Data inventory and gap analysis

Overview

- Objectives of the biological pilot project
- Gap analysis
  - ✓ EurOBIS
  - ✓ Encountered gaps
- Data inventory
  - ✓ Starting from known initiatives and networks
  - ✓ Partners and their contributions
  - ✓ Identifying long-term marine biological monitoring datasets
    - ➢ National
    - ➢ Regional
- Conclusions

Objective of the biological pilot lot

- ***Assemble fragmented and inaccessible marine data*** into interoperable, contiguous and publicly available data streams for complete maritime basins (*Tender*)
  - ✓ Analyse
    - ➢ Coverage and shortcomings

      Taxonomy, geography, temporal
  - ✓ Identify existing marine biological data
  - ✓ Focus on
    - ➢ North Sea, Bay of Biscay and Iberian coast
    - ➢ Different species groups

Gap analysis

- Goal
  - ✓ Determine taxonomical, geographical and temporal gaps
  - ✓ Define the limitations of the data usage
  - ✓ Define possibilities
- Carried out end 2009 (207 datasets)
- Based on public available data in EurOBIS

EurOBIS ?

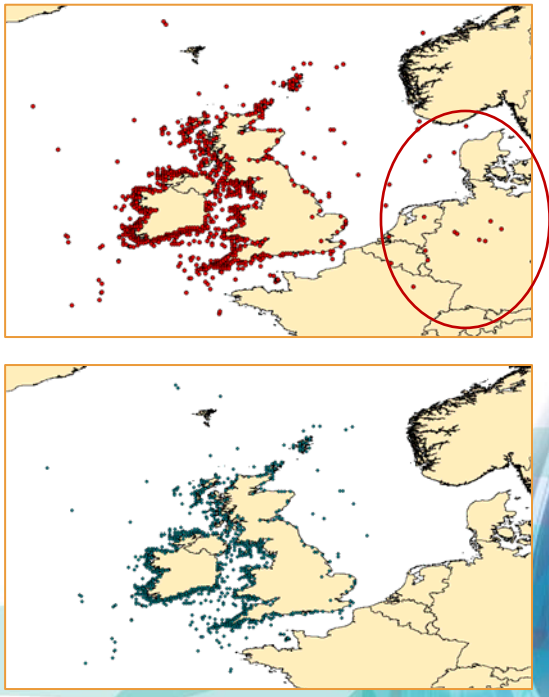
- EurOBIS – European Ocean Biogeographic Information System (2004)
- Biogeographic data on marine species:

  taxon-name, latitude, longitude, date
- Freely available
- Quality controlled data
- International data flow
  - ✓ Ocean Biogeographic Information System (OBIS)
  - ✓ Global Biodiversity Information Facility (GBIF)

Quality control

- Taxonomy
  - ✓ All taxa matched to the World Register of Marine Species
  - ✓ Saveguard originally delivered name
  - ✓ Doubt: feedback to data provider

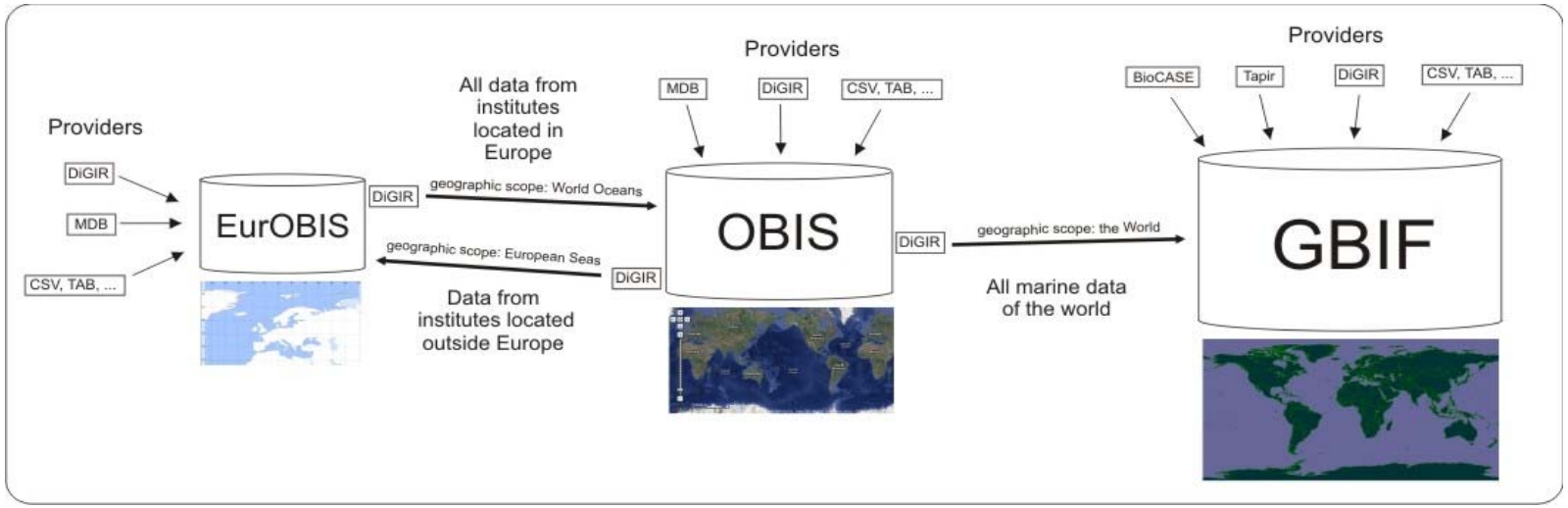
- Geography
  - ✓ Plot sampling locations on a map
  - ✓ Check for odd locations
  - ✓ Doubt or errors: feedback to data provider

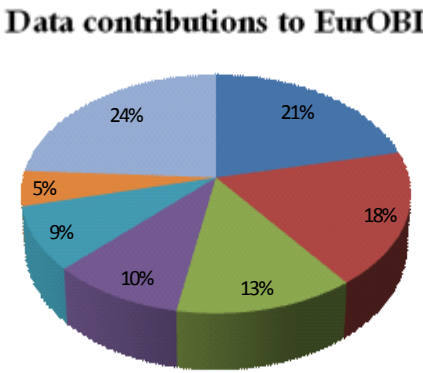


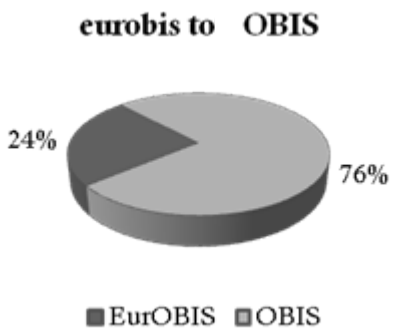
Sightings and strandings of marine turtles around the coast of UK and Ireland
Top: coordinates as received; bottom: corrected. Errors due to missing minus sign

