

Data management for the CCCC Program



Robin M. Brown
Ocean Science and Productivity Division
Institute of Ocean Sciences
P.O. Box 6000, Sidney, B.C.,
CANADA V8L 4B2
E-mail: brownro@pac.dfo-mpo.gc.ca

Robin Brown is Oceanographic Data Manager at the Institute of Ocean Sciences (Department of Fisheries and Oceans) in Sidney, Canada. He graduated from the University of British Columbia in 1976 with B.Sc. in Marine Biology and has participated in many physical, chemical and biological programs in the North Pacific and Arctic Oceans, including the application of remote sensing techniques to oceanography. Since 1995 Robin is the Chairman of the PICES Technical Committee on Data Exchange (TCODE).

The GLOBEC approach is to use a decentralised data management and distribution system. This means that most of the work will be done at the national or regional level, with a small co-ordination role for the International Program Office. Within PICES, the Technical Committee on Data Exchange (TCODE) and the CCCC-Implementation Panel have reviewed the options for CCCC data management and have recommended that the primary responsibility for data management should rest with the national programs, as opposed to some centralised data management office set up in the PICES Secretariat.

Having decided that national programs are to carry most of the responsibilities for data management, it is important to define (in detail) what these responsibilities are! These responsibilities fall into three main areas:

- **Inform interested scientists on the projects that are being carried out as part of national GLOBEC programs.**

This obligation is being carried out quite well by the existing series of CCCC reports and National GLOBEC web sites (Canada, Japan and U.S.A.).

- **Create and maintain inventories of data acquired under GLOBEC (and GLOBEC-like) programs.**

The draft GLOBEC Implementation Plan suggests that “each GLOBEC programme/project should create an inventory of data and data products” and that “These inventories will be collated by the International Project Office to build the centralised database of meta-data and pointers to actual data locations”. This is an area that will need some attention by the national programs - this information is presently missing from all the national programme office web sites. In addition, it is not clear exactly what metadata is required, but a good model can

be found in the Global Change Master Directory (GCMD), which is a large, searchable metadata database that is accessible on the web (<http://gcmd.nasa.gov>). Users can search this inventory using traditional search criteria (latitude and longitude range, date range and parameter) or by keyword/text searches for project or investigator information. For example, it is very simple to search for all GLOBEC data in this inventory (only two entries when I tried!). The search returns a description of each dataset, including area and period of sampling, parameters measured and location of the data.

- **Ensure that data are migrated to a permanent and secure archival centre.**

Permanent, secure archival of data is an important aspect of large international programs such as GLOBEC. The International Geosphere-Biosphere Program (IGBP), which is a sponsor of the GLOBEC program, requires that permanent archival of these data be addressed in the Implementation Plan. This will be a challenge for GLOBEC, as there are few existing data centres that are equipped to handle the diversity of data collected as part of GLOBEC and there is limited experience in the GLOBEC research community with long-term data archival. There may be an important role for the various National Oceanographic Data Centres to provide a permanent home for these data, but handling the GLOBEC data will be a serious challenge for most of these centres as their expertise and experience is focussed on physical and chemical oceanographic data.

The draft International GLOBEC Implementation Plan further suggests that a Data Management Working Group (DMWG) be formed to “review the existing GLOBEC data management systems, issues and problems” and to

“recommend specific strategies for dealing with the shortcomings of existing systems”. The DMWG and IPO staff would also *“monitor migration of data to permanent archives to ensure long term security of the data”*.

It is instructive to consider how other large international programs are handling their data management problems. The Joint Global Ocean Fluxes (JGOFS) program is another IGBP-sponsored marine science program that is more ‘mature’ than GLOBEC. This program is entering the ‘Synthesis and Modelling’ phase and the Scientific Steering Committee has been re-structured with synthesis and modelling in mind. The first objective in the revised JGOFS goals is *“Ensure that all JGOFS observations are lodged with organisations which can guarantee long-term stewardship. Provide web-based information on the availability and access mechanisms to all JGOFS data. Encourage the development of web-based data delivery systems.”*

JGOFS recently (September, 1998) sponsored a Data Management and Synthesis Workshop, which was attended by national JGOFS data managers and researchers who are undertaking aspects of JGOFS data management and synthesis. The present status of data management and the requirements for data synthesis were discussed. The results were as follows:

- Preservation and permanent archival are handled by both migration of data to National Oceanographic Data

Centres (and other data centres, such as the Carbon Dioxide Information Analysis Center) and publication of data compilations on CD-ROM;

- There is no adequate searchable metadata inventory of JGOFS data and this will be a barrier to progress in the synthesis and modelling phase. The JGOFS Data Management Task Team had accepted the responsibility for creating such an inventory, but was unsuccessful in actually completing this task. The present strategy is to explore the possibility of using the Global Change Master Directory system for storing the JGOFS metadata. One large advantage of using the GCMD system is that the “search engine” component is already in place, so it will only be necessary to create the metadata entries (not a trivial task!).

What is the message for the CCCC Program and data management? Staffing for the GLOBEC Program Office is taking place now and there will be a position established for a GLOBEC Data Manager/Co-ordinator. There will be calls for the formation of a GLOBEC Data Management Working Group (as outlined in the Implementation Plan) and the PICES-CCCC Program will be asked to participate in this Working Group. National programmes will be asked to create inventories of GLOBEC data and data products to allow construction of a master inventory of GLOBEC data. In addition, the national programmes will be asked to demonstrate their plans for the permanent archival of GLOBEC data.

(cont. from page 16)

Team will also be developing recommendations about biophysical moorings and zooplankton production.

REX Task Team

The **RE**gional **EX**periments Task Team, which identifies and carries out cooperative research experiments among the PICES regions, made significant progress this year. They published the report of last year’s developmental workshop in PICES Scientific Report No. 9, which forms the basis for their long-term work plan. A highly successful topic session was held jointly with the PICES Fishery Science Committee during PICES VII, highlighting the research findings of GLOBEC and GLOBEC-like programs. This topic session will be continued in future years to ensure a place for GLOBEC researchers to present their findings at PICES scientific meetings. This is an important aspect of scientific networking that provides rapid transfer of information to the GLOBEC research community.

The Task Team held an interesting and successful workshop on climate effects on small pelagic species just prior to

PICES VII in Fairbanks. Nine scientific papers were presented and reviews of existing research programs on small pelagics in each of the PICES nations were provided. Key questions and hypotheses relating to small pelagic species were discussed and several research recommendations for the future were made. A full workshop report will be published in 1999 in the PICES Scientific Report Series. The Task Team plans to hold a workshop in 1999 just prior to PICES VIII in Vladivostok, Russia, on the comparative dynamics of herring and euphausiids. They also plan to compile a summary of the sampling strategies and methods used to assess the stocks of small pelagic species.

*Patricia A. Livingston
Alaska Fisheries Science Center
National Marine Fisheries Service, NOAA
7600 Sand Point Way NE,
Seattle, WA 98115, U.S.A.
E-mail: Pat.Livingston@noaa.gov*