New MapServer interface to US GLOBEC data sets

Robert C. Groman, M. Dicky Allison and Peter H. Wiebe
Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA

A new, geospatial interface to US GLOBEC data sets has been implemented and provides map-based access to the data collected during the US GLOBEC programme. Data from Georges Bank, the Northeast Pacific and the Southern Ocean can now be viewed using MapServer software, modified to access data directly from the JGOFS/GLOBEC data and information management system. Maps and graphs can also be produced and displayed using any standard web browser such as Internet Explorer, Netscape, Firefox and Safari. Additional interoperability features have been added so that data can be downloaded using the Open Geospatial Consortium's Web Mapping Service and Web Feature Service protocols. Data downloads as ASCII files (space, comma or tab separated) and Matlab binary files are also available. Access to the new map-based data interface is via the following URL: http://globec.whoi.edu/map.

Introduction

The JGOFS/GLOBEC data management system has been used by the US GLOBEC programme to manage and serve its scientific investigators' data since the programme's inception in 1994. The system is uncomplicated, flexible and integrates extremely well with the Internet via standard web browsers. It uses a simple directory list of file subdirectories to organise and display available data sets (also called data objects - Figs. 1 and 2). However, times have changed and the need for a map-based interface became clear. The US JGOFS programme used a Live Access Server interface to display its synthesized data sets at the end of its project. However, we wanted a more open-ended approach, one that could be as "data driven" as possible. We also wanted a solution that was not proprietary, but rather open source, and based on recognised standards, so that we could share our efforts with others at little or no cost. Our needs led us to select the MapServer software developed at and available from the University of Minnesota. MapServer software supports a large number of raster and vector plotting formats including Tiff, GeoTiff, EPPL7, ESRI shapefiles, PostGIS, ESRI ArcSDE and Oracle Spatial, but it was its support of the Open Geospatial Consortium (OGC) standard and open web specifications, in particular the Web Mapping Service (WMS) and the Web Feature Service (WFS) protocols that encouraged us to select this software.

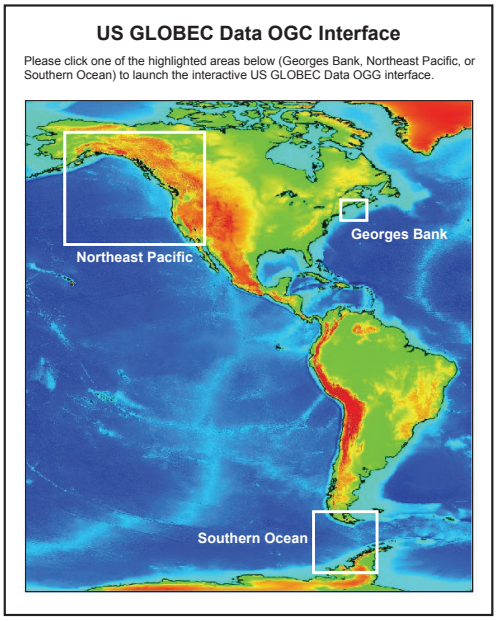

Figure 1. A preliminary version of the map interface to the US GLOBEC data sets. One clicks on the area of interest and is presented with the US GLOBEC cruise tracks available in that area.

Point and click to display data

Once the general area is selected by clicking in the geographic region, an interactive plot is displayed on your browser page showing all the cruise tracks undertaken during the programme (Fig. 2). You can select one or more cruises by highlighting the cruise ID, or select a bounding box around your area of interest using your computer's mouse.

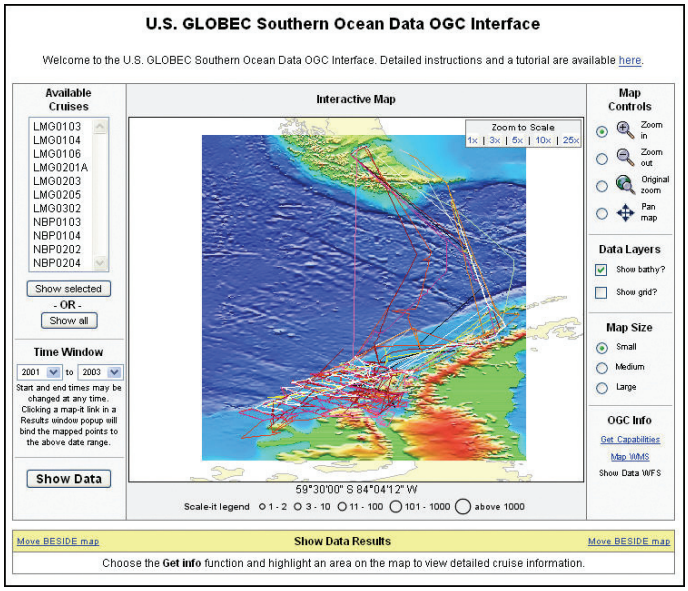

Figure 2. A display of the cruise tracks from the US GLOBEC Southern Ocean project.

