Limited Distribution **IOC/IODE-XVII/28**
Paris, 18 December 2002
Original: English

INTERGOVERNMENTAL OCEANOGRAPHIC COMMISSION
(of UNESCO)

Seventeenth Session of the IOC Committee on International Oceanographic Data and Information Exchange (IODE), Paris, France, 3-7 March 2003

IODE Cooperation with GOOS: Capacity Building and Data Management Aspects

(Greg Reed and Murray Brown)

1. Introduction

The IODE marine data management capacity building program has, up till now, emphasized the use of existing, large-scale data archives as principal sources for constructing national data collections. Students have also been taught methods for adding archived and near-real time satellite data to these archival source materials. Remote sensing data manipulations has, in fact, been rather easy due to the small family of formats used in the satellite community, and to the availability of robust software for data extraction and management. Operational (or real-time) oceanographic data have not, however, been as easy to include in the program, due to the large number of formats involved, and the sporadic nature of the available data transmissions.

2. Assimilating Operational Data

The recent EuroGOOS Conference, held in Athens, highlighted the various projects collecting operational data around Europe. Despite the large amounts of data collected it is often difficult to assimilate these datasets with existing archived data held within the IODE data centres.

As a follow-up to EuroGOOS, the IODE Steering Group for OceanTeacher (SG-OT) was tasked with examining ways to assimilate and synthesise operational data streams within the IODE capacity building programme.

It was decided to focus on a single operational data stream to assimilate with the traditional archived data steams used in IODE/OT. The Argo programme, with a global array of temperature/salinity profiling floats of over 600 active floats collecting T/S profiles that aims to deploy 3000 floats, was selected as the operational data stream. Argo data is transmitted by satellite to data assembly centres in France (Coriolis) and the USA (GODAE). Full resolution profile data, together with quality control flags are encoded into netCDF, a format that is also used by the WOCE programme. Although this de-facto "standardization" on a single format would seem to be a very promising development, in fact the NetCDF standard allows for considerable latitude in individual implementations. There are currently about 10 different versions of NetCDF in use, with both the Coriolis and GODAE centres using different choices. Further complicating the matter is the existence of a third version, the EPIC-COARDS NetCDF standard, which seems to be the best documented and most widely used by WOCE and the DODS live access servers.

The Global Argo Data Repository, GADR, managed by the US NODC has been tasked with providing a permanent archive of all Argo data. They were immediately concerned with the NetCDF variants problem, and have surveyed the available specifications - particularly the EPIC-COORDS standard - with a view toward a "synthesis" standard that can be widely used by software developers and users. The result, informally called the NODC NetCDF, is already being used for daily archiving of Coriolis- and GODAE-derived profile data.

The IODE capacity building program has identified a core group of programs that constitute a toolbox of essential marine software. These programs, covering a wide range of necessary functions, are the subject of lessons and tutorials in IODE training workshops. Three software tools are particularly relevant to the ability to assimilate and synthesise Argo data:

a. Ocean Data View (ODV). ODV is a powerful analysis and display software for ocean station data written by R. Schlitzer of the Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany. ODV has been selected by IODE/OT due to its considerable flexibility and analytical power, and for its ability to re-assign data quality flags "on-the-fly" during quality control work. The ODV data format allows dense storage and direct data access. Extensive data collections containing thousands of stations can easily be maintained and explored on inexpensive desktop and notebook computers. Data from the World Ocean Circulation Experiment (WOCE) and the NODC World Ocean Database 2001 can be directly incorporated into the ODV system.

b. Java Ocean Atlas (JOA). A powerful analysis and display program for ocean station data. JOA was funded by the National Science Foundation and developed in conjunction with the WOCE Hydrographic Office at the Scripps Institution of Oceanography.

c. NCBrowse. Browser/viewer that provides flexible, interactive graphical displays of data and attributes from a wide range of netCDF data file conventions.

The basic problem faced by the IODE training programme was to find a way to bring new operational ocean profile data streams into ODV-based national data collections have been developed in over 20 Member States. Until recently the ODV software could accommodate only NetCDF gridded data and did not have import filters for operational profile data. A pilot experiment by the author of ODV resulted in compatibility with the Coriolis variant of NetCDF, but the proliferation of other NetCDFs represented a serious problem.

To begin a process that might address these problems, the IODE SG/OT has recently been in extensive contact with the authors of the JOA software, the ODV software and the US NODC. A somewhat straightforward solution has been developed, as described below.