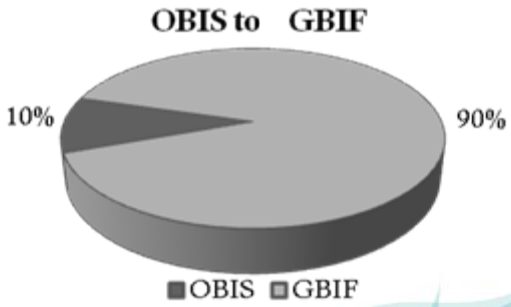
EurOBIS

International data flow



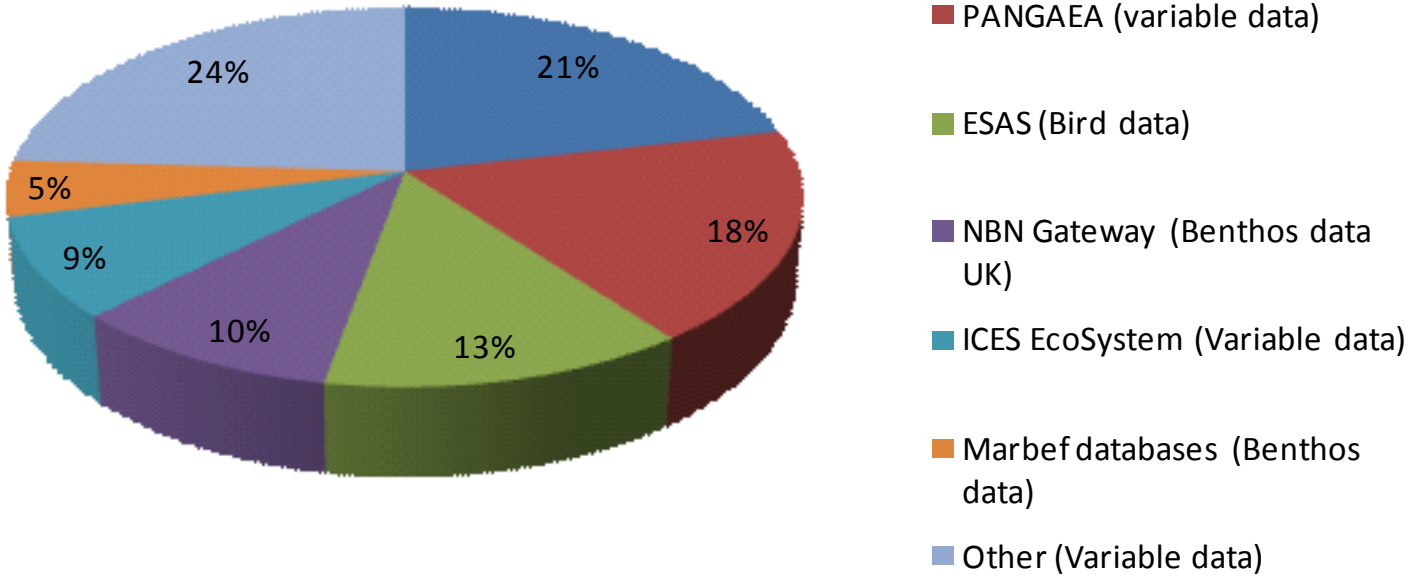




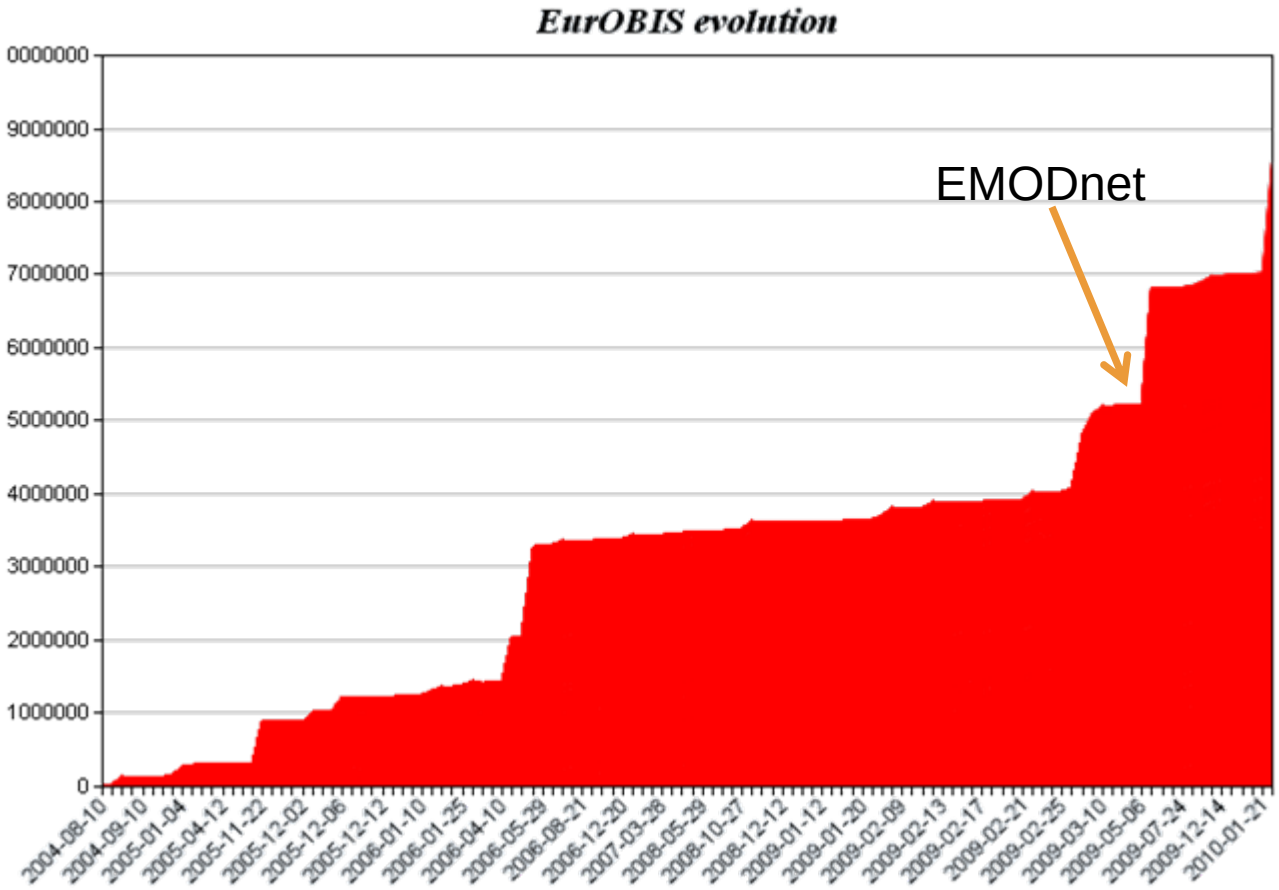


Available data

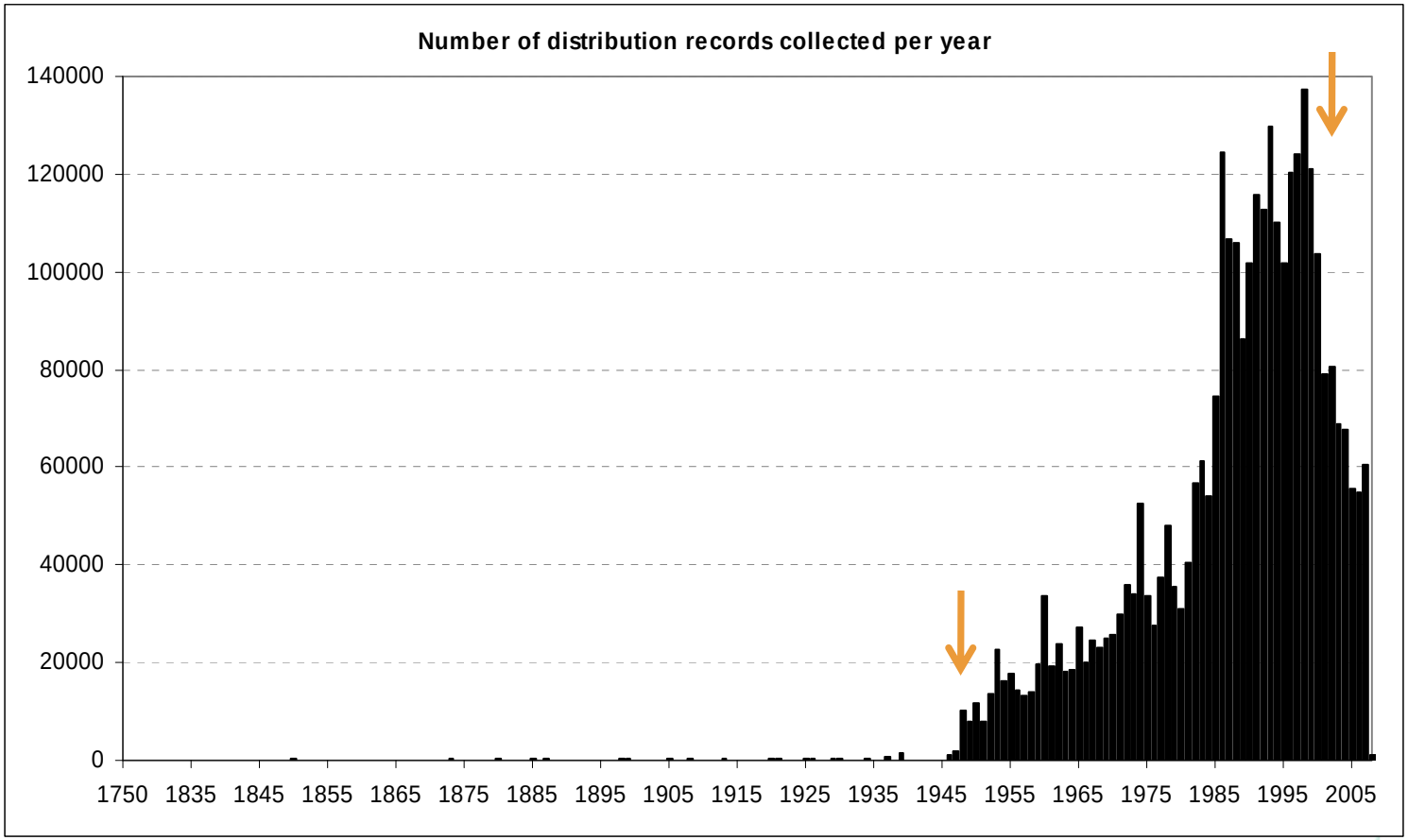
8,560,805 distribution records – 214 datasets



