



SeaDataNet

Biological data handling

Klaas Deneudt, Daphnis Depooter
VLIZ- Flanders Marine Institute

SeaDataNet2 Final meeting 16 September 2015



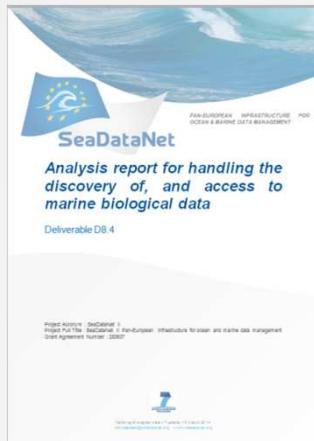
PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

OVERVIEW

- SDN2 deliverables biological data
- Format documentation
- Data published
- Data exchange services
- Conclusions and outlook

SDN2 Deliverables biological data

- D8.4a Analysis report with required adaptions for marine biological data
- D8.4b Format documentation
- Example files



Code	Station	Type	Latitude [degrees.west]	Longitude [degrees.north]	LOCAL_Cat_ID	EMO_code	Rel_Depth(m)	MinimumObservationDepth	QV/SEADATANET	MaximumObservationDepth	QV/SEADATANET
1											
2	/#/sdn:GIVEN/LOCAL_MinimumObservationDepth/subject/objetc#sdn:PO1_MINWODIST/objetc#unit#sdn:POE_ULAN/units#										
3	/#/sdn:GIVEN/LOCAL_MaximumObservationDepth/subject/objetc#sdn:PO1_MAXWODIST/objetc#unit#sdn:POE_ULAN/units#										
4	/#/sdn:GIVEN/LOCAL_SamplingMethod/subject/objetc#sdn:PO1_SAMPLINGMETHOD/objetc#unit#sdn:POE_ULAN/units#										
5	/#/sdn:GIVEN/LOCAL_SamplingEfficiency/subject/objetc#sdn:PO1_AREAENDS/objetc#unit#sdn:POE_ULAN/units#										
6	/#/sdn:GIVEN/LOCAL_SamplingPeriod/subject/objetc#sdn:PO1_SSAMINDIV/objetc#unit#sdn:POE_ULAN/units#										
7	/#/sdn:GIVEN/LOCAL_SamplingEfficiency/subject/objetc#sdn:PO1_SSAMINDIV/objetc#unit#sdn:POE_ULAN/units#										
8	/#/sdn:GIVEN/LOCAL_SamplingPeriod/subject/objetc#sdn:PO1_SSAMINDIV/objetc#unit#sdn:POE_ULAN/units#										
9	/#/sdn:GIVEN/LOCAL_SamplingEfficiency/subject/objetc#sdn:PO1_SSAMINDIV/objetc#unit#sdn:POE_ULAN/units#										
10	/#/sdn:GIVEN/LOCAL_SamplingPeriod/subject/objetc#sdn:PO1_SSAMINDIV/objetc#unit#sdn:POE_ULAN/units#										
11	/#/sdn:GIVEN/LOCAL_ObservationMethod/subject/objetc#sdn:PO1_ENZYME/objetc#unit#sdn:POE_ULAN/units#										
12	/#/sdn:GIVEN/LOCAL_ObservationMethod/subject/objetc#sdn:PO1_OCCURS/objetc#unit#sdn:POE_ULAN/units#										
13	/#/sdn:GIVEN/LOCAL_Density/subject/objetc#sdn:PO1_UPTAKE/objetc#unit#sdn:POE_ULAN/units#										
14	/#/sdn:GIVEN/LOCAL_Density/subject/objetc#sdn:PO1_UPTAKE/objetc#unit#sdn:POE_ULAN/units#										
15	/#/sdn:GIVEN/LOCAL_Density/subject/objetc#sdn:PO1_UPTAKE/objetc#unit#sdn:POE_ULAN/units#										
16											
17											
18	RV Mechelen197606	330	58	1976-06-05T06:02:13	2.92361	51.24411	MACROBELL1351_449	25	21.5	21.5	1
19	RV Mechelen197606	330	58	1976-06-05T06:02:13	2.92361	51.24411	MACROBELL1351_449	25	21.5	21.5	1
20	RV Mechelen197606	330	58	1976-06-05T06:02:13	2.92361	51.24411	MACROBELL1351_449	25	21.5	21.5	1
21	RV Mechelen197606	330	58	1976-06-05T06:02:13	2.92361	51.24411	MACROBELL1351_449	25	21.5	21.5	1
22	RV Mechelen197606	330	58	1976-06-05T06:02:13	2.92361	51.24411	MACROBELL1351_449	25	21.5	21.5	1
23	RV Mechelen197607	2002	58	1976-07-05T07:04:10	3.21113	51.35109	MACROBELL1206_449	28	23.1	23.1	1
24	RV Mechelen197607	2002	58	1976-07-05T07:04:10	3.21113	51.35109	MACROBELL1206_449	28	23.1	23.1	1
25	RV Mechelen197607	2002	58	1976-07-05T07:04:10	3.21113	51.35109	MACROBELL1206_449	28	23.1	23.1	1
26											
27											
28											