When a sub-set of the cruises is selected the resulting image is replaced by the cruise tracks of the selected cruises (Fig.3). At this point, clicking on the "Show Data" button will list all the available data for these selected cruises to be shown at the bottom of the web page (Fig. 4). The types of plots available are determined by the types of data available in each data set. For example, clicking on the CTD data set will show all the stations where these data are available (Fig. 5). Moving the mouse to one of the stations will display another "plotting window", which allows you to create an X-Y plot of water depth versus a scientific parameter, such as salinity or temperature (Figs. 6 and 7).

It should be noted that except for the first map display, all maps and images are created "on the fly", that is, in real-time. While this can result in some delay in displaying data, especially if a large data set is accessed, we felt it best to retain the "data driven" nature of the JGOFS/GLOBEC data management system.

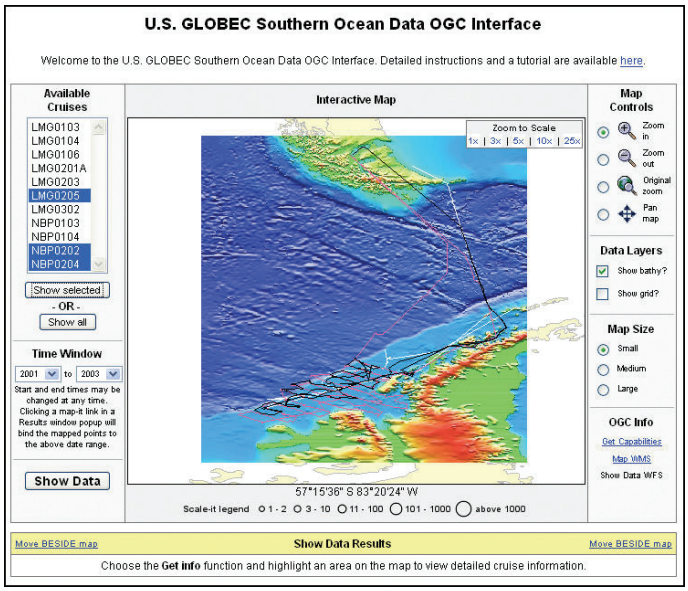

Figure 3. Display of the selected cruises.

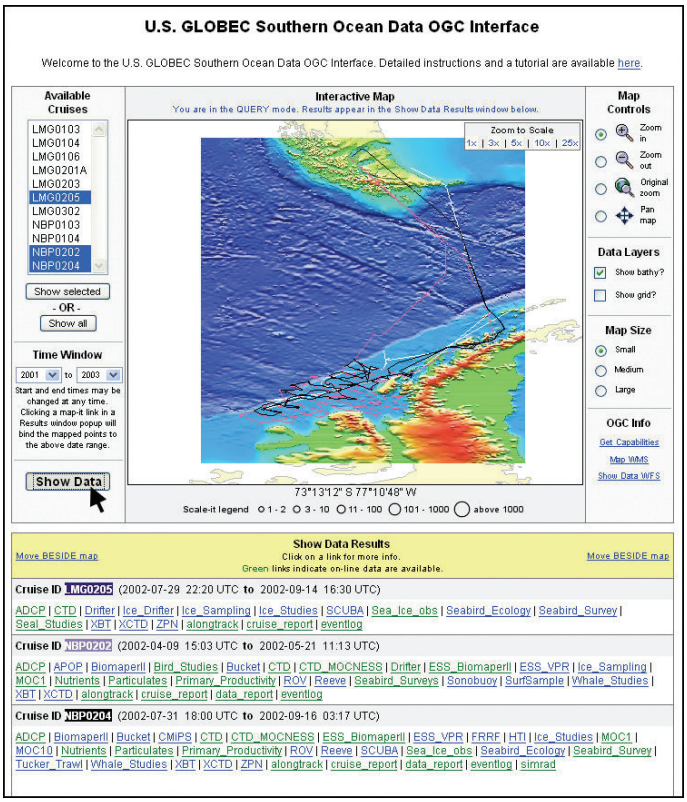

Figure 4. The list of all the data sets available for the selected cruises.

Interoperability

During the development of the map-based interface we also wanted to improve access to the underlying data. The existing JGOFS/GLOBEC system provides a download capability that supports ASCII text-based, tabular downloads as either comma separated or tab separated data files. It also supports downloading the data as a Matlab binary file, organising each of the data values as Matlab data vectors. By selecting the MapServer software we also enabled downloads using the Web Mapping Service and Web Feature Service protocols developed by the Open Geospatial Consortium.