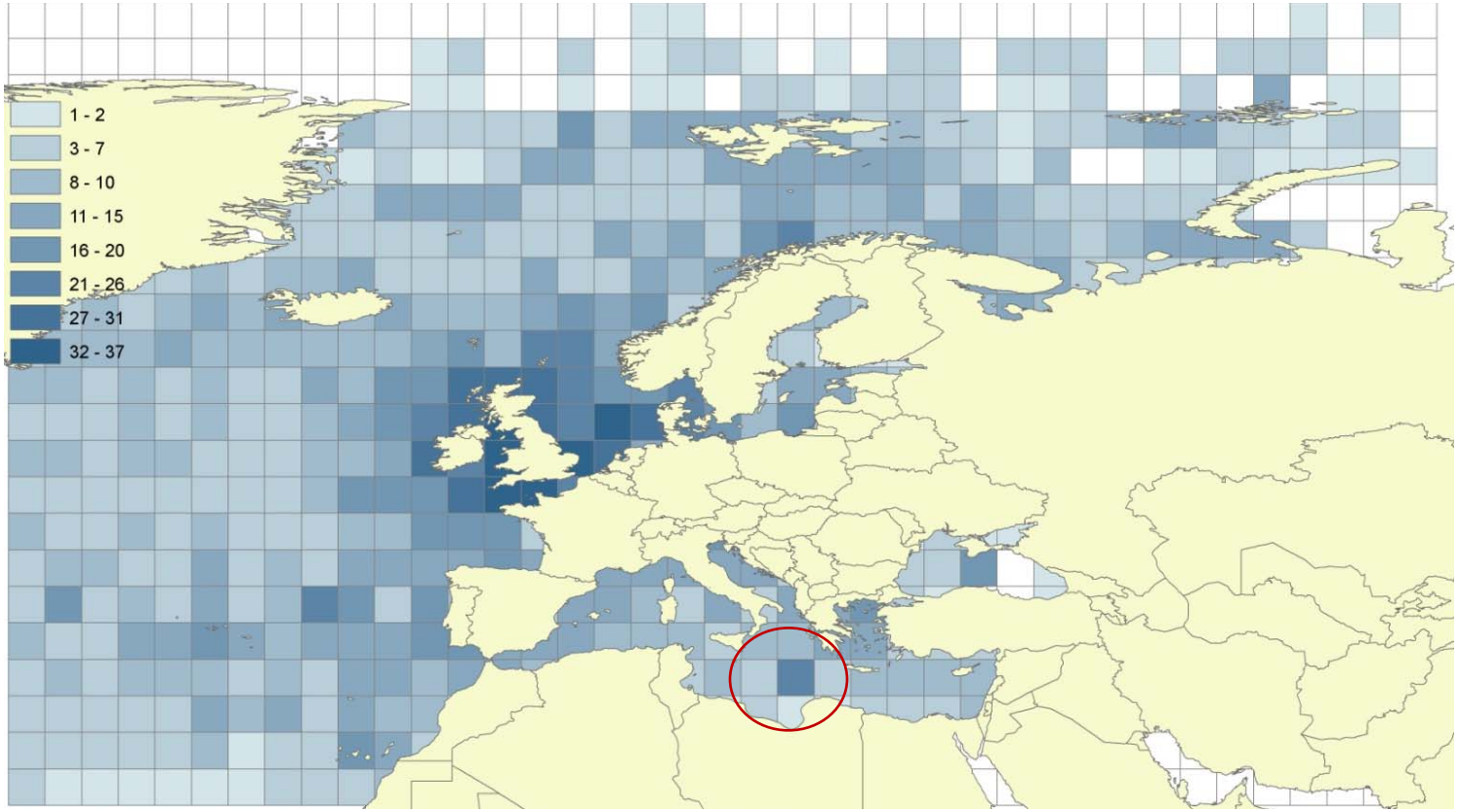
EurOBIS


EurOBIS evolution
0000000
9000000
8000000
7000000
6000000
5000000
4000000
3000000
2000000
1000000
0
EMODnet
2004-08-10
2004-09-10
2005-01-04
2005-04-12
2005-11-22
2005-12-02
2005-12-06
2005-12-12
2006-01-10
2006-01-25
2006-04-10
2006-05-29
2006-08-21
2006-12-20
2007-03-26
2008-05-29
2008-10-27
2008-12-12
2009-01-12
2009-01-20
2009-02-09
2009-02-13
2009-02-17
2009-02-21
2009-02-25
2009-03-10
2009-05-06
2009-07-24
2009-12-14
2010-01-21

Gaps: temporal coverage



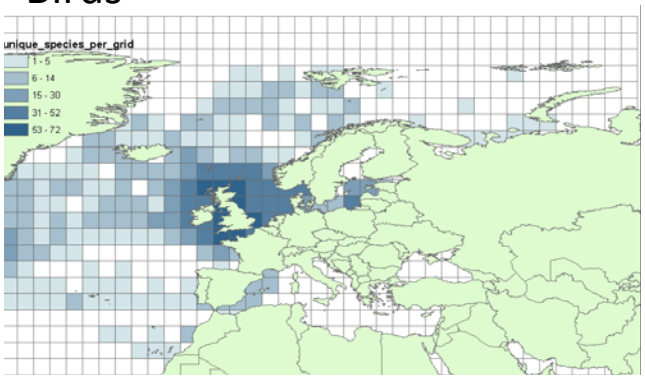
Gaps: geographical coverage



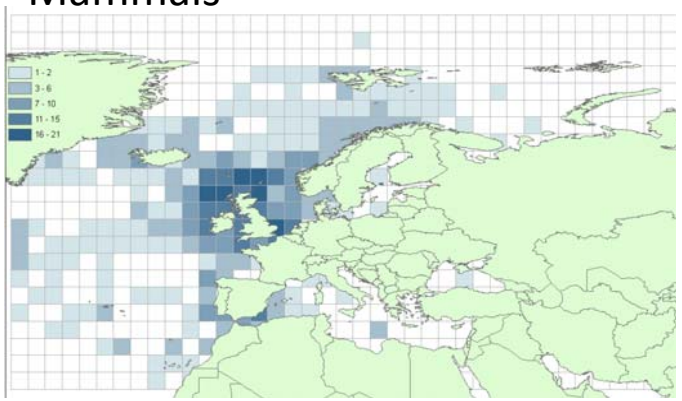
Analysis grid representing the number of distinct larger taxonomic groups (Mammals, Aves, Crustacea, Rotifera, Tunicata...) per grid-cell