Format documentation

- Format description
- Template table listing mandatory, conditional and optional parameters
 - > 55 P01 parameters
- FAQ section
 - 25 questions and answers





SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

Label/column header in datafile	Conceptid	Use
Cruise		mandatory (ODV Default)
Station		mandatory (ODV Default)
Type		mandatory (ODV Default)
yyy-mm-ddThh:mm:ss.sss		mandatory (ODV Default)
Longitude [degrees_east]		mandatory (ODV Default)
Latitude [degrees_north]		mandatory (ODV Default)
LOCAL_CDI_ID		mandatory (ODV Default)
EDMO_code		mandatory (ODV Default)
Bot. Depth [m]		mandatory (ODV Default)
MinimumObservationDepth [m]		
MaximumObservationDepth [m]		
SampleID [#]		
SamplingEffort [unit]		
ScientificName [#]		
ScientificNameID [#]		
Sex [#]		
LifeStage [#]		
ObservedIndividualCount [#]		
EventID [#]		
EventStartTime [#]		
EventEndTime [#]		
EventStartLongitude [#]		
EventEndLongitude [#]		
EventStartLatitude [#]		
EventEndLatitude [#]		
SubsampleID [#]		
SubSamplingCoefficient [#]		
Samplingprotocol [#]		



FOR OCEAN & MARINE DATA
MANAGEMENT

5.3. General questions about ODV bio variant

Q) What should I do when my sample did not contain any organisms?

A) We recommend that you add a record containing the - for the dataset - highest relevant taxonomic level in the field ScientificName (<http://www.marinespecies.org/aphia.php?p=browser&id=2&expand=false#ct>), and add 0 as the ObservedIndividualCount and QF 1 for good value. E.g. Pisces, Animalia,...

Q) What should I do if the taxon was recorded, but I don't have an ObservedIndividualCount?

A) If you don't have an original count but you know the taxon was present, you can assign a NULL value (=leave blank) in combination with a QF of 1 or 2 for (probably) good value. If the original count was lost, but you have a calculated abundance, you can still assign the NULL value to ObservedIndividualCount if the taxon was present, in combination with a parameter "DensityPerUnitEffort". Do not assign the value 0 unless you want to provide absence information. A value of 0 (Zero) in the ObservedIndividualCount would mean that the sample was examined specifically for this taxon, but it was not encountered.

Q) What should I do if my dataset does not contain any information for a required field of the template?

A) The general rule for ODV is that you can leave empty mandatory fields setting the QC flag



SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

Data published: Overview

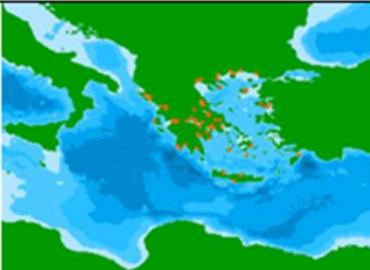
Spatial Coverage	CDI partner	Dataset	Size	DM
	Aarhus University	Time series of zoobenthos in Danish waters (1979 onwards) Danish benthic marine monitoring data from ODAM (1911 onwards)	18.048 CDI files 794.717 data records	Yes
	IFREMER	<u>Rephy - Réseau de Surveillance phytoplanctonique</u> 1984 –	495 CDI files 52.598 data records	Yes
	SOVON Dutch Centre for Field Ornithology	<u>Lauwersmeer</u> 1970 – 2012; Bird counts	323 CDI records 4.446 data records	No
	IMARES	1970 – 2013: Demersal (young) fish survey 1960 – 2013: Monitoring Seals in the Dutch Wadden Sea 2007 -2012: Inventory of shellfish species in the Dutch Wadden sea	333 CDI files 78.400 data records	Yes



