making Marine Biological Data Interoperable





# OBIS Env datascheme ~ Darwin Core



Biodiversity Data Journal 5: e10989  
doi: 10.1093/bdj/5/e10989

**Research Article**

**Toward a new data standard for combined marine biological and environmental datasets - expanding OBIS beyond species occurrences**

Daphnis De Potter<sup>1</sup>, Ward Appeltans<sup>1</sup>, Nicolas Bailly<sup>1</sup>, Sky Brisco<sup>1</sup>, Kiasa Deneuot<sup>1</sup>, Menashé Eliezer<sup>2</sup>, El Fujikura<sup>3</sup>, Alessandro Giorgini<sup>4</sup>, Philip Goodstein<sup>5</sup>, Mirna Lewis<sup>6</sup>, Maria Lipton<sup>7</sup>, Kevin Mackay<sup>8</sup>, Maria Marin<sup>9</sup>, Owenella Moncalvo<sup>10</sup>, Stamatina Nikolopoulou<sup>11</sup>, Pieter Provost<sup>12</sup>, Shantenn Raach<sup>13</sup>, Andres Roubicek<sup>14</sup>, Carlos Torres<sup>15</sup>, Anton van de Putte<sup>16</sup>, Leen Vandepitte<sup>17</sup>, Bart Vanhoose<sup>18</sup>, Matteo Vinci<sup>19</sup>, Nina Wambui<sup>20</sup>, David Watts<sup>21</sup>, Eduardo Klein Salas<sup>22</sup>, Francisco Hernandez<sup>23</sup>

<sup>1</sup> Vlaams Instituut voor de Zee (Flanders Marine Institute, EvOBS), Oostende, Belgium  
<sup>2</sup> UNESCO Intergovernmental Oceanographic Commission, Ocean Biogeographic Information System, UNESCO-IOC Project Office for ODS, Oostende, Belgium  
<sup>3</sup> Institute Center for Marine Research, MacDiarmid, Gosport, Greece  
<sup>4</sup> United States Geological Survey, OBIS-USA, Virginia, United States of America  
<sup>5</sup> United States Geological Survey, OBIS-USA, Virginia, United States of America  
<sup>6</sup> Italian National Oceanographic Institute, Oceanographic Department (ODI), Trieste, Italy  
<sup>7</sup> Marine Geospatial Ecology Laboratory, Nicholas School of the Environment, Duke University, OBIS-SEAMAP, Durham, NC, United States of America  
<sup>8</sup> University of Colorado Museum of Natural History, OBIS-USA, Colorado, United States of America  
<sup>9</sup> Centro Nacional Patagónico, OBIS, Puerto Madryn, Argentina  
<sup>10</sup> National Institute of Water and Atmospheric Research, Southwest Pacific OBIS, Wellington, New Zealand  
<sup>11</sup> British Oceanographic Data Centre, Liverpool, United Kingdom  
<sup>12</sup> Woods Hole Oceanographic Institution, Woods Hole, United States of America  
<sup>13</sup> CSIRO Oceans and Atmosphere, OBIS-Australia, Hobart, Australia  
<sup>14</sup> Universidad Autónoma de Baja California, Mexcala, Mexico  
<sup>15</sup> CD Nature, Royal Belgian Institute for Natural Sciences, Antarctic OBIS, Brussels, Belgium  
<sup>16</sup> King's College London, OBIS, London, United Kingdom  
<sup>17</sup> Universidad Simon Bolivar, Caribbean OBIS, Caracas, Venezuela

Corresponding author:  
Academic editor: Christos Avranitis

Received: 31 Oct 2016 | Accepted: 07 Jan 2017 | Published: 09 Jan 2017

Citation: De Potter D, Appeltans W, Bailly N, Brisco S, Deneuot K, Eliezer M, Fujikura E, Giorgini A, Goodstein P, Lewis M, Lipton K, Mackay K, Marin M, Moncalvo O, Nikolopoulou S, Provost P, Raach S, Roubicek A, Torres C, van de Putte A, Vandepitte L, Vanhoose B, Vinci M, Wambui N, Watts D, Klein Salas E, Hernandez F (2017) Toward a new data standard for combined marine biological and environmental datasets - expanding OBIS beyond species occurrences. Biodiversity Data Journal 5: e10989. <https://doi.org/10.1093/bdj/5/e10989>