EurOBIS

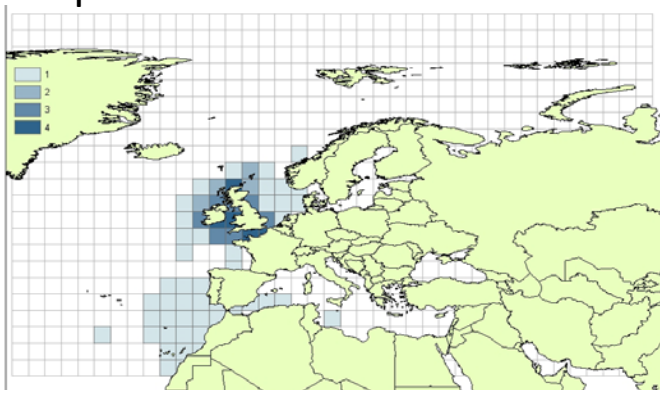
Birds



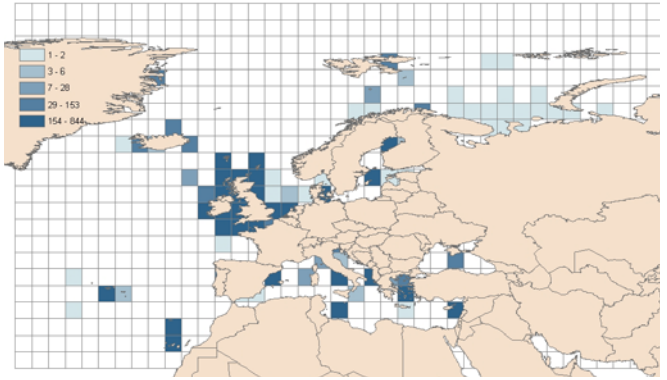
Mammals



Reptiles



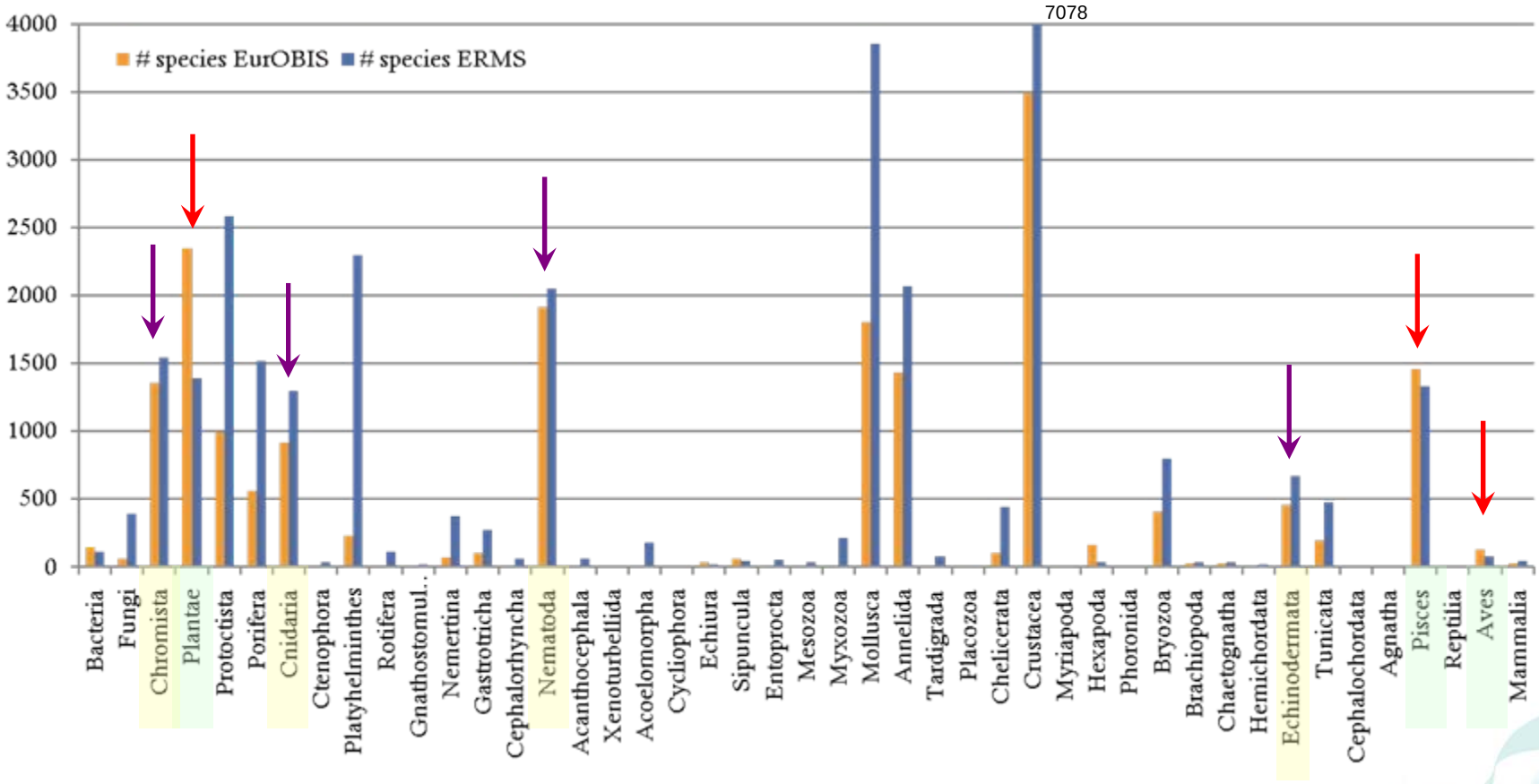
Plants



No examples of general biodiversity for a group within Europe!

Representation of data coverage for each group within Europe.

Gaps: taxonomic coverage



Number of species per higher taxon group in EurOBIS compared to ERMS

Some statistics

- 69 % species distribution records
- Almost 19 000 accepted species names (31 518 in ERMS)
- 69 % of species distribution records with indication of sampling year
- 4.85% species distribution records have abundance information
- 14 % species distribution records with life-stage information

- 40 datasets with temporal span larger than 10 years
- 30 uninterrupted time-series of more than 10 years

Data inventory

Starting from known initiatives and networks

- Goal of inventory
  - ✓ Identify existing marine biological datasets
- How to accomplish?
  - ✓ Contact partner networks
    - EurOBIS, OBIS and GBIF *(see gap analysis)*
