ROYAL BELGIAN INSTITUTE OF NATURAL SCIENCES, OD NATURAL ENVIRONMENT

BACK TO THE SEA OF THE PAST TO EVALUATE THE STATUS OF TODAY: INTEGRATION AND VALORISATION OF HISTORICAL DATA

R. Lagring, K. De Cauwer, M. Devolder, Y. Stojanov, S. Scory
Royal Belgian Institute of Natural Sciences, OD Natural Environment, Belgian Marine Data Centre, bmdc@mumm.ac.be

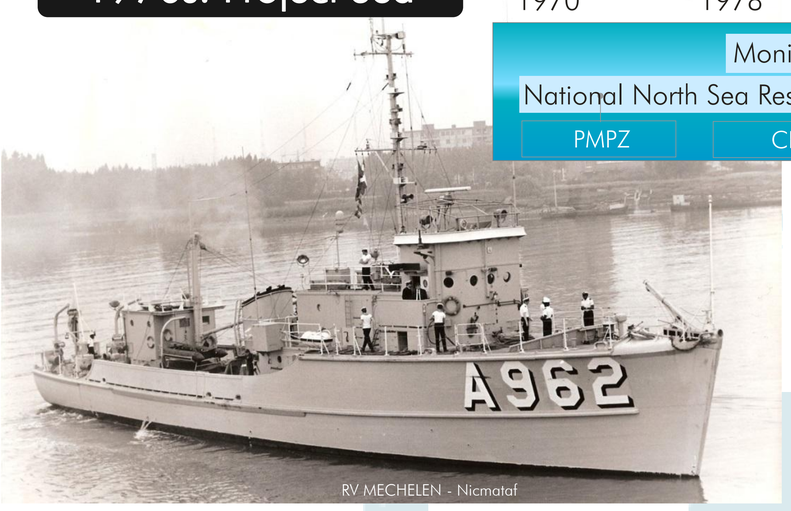
To evaluate the status of our seas today we have to go back to the status of the past, therefore historical marine data are an important scientific heritage. In Belgium, systematic measurement campaigns started with Project Sea (1970-1976). A massive amount of data was produced from various disciplines, but results were not centralised in modern databases and therefore these datasets were not readily available. During data rescuing in the project PMPZ-DBII, historic data from 1970 until 1982 have been recovered and long-term datasets (40 years) are now online available. The historical data needs to be normalised and integrated with more recent results before environmental change can be determined.

LONG-TERM DATA COLLECTION

1970 — 1978 — TODAY

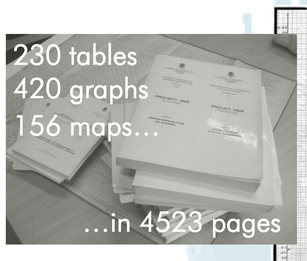
- Monitoring (National, OSPAR,…)
- National North Sea Research Programme
  - PMPZ
  - CRA
  - SMNS
  - SPSD
  - SSD
  - BRAIN

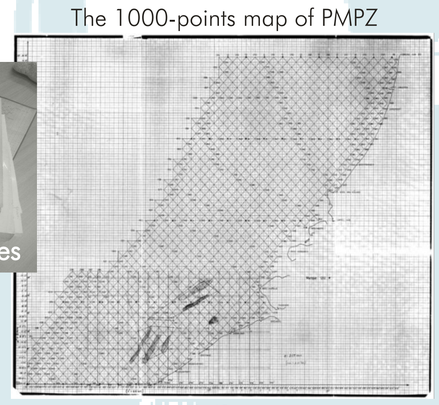
1970s: Project Sea



RV MECHELEN - Nicmataf

The 1000-points map of PMPZ





Project Sea (PMPZ): Start of modern oceanography in Belgium with systematic measurement campaigns

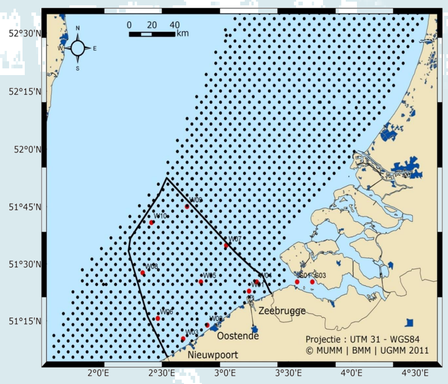
- Environmental problems due to human activities under world wide attention (first Earth Day) → pressure EEC.
- Extensive national marine research with **multidisciplinary** character: >200 physicists, chemists, biologists and geologists.
- Results: massive amount of data on paper, not centralised and **not readily available.**

Today: Data rescuing

The process of data rescuing is a three-step procedure

1. INVENTORY

Create full **overview** of the project with all data to be expected, via screening and tracing back all possible sources (final reports, maps, technical reports, scientific publications etc.)