Biodiversity  
Information  
Standards  
T D W G



# WORMS

## World Register of Marine Species

Global effort to register all marine species names  
(incl invalid synonyms)

270 taxonomic editors, 185 institutions, 38 countries  
Standard for OBIS, IODE, ICES, BODC, SeaDataNet  
> 40 global or regional lists

<http://www.marinespecies.org>

Developed and hosted by VLIZ





# WoRMS

## World Register of Marine Species



*Halichondria panicea* (Pallas 1776)

Alcyonium manusdiaboli sensu Esper, 1794  
Alcyonium medullare Lamarck, 1815  
Alcyonium paniceum (Pallas, 1766)  
Amorphina appendiculata Schmidt, 1875  
Amorphina grisea Fristedt, 1887  
Amorphina paciscens Schmidt, 1875  
Clathria (Microciona) seriata (Grant, 1826)  
Clathria (Microciona) tumulosa (Bowerbank, 1882)  
Clathria seriata (Grant, 1826)  
Halichondria albescens Rafinesque, 1818  
Halichondria ambigua Bowerbank, 1874  
Halichondria brettii (Bowerbank, 1866)  
Halichondria caduca Bowerbank, 1866  
Halichondria coccinea Bowerbank, 1861  
Halichondria coralloides Bowerbank, 1882  
Halichondria edusa Bowerbank, 1874  
Halichondria firmus (Bowerbank, 1874)  
Halichondria glabra Bowerbank, 1866  
Halichondria grisea (Fristedt, 1887)

Halichondria incerta Bowerbank, 1866  
Halichondria lactea (Bowerbank, 1866)  
Halichondria membrana (Bowerbank, 1866)  
Halichondria paciscens (Schmidt, 1875)  
Halichondria panicea (Pallas, 1766)  
Halichondria pannosus Verrill, 1874  
Halichondria papillaris (Linnaeus, 1791)  
Halichondria reticulata Lieberkühn, 1859  
Halichondria sevesa Johnston, 1842  
Halichondria topsenti de Laubenfels, 1936  
Halichondriella corticata Burton, 1931  
Halina panicea (Pallas, 1766)  
Hymeniacion brettii Bowerbank, 1866  
Hymeniacion coccinea (Bowerbank, 1861)  
Hymeniacion fallaciosus Bowerbank, 1866  
Hymeniacion firmus Bowerbank, 1874  
Hymeniacion fragilis Bowerbank, 1866  
Hymeniacion lactea Bowerbank, 1866  
Hymeniacion membrana Bowerbank, 1866  
Hymeniacion parfitti Parfitt, 1868  
Hymeniacion reticulatus Bowerbank, 1866  
Hymeniacion solida Bowerbank, 1874  
Hymeniacion tegeticula Bowerbank, 1874  
Hymeniacion thomasii Bowerbank, 1866  
Isodictya crassa Bowerbank, 1882  
Isodictya perplexa Bowerbank, 1882  
Menanetia minchini Topsent, 1896  
Microciona tumulosa Bowerbank, 1882  
Pellina bibula Schmidt, 1870  
Seriatala seriata (Grant, 1826)  
Spongia compacta Sowerby, 1806  
Spongia cristata Ellis & Solander, 1786  
Spongia panicea Pallas, 1766  
Spongia seriata Grant, 1826  
Spongia tomentosa Linnaeus, 1767  
Spongia tubulosa Ellis & Solander, 1786  
Spongia urens Ellis & Solander, 1786  
Spuma borealis var. convoluta Miklucho-Maclay, 1870  
Spuma borealis var. tuberosa Miklucho-Maclay, 1870  
Spuma borealis var. velamentosa Miklucho-Maclay, 1870  
Trachypsilla glaberrima Burton, 1931



**EMODnet**



# Making Marine Biological Data Reusable



- 🎯 Conservation & Management Organisations & NGO's
- 🎯 Transatlantic/Global Initiatives
- 🎯 Regional Sea Commissions





**EMODnet**



---

# EMODnet biology data products

- ① It is using species observation data to create biological data products that showing changes in species, communities and functional traits over time
- ① A combination of tools, models and spatial maps that allow to visualise biological data to different users
- ① Can be reused and anyone can further build upon the products for any purpose, even commercially
- ① Products freely available (as webservice); workflow as markdown document (in R, in Jupyter)



