

Back to the sea of the past to evaluate the status of today: Integration and valorisation of historical data

Ruth Lagring, Karien De Cauwer, Mia Devolder, Serge Scory

Royal Belgian Institute of Natural Sciences
Management Unit of the North Sea Mathematical Models (RBINS-MUMM),
ruth.lagring@mumm.ac.be (Belgium)

Historical marine datasets are an important scientific heritage. Not only do they allow us to give an idea on the status of the environment at the time of sampling, more importantly, long-term dataserieS enable the assessment of global climate change providing insight in the evolution over the years. In the frame of the EU Marine Strategy Framework Directive, the ‘Good Environmental Status’, for which descriptors like eutrophication and contaminants were defined, has to be determined. To increase scientific knowledge of these descriptors, we have to go back to the data compiled in the past. However, long-term dataserieS are not readily available and historical data, which are at risk of being lost, need to be recovered, integrated and quality checked before they can be compared with current data. In project PMPZ-DBII (2009-2012), data of 1970-1982 from the Belgian Continental Shelf have been recovered and integrated in the existing database and studied to assess usefulness for long-term studies.

In Belgium, modern oceanography with systematic measurement campaigns for assessing the quality of the marine environment, started in 1970 with the ‘Projet Mer/Projekt Zee’ or PMPZ (1970-1976). All compartments of the ecosystem were studied with the collaboration of over 200 scientists to explore the Belgian Continental Shelf (BCS). This initial program was followed by other research actions, programs and monitoring campaigns, resulting in a vast quantity of scientific data of the BCS since 1970. Techniques for digitization were not as advanced as today and data that were obtained via measurement campaigns aboard ships were not centralised in adequate databases. In the frame of PMPZ-DBII, the objective was to centralise and integrate the PMPZ-data with current data originating from monitoring and research programs, and to make them publically available. The IDOD database or “Integrated and Dynamical Oceanographic Data management system” was specifically developed to deal with marine environmental data and is as such suitable for these long-term datasets.

Original data and metadata from the years 1970 until 1982 could be recovered from technical reports, cruise reports and grey literature, digitized and imported. The reconstruction of the extensive historical research actions was very complex and many obstacles during the import process, which will be further described, needed to be overcome, like:

- A high number and high variety of data was dealt with which were scattered over the country and mostly in paper format.

- Final reports contain aggregated values and maps which are difficult to import in a database. Original values, which could be imported, were found in the technical reports.
- Available information was limited and metadata was often missing, making it very time consuming to recover.
- Information needed to be cross-linked to reconstruct the complete data and their metadata, like values, samples, campaigns, positions, dates, depth, sampling gears, analyses methods, data-origins, platforms, etc.

The IDOD database has been extended with data from over 120 campaigns and about 1600 sampling events, like water values for pigments (5864), nutrients (3298), heavy metals (2302), physical measurements (3295) and sediment values for heavy metals (1649). As a result, today the IDOD database has dataseries available for over 40 years, accessible via the BMDC website (<http://www.mumm.ac.be/datacentre/>).

Finally, as a first attempt to valorise the PMPZ-data, the evolution of heavy metals and pesticides in the sediment of the BCS from 1970 until today has been studied. During this assessment, a lot of limiting factors had to be dealt with and it was found that important meta information was lacking, which could be retrieved by consulting the data-originator. For example, the sediment fraction analyzed was not clearly described in the reports. After cross-checking data of different tables and testing the possible hypotheses, the fraction could be derived and corrected. However, it appeared to be smaller than the fraction used in the current monitoring practices (63 μm). Recalculated concentrations of lindane and mercury, appeared to be higher than the recent concentrations. It was concluded that a thorough study on normalization of these historical data in the BCS is needed to obtain a reliable comparison with current data. The project has learned us to cope with characteristics of historical datasets and that, in order to valorise and use them for long-term assessments, specific additional efforts are required to quality check and intercalibrate the data. The results of this study will be presented.