2. DIGITISATION

- Assign **database** for long-term depository: **IDOD** (Integrated and Dynamical Oceanographic Data management system) at BMDC.
- After **prioritisation: digitise** old paper documents or transform old databases.
- Trace back **missing metadata** by crosslinking information or screening other publications.
- Perform **quality control** and **import** rescued data.
- **Normalisation:** Render historical data comparable to more recent results.

3. DISSEMINATION and VALORISATION

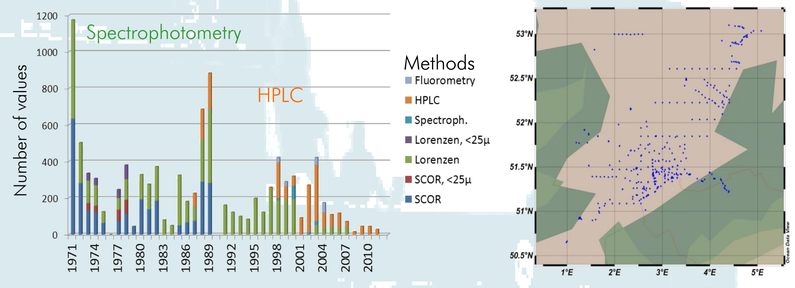
http://www.mumm.ac.be/datacentre

Extended datasets

Historical data from 1970 until 1982 have been recovered

- Datasets in IDOD (current monitoring and research programs) have been extended and the data has been made internationally available via SeaDataNet Services.
- >120 Campaigns and ca. 1600 sampling events.
- Water values for pigments (5864), nutrients (3298), heavy metals (2302), physical measurements (3295).
- Sediment values for heavy metals and pesticides (1649).

Chlorophyll a values in IDOD



Obstacles

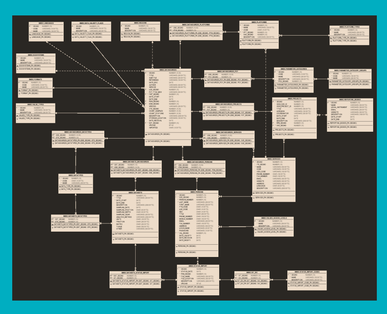
Reconstruction of extensive historical research actions is complex and many obstacles need to be overcome:

- High variety of data, scattered over the country while many laboratories have been closed down or reorganised.
- Aggregated values and maps were published in final reports, need for original values.
- Missing metadata and gaps, and lack of documentation slow down the process.

Data-tracking system

For an efficient data recovery process, a data-tracking system is being designed to assess the work, identify the gaps and evaluate the progress:

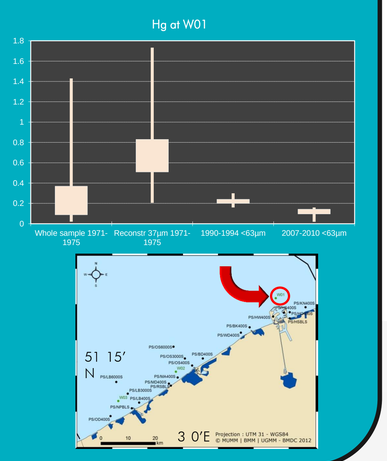
What (meta)data is missing? Are methodologies of analyses and sampling fully documented? Where are the documents located and who is contact person? What datasets are complete and quality checked? What datasets have been digitised and imported?


Oracle Database

Case study on heavy metals in coastal sediment

Hg is one of the parameters that was systematically measured during Project Sea by IHE at 28 coastal locations

- Time series were investigated testing different methods for trend detection, *ia*. ROB-LAR.
- Original values are available in final reports, however sampling and analytical methodology, like analysed sediment fraction, were not clearly described: concentrations refer to the whole sample and not to the fraction <37 μm as initially assumed.
- Hg appeared to be significantly decreasing between 1990 and 2010 at station W01. For Project Sea data, only an indication of temporal evolution could be given, taking into account the possible ranges.



CHALLENGES FOR COMPARING HISTORICAL AND CONTEMPORARY DATA

- Importance of **metadata**: Missing or incomplete documentation of metadata (e.g. analyses methodology, normalisation) will result in erroneous data, which should be detected before data import. During data analyses and interpretation in long-term studies, the metadata always has to be taken into account.
- Importance of **data-tracking system**: For an efficient data recovering process with progress assessment of the work and possibility to identify suitable datasets for long-term trends (e.g. no gaps over time, enough data, localisation of data).
- Importance of **data consistency**: Before historic data can be compared with current data in long-term trends, each dataset requires thorough quality control, intercalibration and normalisation. Protocols need to be established to deal with changes in sampling strategies and methodologies and to determine conversion factors and error estimates for data intercalibration.

Design by Isabelle Bachy, Royal Belgian Institute of Natural Sciences, OD Natural Environment

Projects aiming at digitising results of past marine research programmes are funded by the Belgian Science Policy Office