**EMODnet**



---

# EMODnet biology data products

- ① Structured around Essential Ocean Variables for Biodiversity (EOV)
  - ① Phytoplankton biomass and diversity
  - ① Zooplankton biomass and diversity
  - ① Fish abundance and distribution
  - ① Benthic invertebrate abundance and distribution
  - ① Marine turtles, birds, mammals abundance and distribution



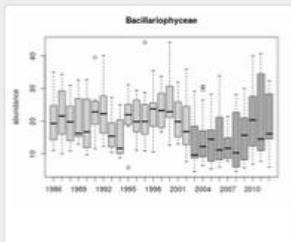


EMODnet



European Marine  
Observation and  
Data Network

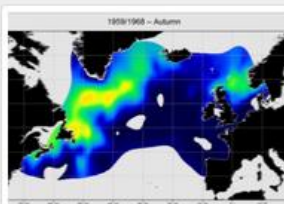
# Data product portfolio



Phytoplankton community analysis in the Northern Adriatic

Source: OGS & Deltaris

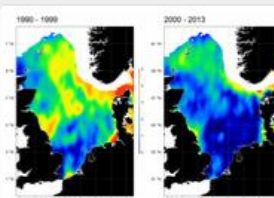
Visualize



Gridded abundance maps of most common Atlantic Copepod species

Source: VLIZ, Ulg And SAHFOS

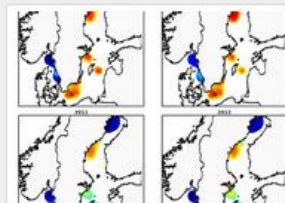
Visualize



Trends in abundance of fish species in the North Sea

Source: IBTS

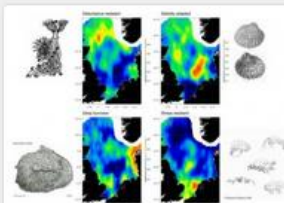
Visualize



Data workflow analyzing trends of Swedish zooplankton species

Source: EMODnet Biology

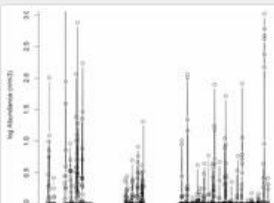
Visualize



Benthic trait analysis showing benthic communities of the Atlantic

Source: Beauchard, Olivier

Visualize



Long term zooplankton time series analysis from the West Med Sea

Source: CNRS, Deltaris

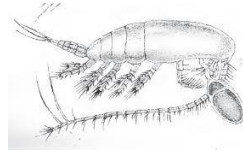
Visualize



EMODnet

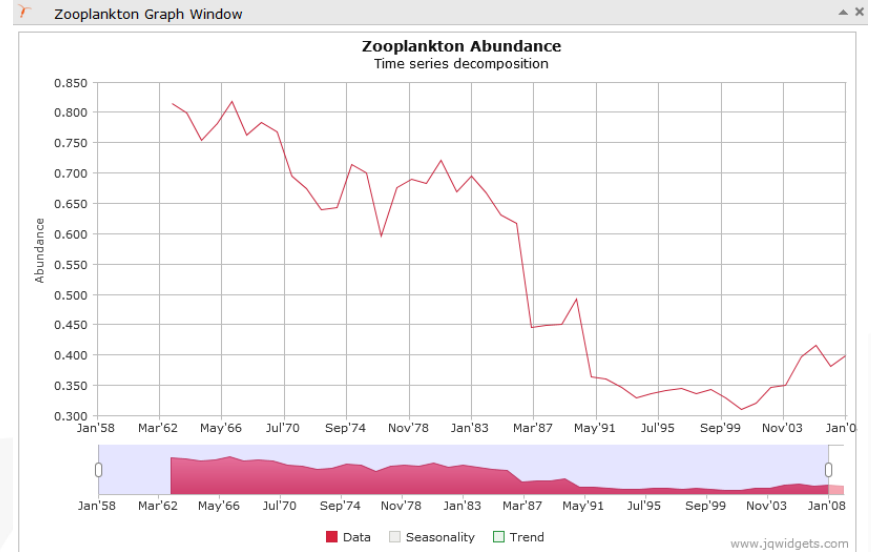
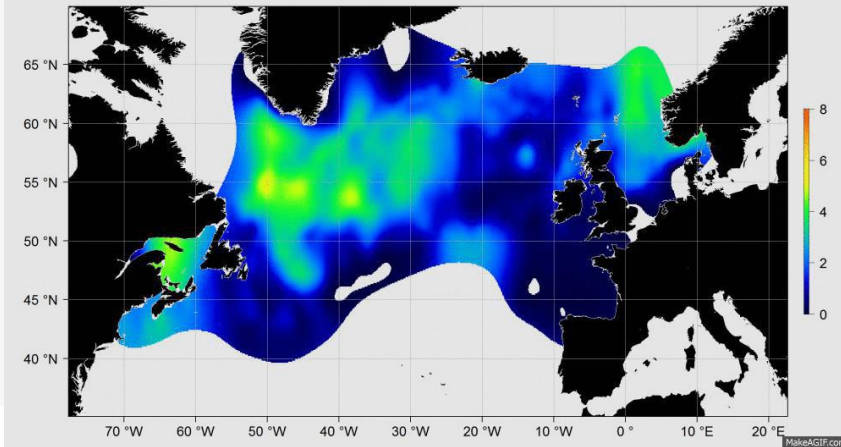