    - MARS Network (European Biological Research Stations and Institutes) and data legacy FP6 MarBEF
    - ICES data centre
    - Pangaea World Data Centre
    - IBSS and Black Sea Biological Network
    - SeaDataNet

Data inventory

MARS Network and data legacy FP6 MarBEF

- Data collected through
  - ✓ FP5 Biomare (2000-2002)
  - ✓ FP6 MarBEF (2004-2009)
  - ✓ MARS partners (2009-2010)

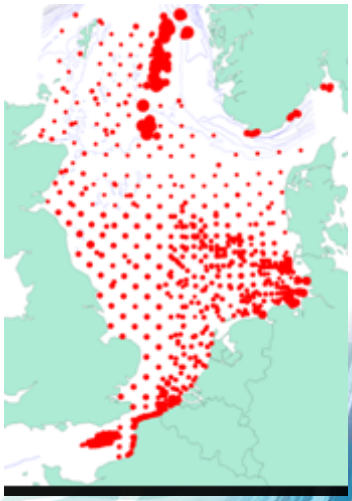
Location of MARS member institutes

- Data inventory:
  - ✓ Total of 257 datasets
  - ✓ Sent to over 100 people
  - ✓ Question 1: knowledge of additional data?
  - ✓ Question 2: data available for EurOBIS-EMODNet?
  - ✓ Average response (±30%)

Data inventory

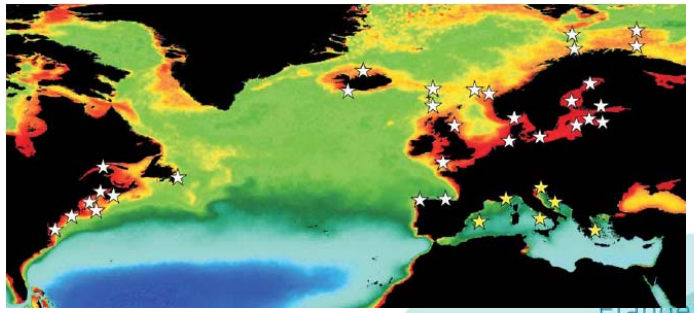
ICES

- Data centre
  - ✓ DATRAS trawl surveys, 45 years
  - ✓ Biological community & contaminants and biological effects, 30 years
  - ✓ Predator-prey, 12 years
  - ✓ STATLANT official catch statistics, 60 years
  - ⇒ *All will contribute to EMODNET, over 6 million records*



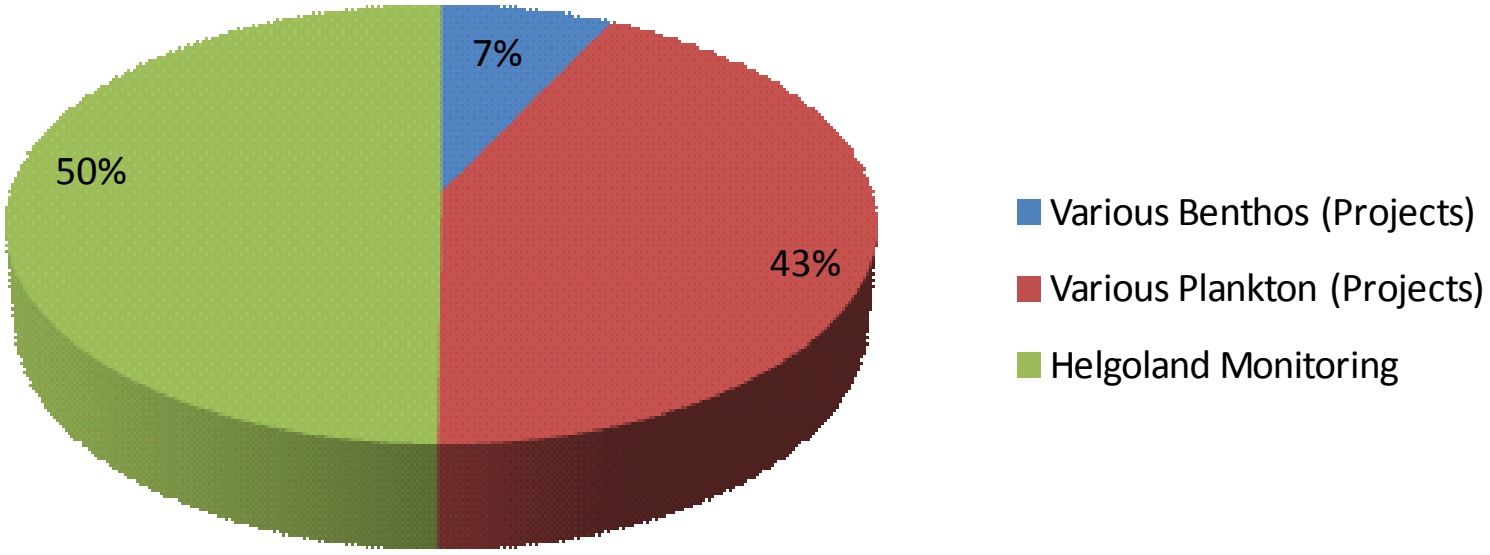
- Working groups
  - ✓ … on benthos ecology
  - ✓ … on zooplankton ecology