SeaDataNet

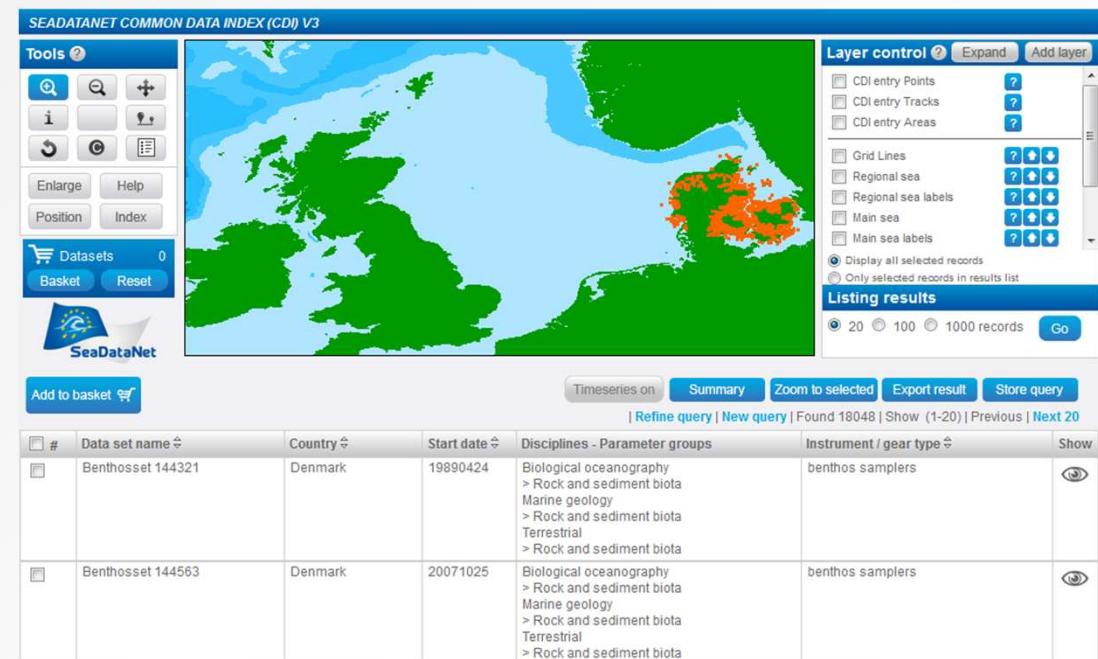
PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

New data published: Overview

Spatial Coverage	CDI partner	Dataset	Size	Access
	HCMR	WFD_HCMR_Zoobenthos 2012 - 2013	89 CDI's	By Negotiation
	NIOZ	Fyke data Fish caught in 24-48 hour cycles 1960 - 2014	107 CDI's 4600 records	SeaDataNet licence
	ICES	DOME phytoplankton 1983 - 2014	1400 CDI's 30.000 data records	SeaDataNet licence

Data published: Aarhus University

- Time series of zoobenthos in Danish waters (1979 onwards)
- Sediment biota
- 18.048 CDI records