# Gridded abundance maps



*Calanus finmarchus*

1958/1967 – Summer



Service for ICES Ecosystem Overviews





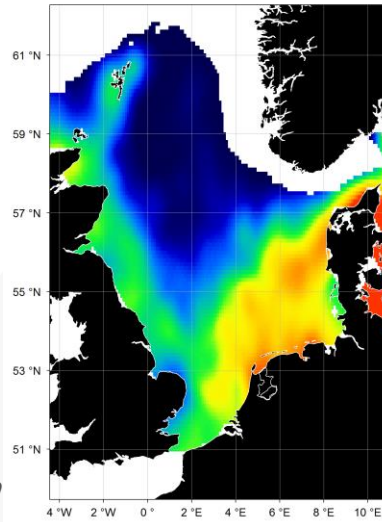
EMODnet



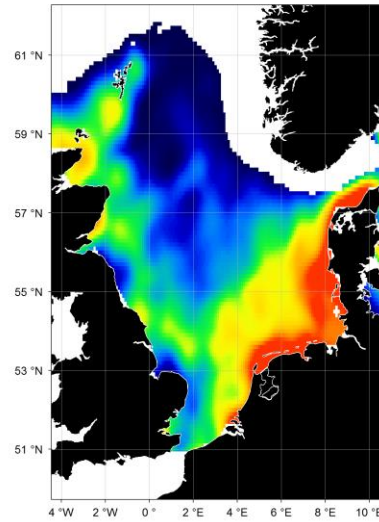
*Plaice –  
Pleuronectes platessa*

# Gridded abundance maps

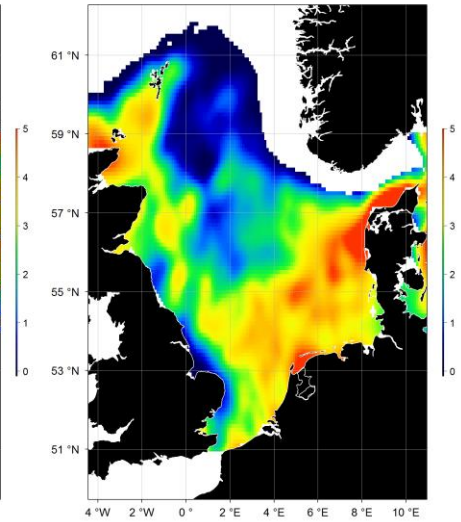
1980 - 1989



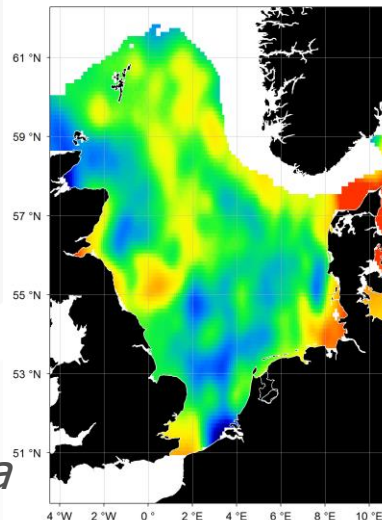
1990 - 1999



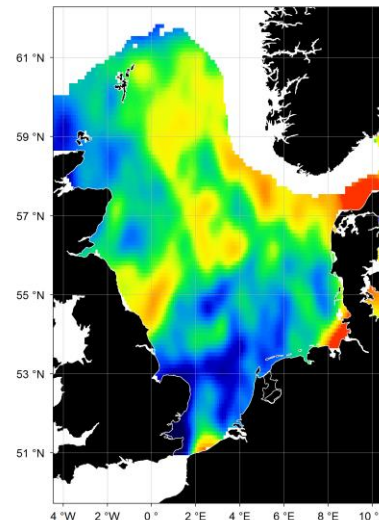
2000 - 2013



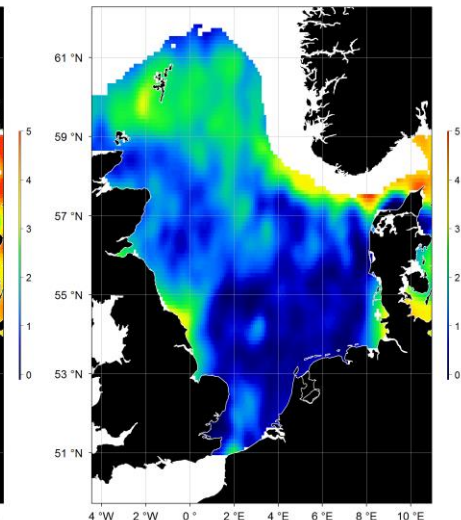
1980 - 1989



1990 - 1999



2000 - 2013