  => *Partly available*



Data inventory

World Data Centre MARE / PANGAEA



- 1.5 million distribution records
- Important plankton contribution

=> *Already available to EMODnet*

Data inventory

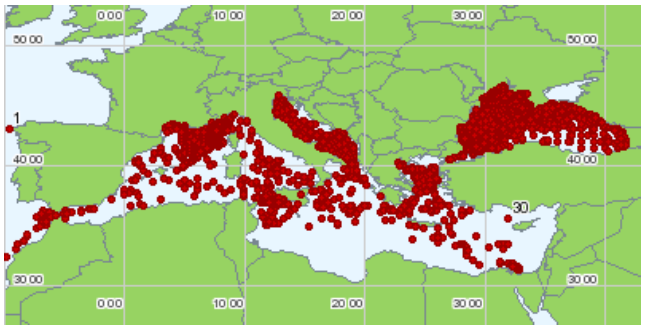
IBSS and 'Black Sea Marine Biological Network'

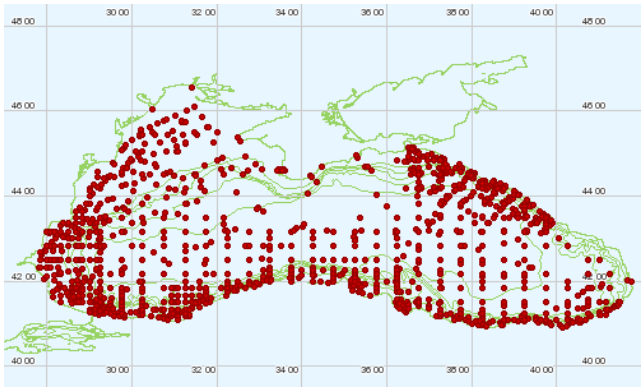
- Data inventory:

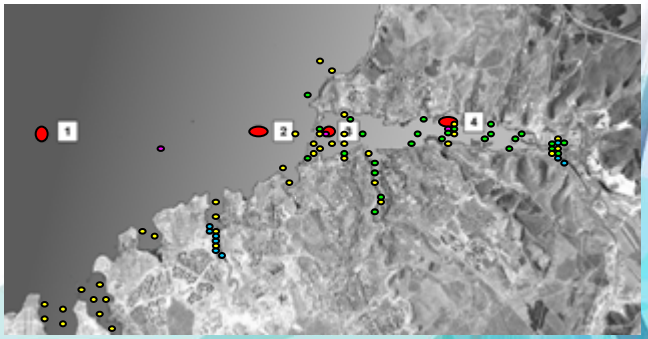
  Total of 17 Black Sea datasets

  - ✓ IBSS historical data from different cruises (1877- …)
  - ✓ Zooplankton (1954 - …)
  - ✓ Phytoplankton (1968 - …)
  - ✓ Cetacea (1988 - …)
  - ✓ Black Sea *Mnemiopsis leidyi* and *Beroe ovata* database
  - ✓ Stock assessments Black and Azov Sea
  - ✓ Sebastopol Bay plankton monitoring
  - ✓ NATO TU – Black Sea (1954-1996)

  => Some data already available







Data inventory

SeaDataNet

- FP6 Network of NODC's
- 49 partners (European NODC's)
- Link to physical-chemical data
- Metadata inventory: in preparation

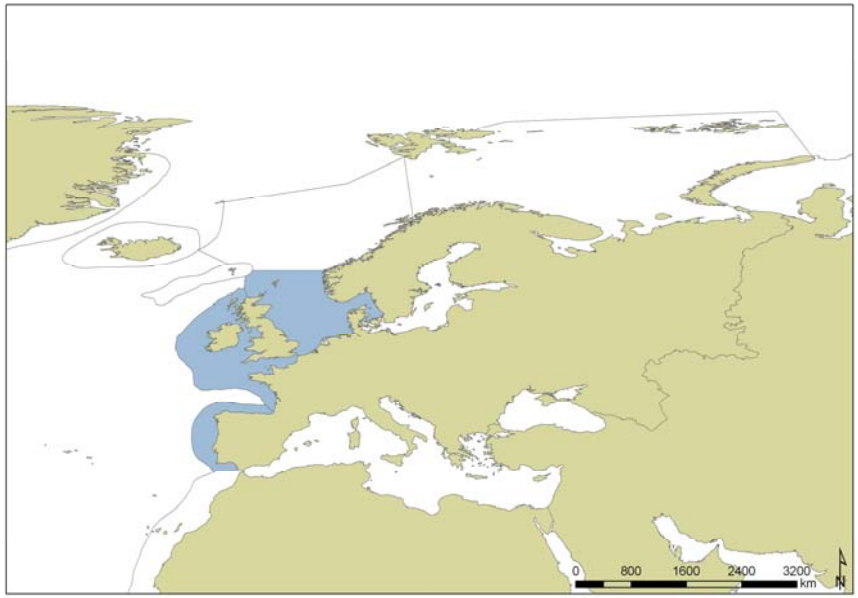
Data inventory

Identifying long-term marine monitoring datasets

- Additional search for long-term biological data series
- Two approaches
  - ✓ National marine biological monitoring
  - ✓ Regional marine biological monitoring
- Work in progress

Data inventory

Focus



Indication of the focal area of the Biological lot (blue)

- Area
  - ✓ Bay of Biscay
  - ✓ Iberian coast
  - ✓ Greater North Sea, including
    - ➢ Kattegat
    - ➢ English Channel
- Groups
  - ✓ Benthos
  - ✓ Plankton
  - ✓ Macro-algae
  - ✓ Reptiles
  - ✓ Birds
  - ✓ Mammals