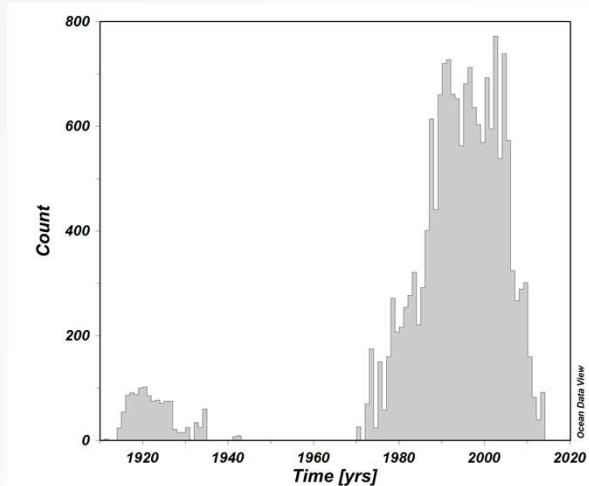
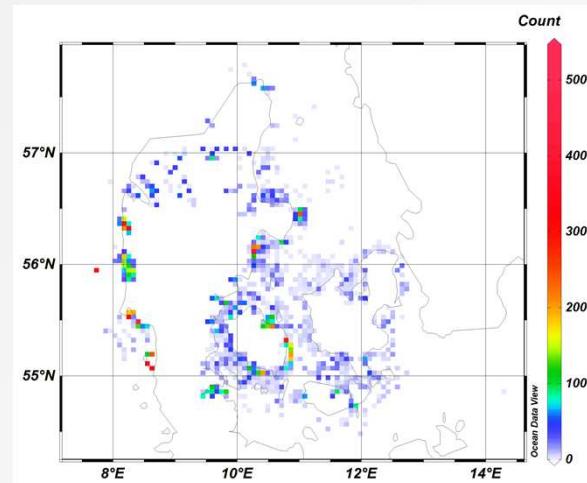
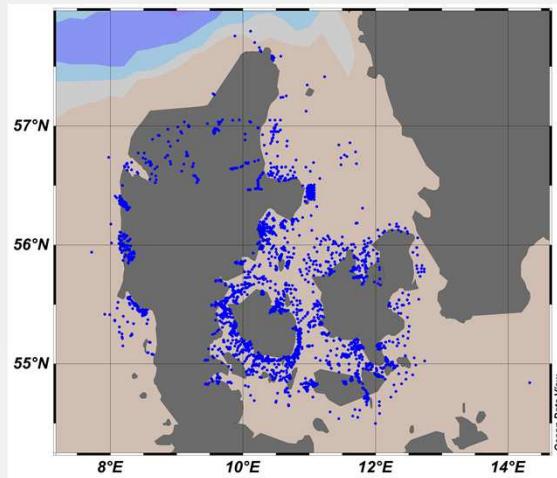


Data published: Aarhus University

- Use of download manager to generate ODV bio upon request

```
<?xml version="1.0" encoding="UTF-8"?>
<- <root>
    <updated>2014-08-29T07:00:00+02:00</updated>
    - <codes type="odv">
        <code local="SurveyNo" to="SDN:LOCAL:Cruise" from="" />
        <code local="Station" to="SDN:LOCAL:Station" from="" />
        <code local="Date" to="SDN:LOCAL:yyyy-mm-ddThh:mm:ss.sss" from="" />
        <code local="Longitude" to="SDN:LOCAL:Longitude, degrees_east" from="" />
        <code local="Latitude" to="SDN:LOCAL:Latitude, degrees_north" from="" />
        <code local="BotDepth" to="SDN:LOCAL:Bot. Depth, m" from="" />
        <code local="Type" to="SDN:LOCAL>Type" from="" />
        <code local="MinDepth" to="SDN:LOCAL:MinimumObservationDepth, m" from="SDN:P01::MINWDIST, SDN:P06::ULAA" qflag="fMinDepth"/>
        <code local="MaxDepth" to="SDN:LOCAL:MaximumObservationDepth, m" from="SDN:P01::MAXWDIST, SDN:P06::ULAA" qflag="fMaxDepth"/>
        <code local="SampleID" to="SDN:LOCAL:SampleID, #" from="SDN:P01::SAMPID01, SDN:P06::UUUU" qflag="fSampleID"/>
        <code local="Area" to="SDN:LOCAL:SamplingEffort, m^2" from="SDN:P01::AREABEDS, SDN:P06::UMSQ" qflag="fArea"/>
        <code local="SubSampleID" to="SDN:LOCAL:SubsampleID, #" from="SDN:P01::SSAMID01, SDN:P06::UUUU" qflag="fSubSampleID"/>
        <code local="SubCoef" to="SDN:LOCAL:SubSamplingCoefficient, none" from="SDN:P01::SSAMPC01, SDN:P06::UUUU" qflag="fSubCoef"/>
        <code local="LatinName" to="SDN:LOCAL:ScientificName, none" from="SDN:P01::SCNAME01, SDN:P06::UUUU" qflag="fLatinName"/>
        <code local="WormsID" to="SDN:LOCAL:ScientificNameID, #" from="SDN:P01::SNAMID01, SDN:P06::UUUU" qflag="fWormsID"/>
        <code local="Sex" to="SDN:LOCAL:Sex, none" from="SDN:P01::ENTSEX01, SDN:P06::UUUU" qflag="fSex"/>
        <code local="Stage" to="SDN:LOCAL:LifeStage, none" from="SDN:P01::LSTAGE01, SDN:P06::UUUU" qflag="fStage"/>
        <code local="Count" to="SDN:LOCAL:ObservedIndividualCount, #" from="SDN:P01::OCOUNT01, SDN:P06::UUUU" qflag="fCount"/>
        <code local="WetWeight" to="SDN:LOCAL:WetWeight, g" from="SDN:P01::SDBIOL05, SDN:P06::UGRM" qflag="fWetWeight"/>
        <code local="DryWeight" to="SDN:LOCAL:DryWeight, g" from="SDN:P01::SDBIOL08, SDN:P06::UGRM" qflag="fDryWeight"/>
    </codes>
</root>
```