*Cod – Gadus morhua*



EMODnet



# Plankton time series analysis

## ILTER North Adriatic plankton series

Currently restricted to  
Phytoplankton

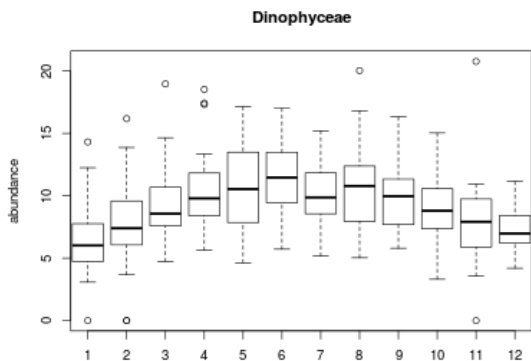
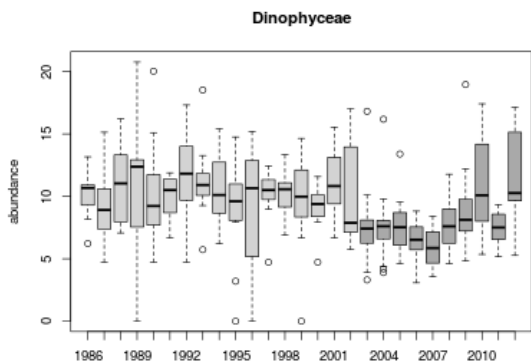
Phytoplankton

Choose a species (group)

Dinophyceae

value double sqrt transformed

Observations Multiv 1 about



## ILTER North Adriatic plankton series

Currently restricted to  
Phytoplankton

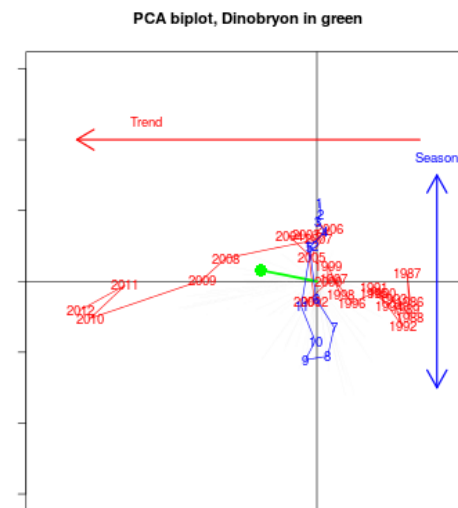
Phytoplankton

Choose a species (group)

