



SeaDataNet

*PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT*

What more could be done ...

Report from the ICES Data Centre

Status - ICES data in SDN

Dataset	Number of stations in SDN
Contaminants in Sediment	414
Contaminants in Seawater	789
Contaminants in Biota	1273
Phytoplankton	1424
HydroChemistry*	19902
Total	23802

*Low resolution CTD and Bottle data with Temperature, Salinity, Oxygen, Nutrients, Chlorophyll a and pH/Alkalinity.

Challenges – Duplication and Mapping

- **Duplication**
 - As a regional data centre ICES databases contains data from some of the national data centres also submitting data directly to SDN as a partner.
 - As SDN obviously wants to avoid duplicates only a subset (known not to be in SDN) of the potential data available in ICES have been made available for SDN so far.
- **Mapping**
 - Mapping to P01 and EDMO has proven not to be straight forward.

Challenges – Duplication

- **Contaminants**
 - In order to avoid duplicates for Contaminants only P01 entities that have been created specifically per request from ICES are included.
- **HydroChemistry**
 - In order to avoid duplicates for HydroChemistry only data from one institute have currently been included.

Challenges – Mapping

- P01
 - Only EMODNET Chemistry priority substances have been mapped so far.
 - For Contaminants in Biota 379 chemicals in 115 species = 9061 P01 codes (combinations of Species, Matrix, Substance, Basis of determination).
- EDMO
 - (ICES) RLABO to EMDO.
 - Provider/Originator versus Platform.

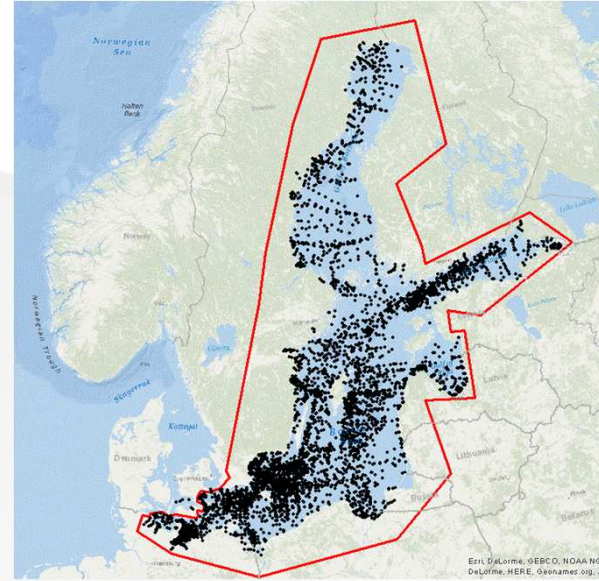
Potential – A case study in the Baltic

- In order to estimate the potential number of ICES data not yet in SDN, ICES received the EMODNET Chemical data buffer with Nutrients in the Baltic from MARIS and compared them to Baltic Nutrients stations within the ICES database.

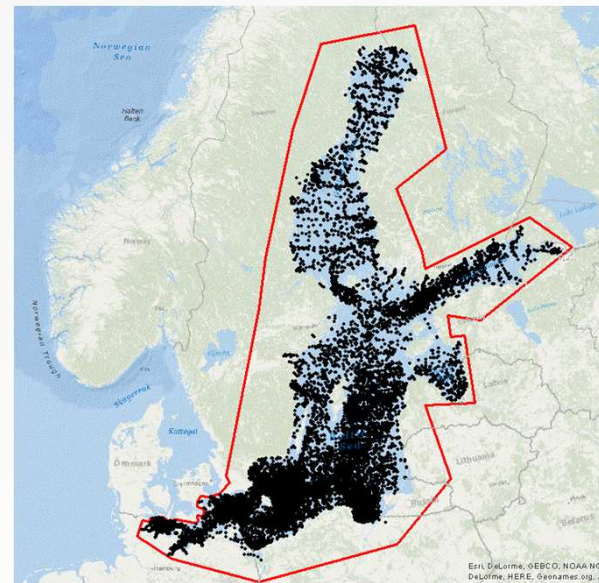
Potential – a case study in the Baltic



SDN



ICES



+ parameter criteria

Potential – A case study from the Baltic

- SDN chemical stations (extracted 2014-10-28):
58778 stations.
- ICES chemical stations (extracted 2015-04-31):
117995 stations.
- 59217 stations more in ICES not in SDN.
- When taken duplicates into account e.g. within 3
minutes and approximately 100 m:
- 62063 stations in ICES not in SDN.

What more could be done ...

- Due to the duplication issue only a subset of the potential data in ICES can currently be made available for SDN.
- Implementation of a central CDI duplication identification routine and flagging system would remove this bottleneck.
- Better safe than sorry!

What more could be done ... SDN III