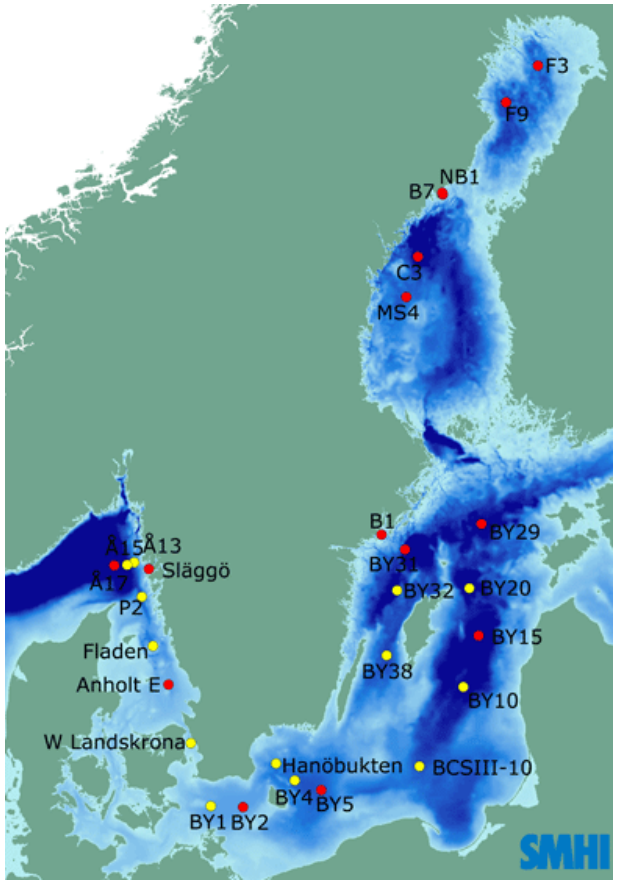
Data inventory

Identified national marine biological monitoring data

Country	Groups	Temporal scope
Sweden	benthos, plankton, mammals	1971 - present
Denmark	benthos, plankton, algae	1979 - present
Germany	benthos, plankton, birds, reptiles, algae	1973 - present
Netherlands	benthos, plankton, birds, plants, mammals, bacteria	1948 - present
Belgium	benthos, birds	1979 - present
UK	benthos	1990's - present
Ireland	plankton, mammals	1990's - present
France	benthos, plankton	1987 - present
Spain	plankton	1987 - present
Portugal	*No specific national program*	

Data inventory

Sweden

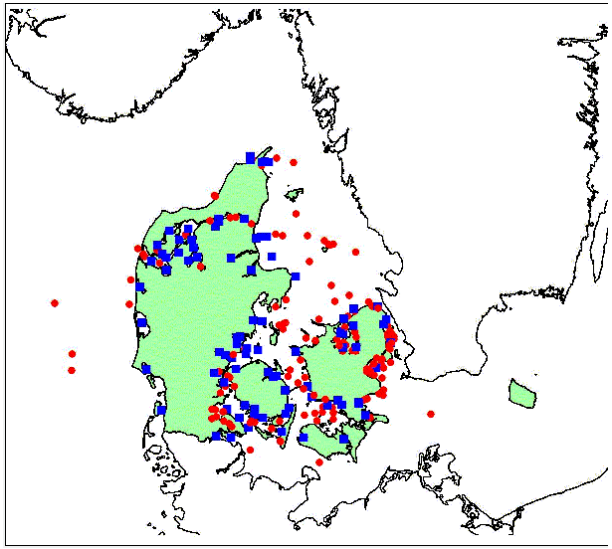


- Marine biological monitoring data – SHARK
- Swedish Meteorological and Hydrographical Institute - SMHI
- Phytoplankton, zoobenthos, phytobenthos and seal observations
- 1971 onwards
- Online available

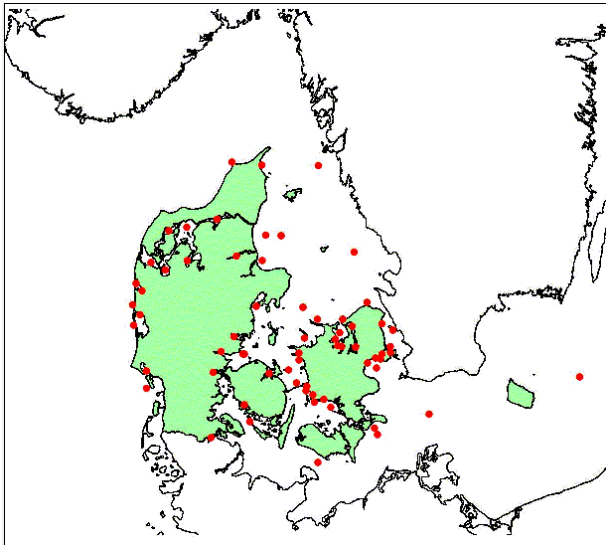
Data inventory

Denmark

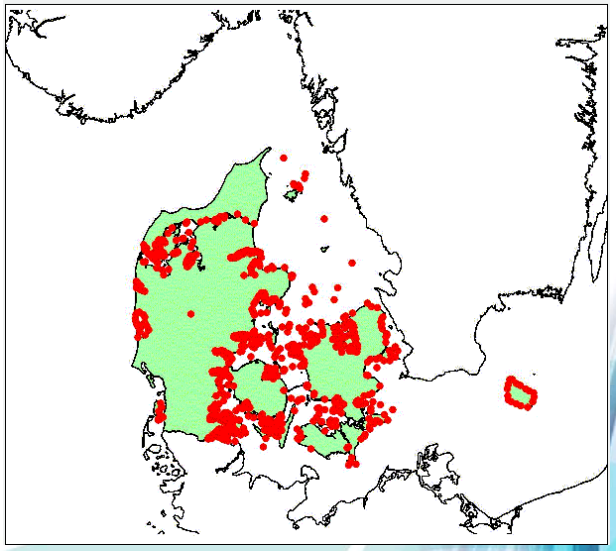
- The Danish national database for marine data – MADS
  - ✓ National Environmental Research Insitute (NERU)
  - ✓ From 1979 onwards: 700 sites, 1,000 species
  - ✓ Plankton, benthos, algae and plants
  - ✓ Online available



benthos



phytoplankton

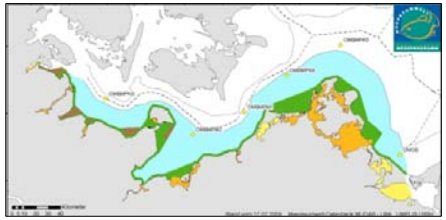


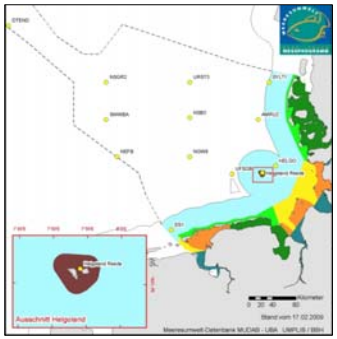
macro-algae & plants

Data inventory

Germany

- German Marine Monitoring Programme – GMMP (BLMP)
  - ✓ Federal Environmental Agency
  - ✓ Phyto-, zooplankton, macrophytes, macrobenthos, fish, birds, mammals
  - ✓ From 1973 onwards
  - ✓ National waters: North Sea & Baltic