Dinobryon

- Dictyocha
- Dimeregramma
- Dinobryon
- Dinophyceae indet.
- Dinophysis
- Diploneis
- Diplopsalis
- Emiliana

Observations Multiv 1 about





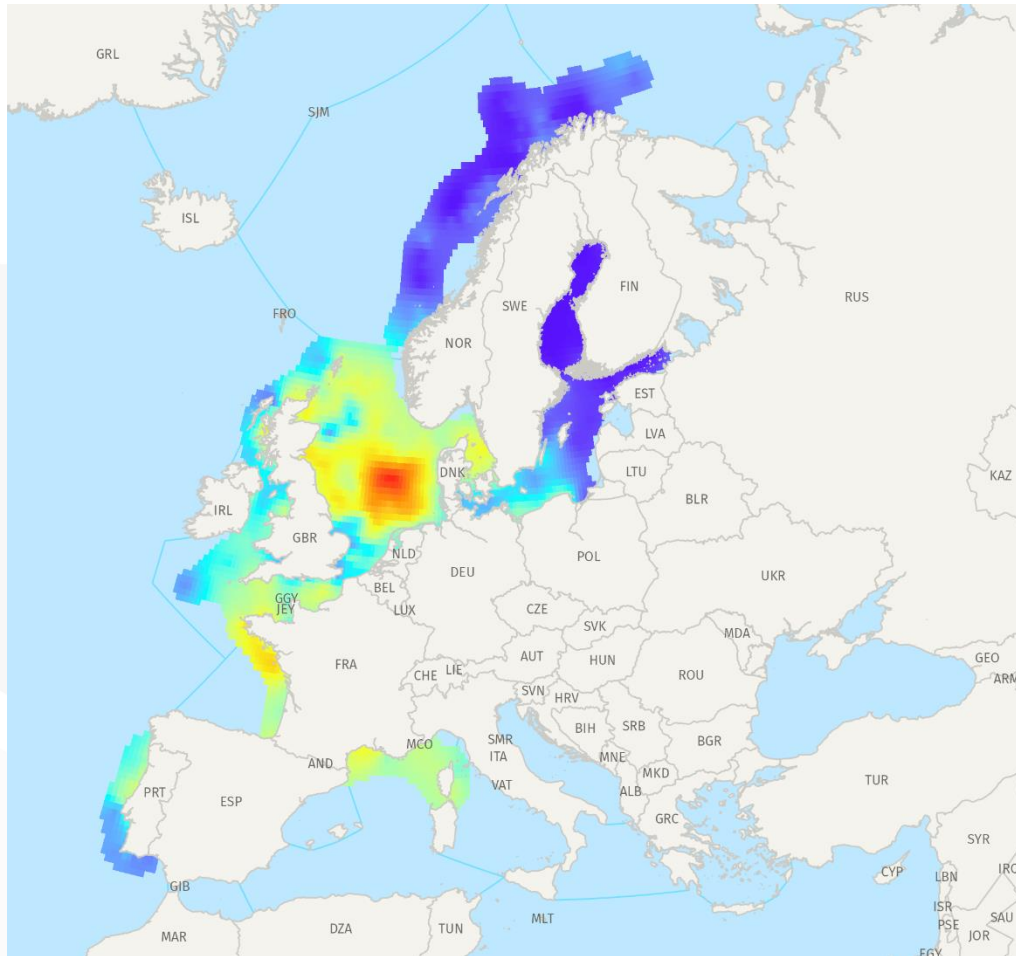


EMODnet

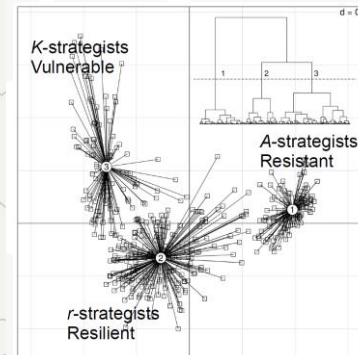


European Marine  
Observation and  
Data Network

# Macrobenthos '*vulnerability to physical damage*'



Functional types of sea floor macroinvertebrates derived from 13 life history traits defined on 617 abundant benthic taxa