The most flexible medium of exchange between JOA and ODV is the Java Ocean Atlas spreadsheet format (JOS), which can be imported/exported by JOA and imported by ODV. The JOS files can also be imported into typical spreadsheet programs (e.g. Excel), so they are extremely easy to use. The US NODC has decided that, in addition to the NetCDF files they post daily, they will also post JOS versions of the same file. Early experiments with these JOS files indicated that certain additions to the format would better accommodate the new data streams (e.g. separate TIME field, quality flags for individual measurements), and these upgrades have been supported by the authors of the JOA and ODV software. Currently, daily Argo data can be retrieved from a new website, imported into either JOA or ODV, and immediately added to national data collections. These relationships and compatibility channels are illustrated in the accompanying Figure 1.

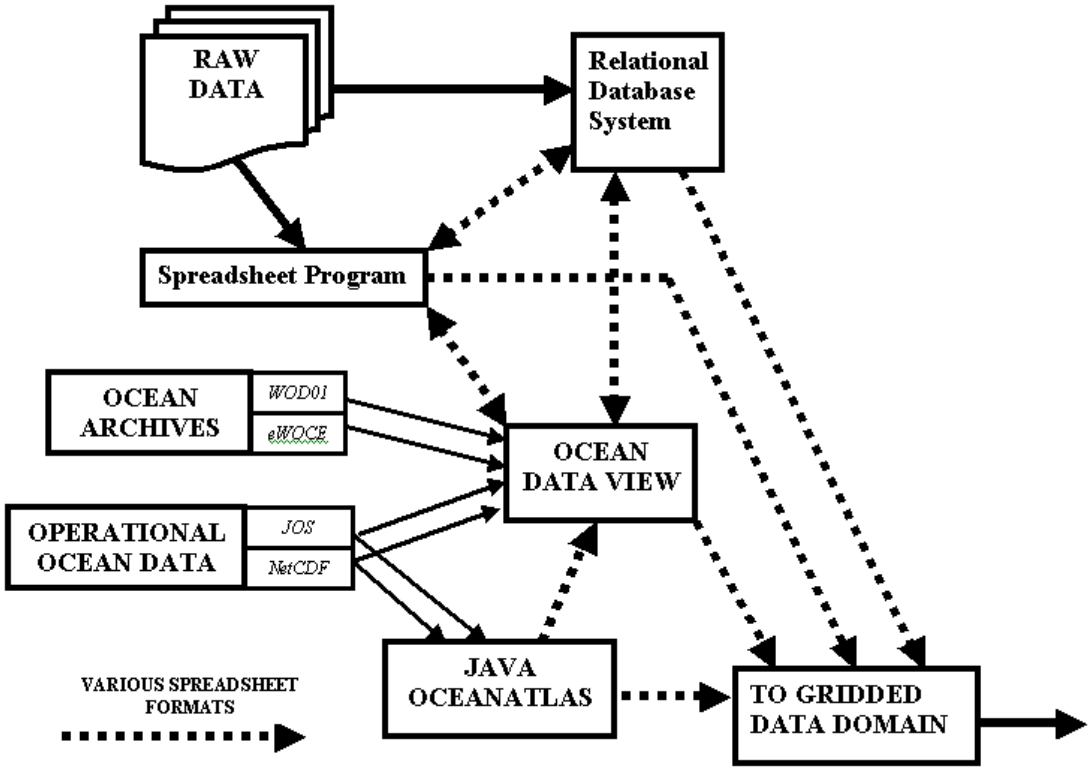

Figure 1. Assimilation of Argo Data into IODE National Data Collections

3. Summary

Recent developments, involving an informal community of software developers, data archivists and the IODE SG/OT, have led to a significant breakthrough in access to operational oceanographic data for training, national data collecting, and analytical synthesis work. However, due to their very "operationality" the data have received only basic quality-control, and new operational data cannot be simply added to existing collections without ensuring they meet the specified QC criteria. Instruction on data quality control concepts and procedures is already covered during training workshops and the tools are available to visually inspect new data.

Work is underway to add supplementary tutorials to the IODE OceanTeacher to demonstrate the exact methods involved and, from 2003 these procedures will be included in the IODE capacity building program. This will provide a relatively easy procedure to assimilate operational data into traditional IODE national data collections.