- Helgoland Roads Time Series
  - ✓ AWI
  - ✓ Phyto-, zooplankton, nutrients
  - ✓ From 1965 onwards

- Trilateral Monitoring and Assessment Program (TMAP)
  - ✓ Germany – Denmark – Netherlands: Wadden Sea
  - ✓ Phytoplankton, plants, macrobenthos, fish, birds, mammals
  - ✓ From 1991 onwards

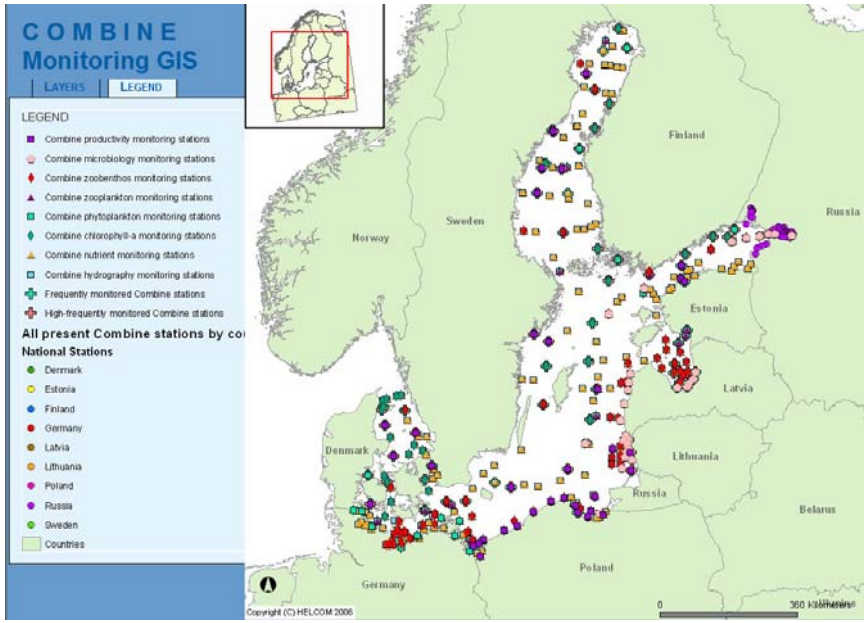
Data inventory

Identified regional marine biological monitoring data

	Groups	Temporal scope
HELCOM	benthos, plankton, chlorophyll	1979 - present
EEA	chlorophyll	1980 - present
ICES	*discussed earlier*	

Data inventory

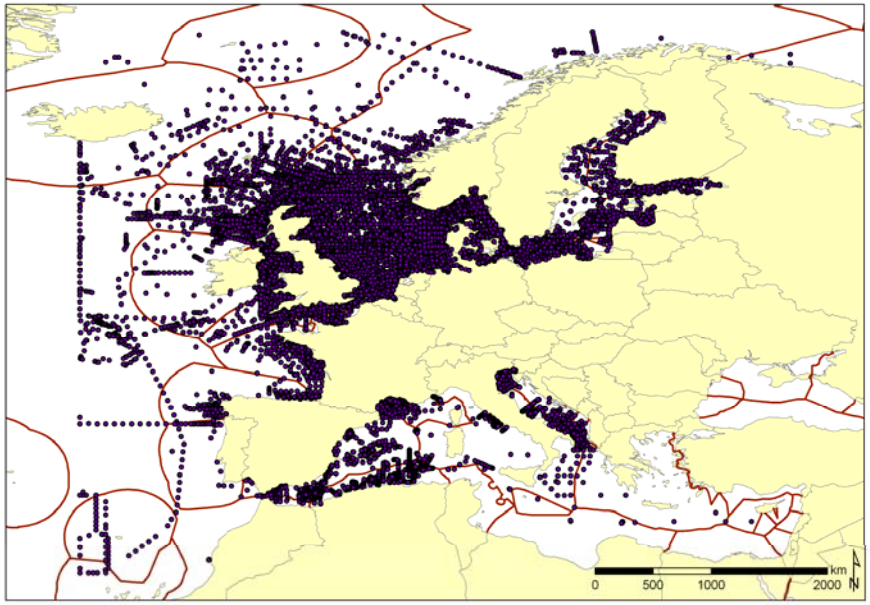
HELCOM



- Co-operative monitoring in the Baltic environment (COMBINE)
  - ✓ 9 countries, 459 stations
  - ✓ Benthos, plankton
  - ✓ Since 1979
  - ✓ Biological data through ICES
- Regional coastal fish monitoring (HELCOM Fish)
- Thematic databases:
  - ✓ Alien species
  - ✓ Species occurences in Baltic Sea Protected Areas (BSPA)

Data inventory

EEA



Chlorophyll a sampling stations within EEA database

- Waterbase – transitional, coastal and marine waters
  - ✓ Chlorophyll a and oxygen
  - ✓ Data are (sub-)samples of national data from 16 countries, 1980
  - ✓ Online available
- Biological data: ECOSTAT
  - ✓ Benthos, macro-algae, plants
  - ✓ Compiled under the implementation of the Water Framework Directive

Conclusions – gap analysis

Data limitations

- ✓ Missing information
  - ➢ Abundance, biomass, …
  - ➢ Sampling methodology : necessary for standardisation
  - ➢ Not georeferenced records
- ✓ Taxonomy: species not yet represented
- ✓ Geography:
  - o Mediterranean data underrepresented
- ✓ Time:
  - o Day-month-year essential for temporal analyses
  - o Pre-fifties and post 2002

Conclusions – gap analysis

Data limitations (continued)

- ✓ Diverse information
  - ➢ Research versus monitoring
  - ➢ Literature versus field
  - ➢ Aggregated versus species

Conclusions – gap analysis

However…

- Reliable data available
- "Endless" possibilities with public data
  - ✓ Presence-absence maps
  - ✓ Area-comparisons
  - ✓ Calculation of diversity indices
  - ✓ …
  - ⇒ Requires critical mind and thorough selection of data
  - ⇒ Never possible on all data combined, only on subsets
- Derived dataproducts from non-public data
  - ✓ Abundance data, …

Conclusions – data inventory

- Good progress towards complete biological data inventory for EU
- Ongoing => more information expected
- Need to bring actual data together

  Keep in mind…
  - Embargo period for research data
  - Regional reporting of national data => double entries!
  - Presence versus abundance/biomass data
  - Aggregated information on regional level
  - Certain information needed for good integration

Thank you