# EMODnet



European Marine  
Observation and  
Data Network

# Products, application, stories & markdown

```

Reading in the basic data files (first the main file, then the measurements and facts)
bd<-read.csv("20180425_155218_15ae08812c7c53.csv",header = T,
             stringsAsFactors = F)
mf<-read.csv("20180425_155240_15ae088289e138.csv",header = T,
             stringsAsFactors = F)

Inspection of the data shows that there is a problem with dates, but also with abundance (fields are empty)
for all data that have "1994Cnid" in measurementremarks. We temporarily drop this part of the files to avoid
those problems. First we have to merge the two files. VLIZ: merging observation files with 'measurement and
facts' files is impossible without first correcting the name 'occurrenceid' (with single r) into 'occurrenceid'
(double r) in 'measurement and facts'. I now know this error, but it has taken me a lot of searching before I
first saw the difference. It is a nuisance you have to remove

names(mf)[2]<-"occurrenceid"
bd<-merge(bd,mf,"occurrenceid")
bd<-bd[bd$measurementremarks!="1994Cnid",]

There are a few records that do not have startdaycollected, but do have enddaycollected, and a few where it
is the other way round. We will give the available value to the other, nonavailable one.

bd[is.na(bd$startdaycollected),"startdaycollected"]<-
  bd[is.na(bd$startdaycollected),"enddaycollected"]
bd[is.na(bd$enddaycollected),"enddaycollected"]<-
  bd[is.na(bd$enddaycollected),"startdaycollected"]

We want a starting year for all observations. When it is unavailable, we take yearcollected as the starting year
bd$startyearcollected[is.na(bd$startyearcollected)]<-
  bd$yearcollected[is.na(bd$startyearcollected)]
bd$endyearcollected[is.na(bd$endyearcollected)]<-
  bd$yearcollected[is.na(bd$endyearcollected)]

After checking, we now have startyearcollected and endyearcollected for all records.

We do the same trick with months
bd$startmonthcollected[is.na(bd$startmonthcollected)]<-
  bd$monthcollected[is.na(bd$startmonthcollected)]

+
> mof <- read.csv("http://geo.vliz.be/geoserver/Dataportal/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=
Dataportal:euobis_measurementandfacts&viewParams=where:dataportaldid=1086&outputformat=csv", header = T,
+               stringsAsFactors = F)
> names(mof)[3]<-"occurrenceid"
> all<-merge(bd,mof,"occurrenceid")
> ggplot() + geom_point(data = all, aes(x = measurementvalue, y = scientificname))
Error in ggplot() : could not find function "ggplot"
> library(ggplot2)
> ggplot() + geom_point(data = all, aes(x = measurementvalue, y = scientificname))
+

```

EMODnet  
Your gateway to marine data in Europe  
http://www.emodnet.eu/ info@emodnet.eu

Repositories 9 People 0 Projects 0

**Grow your team on GitHub**  
GitHub is home to over 28 million developers working together. Join them to grow your own development teams, manage permissions, and collaborate on projects.

Sign up

scientificname	measurementvalue
Triplos lineatus	~10000
Scrippsiella	~10000
Prorocentrum triestinum	~10000
Prorocentrum micans	~10000
Lingulodinium polyedra	~10000
Dityocha speculum	~10000
Dityocha	~10000
Asterionellopsis glacialis	~10000
Alexandrium pseudogonyaulax	~10000



**EMODnet**



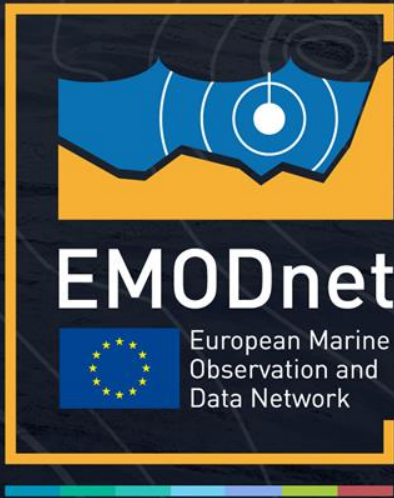
---

To be expected soon

[www.emodnet-biology.eu](http://www.emodnet-biology.eu)







[www.emodnet.eu](http://www.emodnet.eu)

*Your gateway to marine data in Europe*