

MarBEF data management

By Edward Vanden Berghe

Data integration is one of the main objectives of the MarBEF Network. The data management team is actively looking for datasets, to increase the mass of data available through the network. Also, serving the information needs of the MarBEF community remains an active point of attention.

Too many good datasets are lost. Either data are lost through lack of proper management – failing back-ups or no back-ups at all – or data just ‘disappear’ at the end of a project, when the people actively collecting the data change jobs or leave the institute. One of the most important objectives of the MarBEF network has been, from the outset, to prevent this from happening. The data management team actively works with scientists in the MarBEF community, to assist them with data management tasks.

Internally, we use a system for archiving any datasets that are contributed by our scientists. To keep track of the datasets, and to document where they came from and what data they contain, we use the ‘Integrated Marine Information System (IMIS),’ a system that is also used for the ‘Register of Resources,’ which most MarBEF members will be familiar with. The metadata part is also visible online through the RoR, and can, to a certain extent, be edited online (through Data>Datasets>Submit on the MarBEF website).

We plan to open up the system further, offering the MarBEF community a facility to upload data files, and having them archived on our servers.

More about this in a future newsletter or weekly news mail.

Often scientists are not aware of proper data management procedures. Unlike statistics or other ‘supporting’ sciences, data management is only very rarely part of the university curriculum. To help alleviate this problem, MarBEF will organise a training workshop on marine biodiversity data management. The workshop has been held already once, and is now planned for a second time, and again in collaboration with the International Oceanographic Data and Information Exchange (IODE) programme of the Intergovernmental Commission (IOC). More information can be found on the Training pages of the MarBEF website.

But the daily work of the data management team remains integration of data. In this, there are two streams of activity. A first series of activities is related to ERMS and EurOBIS – data that becomes publicly available as soon as we receive it from the data providers. Then there is a second set of activities, dealing with data that are only made available within the context of a specific project: these are the soft-bottom Macrobenenthos database (MacroBen)

of Theme 1, the meibenthos database of Manuela RMP, and several databases that are planned now in the framework of the LargeNet RMP.

Usually the data in this second type of database are made available to fellow data contributors, for the sake of joint analysis. These two activities have common points, of course: they use the same standard taxonomic lists, the same georeferencing... And, most importantly, most data providers agree on making the data public through EurOBIS as soon as the analysis has been done, and the results published.

Some of the data management activities have been reported in international journals and conferences; several presentations were held at the last ICES Annual Science Conference in Maastricht. Also, the presentations made during the Ocean Biodiversity Informatics conference held in Hamburg are now published as a special issue of MEPS, and are available through the MarBEF website.

We hope to be able to continue with this work. As a data management team, we have the feeling that more and more people are becoming aware of the need for proper data management, and of the advantages of data sharing. We look forward to continued collaboration.

DATA INTEGRATION

WP-RMP code and title

1-1 Bringing biogeographical data online (BBDO) (www.marbef.org/projects/bbdo/index.php)

Principal Investigator

Edward Vanden Berghe