Implementation difficulties and current limitations

Currently, the map-based interface has several limitations. One major limitation is that geospatial display of data cannot be accomplished unless latitude and longitude values are available. While this limitation is "obvious to the casual observer" it took some data contributors by surprise since they thought that specifying a station number as the "location" would be sufficient. While a mapping between station number and a latitude and longitude ordered pair is clearly possible, the MapServer software prefers to get its position information in a direct simple way.

Another limitation is that data retrieval by date is limited to specifying only the start and end years. This restriction will be removed as we unify the format of the data and time information. Here we are paying the price of allowing data contributors to provide date and time information in whatever form they wish. While coding for all the different date and time options is possible, it adds significantly to the implementation costs. Since the JGOFS/GLOBEC data management system already has the ability to convert date and time into several different formats, including year-day and number of seconds since "the epoch", we decided to take advantage of the existing software by modifying how the existing data are served and, without changing the underlying data files, add the necessary date and time fields automatically.

A significant limitation that has slowed our map-based implementation is that metadata information about each of the data sets served are currently stored and served as text files (actually as html files). Initially, this approach was very flexible and fostered easy data sharing among the scientific investigators. However, in order for computer software to decide what types of plots are possible, it is necessary for the metadata to be encoded into a machine readable form. Currently, we have implemented simple tables describing our served data sets, but we will replace this approach with a more robust method that accesses information about each data set from a separately maintained database. This effort is currently underway.

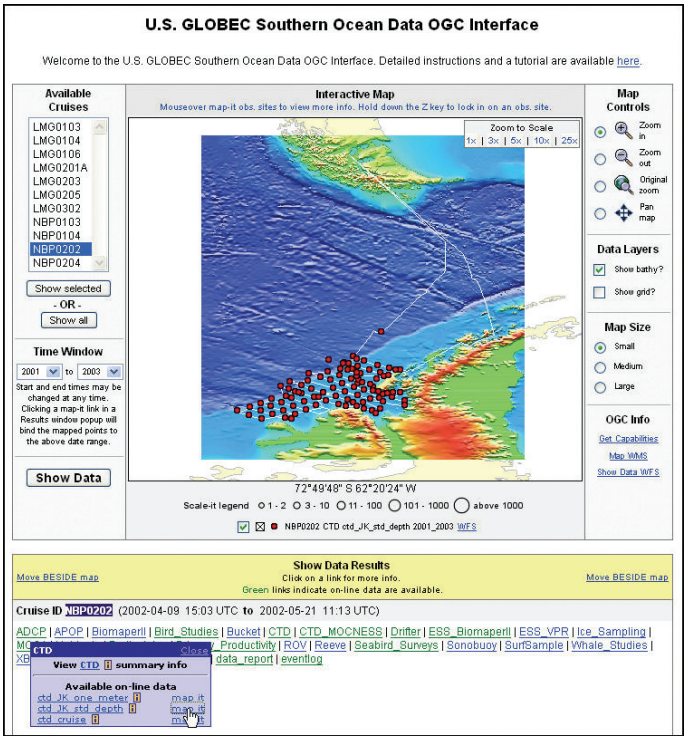

Figure 5. A selected cruise track and the location of stations where data are available.

Figure 6. The data selection window.

Conclusion

A map-based interface to the US GLOBEC data sets is available on-line using any standard web browser at http://globec.whoi.edu/map. The interface is still under development, but initial improvements in data discovery, display and accessibility have been achieved. We encourage feedback from people using the new interface so that we can continue to improve its functionality and ease of use.

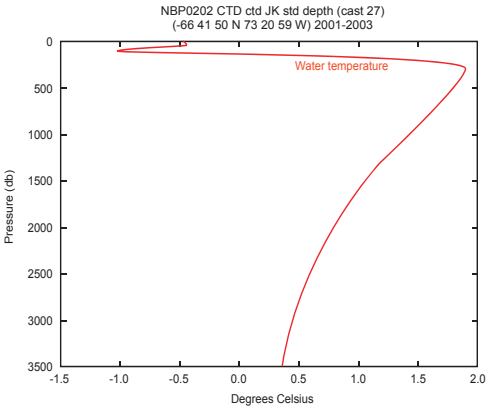

Figure 7. An X-Y plot of the data from the selected station's parameters.

Acknowledgments

We would like to thank Charlton Purvis for implementing the MapServer interface to the JGOFS/GLOBEC data sets. We also would like to thank all the (unnamed) people who contributed their ideas and suggestions for improving the map-based interface.