Data published: Aarhus University





Data to be published soon

- IFREMER
 - Rephy
 - Phytoplankton
 - Update expected in October
 - REBENT - Réseau de surveillance du benthos
 - 1996 - ongoing
 - + 141.000 data records
 - RSL - Réseau de Surveillance Lagunaire
 - Benthos + Macroalgae
 - 1999 - ongoing
 - + 10.000 records



Data to be published soon

- **IMGW, Poland**
 - Polish-Russian Vistula Lagoon Monitoring
 - Phytoplankton, Phytoplankton, Zooplankton, Zoobenthos (>2008)
- **IMARES**
 - Monitoring of macrozoobenthos in the Voordelta (2004- 2011)
- **Aarhus University other datasets**
 - Macroalgae on stone reefs (1979 – ongoing)
 - Marine phytoplankton (1979 – ongoing)
 - Marine zooplankton (1979 – ongoing)

Data exchange

- Buffer created for data exchange with EMODNet Biology
 - Testing with DK Aarhus University
 - Selected data moved to data buffer by MARIS
 - Search filter with 'benthos' set as string search on Local_CDI_ID and CDI partner = 729
- Harvesting by EMODnet biology. Currently manual process
- VLIZ is working towards more automated process using full functionality of buffer set-up



PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

Data exchange

Buffer used to harvest the Benthos dataset Aarhus University

<http://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=4494>



The screenshot displays the EMODnet Biology Data Discovery and Access Service. At the top, there's a navigation bar with links for Home, Data portal, Data Catalog, Data products, Partners, Project, Contribute, and Contact. The main area features a map of Europe with a specific region highlighted in orange, representing a buffer around the British Isles. To the left of the map is a sidebar with a 'Search' field, a 'Active layers' tab (which is selected), and an 'All Layers' tab. Below the search field is a list of layers: 'Google Satellite' (radio button selected) and 'Danish benthic marin...' (checkbox checked). Underneath these are several small icons. Further down the sidebar is a legend for 'records/grid' with categories: ≤ 1 (light orange), ≤ 10 (orange), ≤ 100 (dark orange), ≤ 1000 (red-orange), and > 1000 (dark red). At the bottom of the sidebar is a scroll bar. The bottom part of the screenshot shows a data catalog table with columns for Dataset Name, RecordCount, RecordCount QC, and View. The first row in the table is 'Danish benthic marine monitoring data from ODAM' with values 794,717 and 746,403 respectively.

Dataset Name	RecordCount	RecordCount QC	View
Danish benthic marine monitoring data from ODAM	794,717	746,403	

Conclusions and outlook

- Specific format developed, described and tested
 - Feasible but not trivial
 - >1M data records/ >20k CDI 's made available
 - >500k data records in preparation
- Exchange with EMDN BIO made operational
 - Upgrade towards full machine-machine yet to be realised
- More specific services for QC and analysis to be developed
- Expanding # of datasets & providers
- Looking at some additional data types (bio sensors, ...)