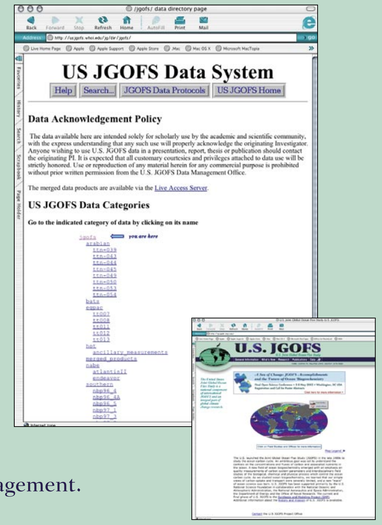
U.S. JGOFS Data Management

Author: Cyndy Chandler (U.S. JGOFS DMO, Woods Hole Oceanographic Institution)

The Joint Global Ocean Flux Study (JGOFS), organized under the aegis of the Scientific Committee on Oceanic Research (SCOR), was a core project of the International Geosphere-Biosphere Programme (IGBP). U.S. JGOFS, a national component of JGOFS, included field studies conducted between 1989 and 1998 and brought together biological, chemical, physical and geological oceanographers and modelers in a multidisciplinary investigation of the pools and fluxes of carbon and associated biogenic elements in the ocean. The U.S. JGOFS research program comprised four basin-scale process studies, two long-term time-series programs, participation in a global survey of carbon dioxide (CO_2), a synthesis and modeling project and data management.



U.S. JGOFS Data Management Office

The United States JGOFS Data Management Office (U.S. JGOFS DMO) was formed at Woods Hole Oceanographic Institution specifically to meet the needs of U.S. JGOFS and was funded by the United States National Science Foundation (NSF). A prominent aspect of U.S. JGOFS was the sharing of information between program participants as well as with the larger scientific community in the U.S. and abroad. Data and documentation were stored and organized in the JGOFS database system and the U.S. JGOFS data server utilized the World Wide Web (WWW) to allow investigators to easily exchange information. The U.S. JGOFS WWW interface was developed to support dynamic subselection of data, access to metadata documents, customized merging of datasets and extraction of user-specified data for subsequent incorporation into data analysis applications (i.e. MATLAB) or data assimilation models.

Data Management Office Personnel:

David M. Glover
Director, U.S. JGOFS DMO

Cyndy Chandler
System administrator and database manager

Kathryn Elder
Data analyst: 1994-95

Christine Hammond
Manager, 1994-2001

George Heimerdinger
Data analyst: NABE and EqPac

Terri MacEachern
Data report: 2002-03

Dave Schneider
Data analyst: Arabian Sea and AESOPS

Lt. Raymond Slagle
Data analyst:1989-1991

WOODS HOLE OCEANOGRAPHIC INSTITUTION 1930

NSF

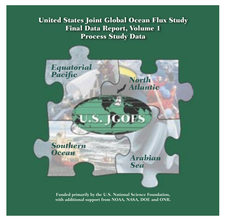
U.S. JGOFS Merged Data Product Synthesis

Selected data submitted by process study cruise participants have been combined into a collection of comprehensive data objects. Initially, the main tasks of the U.S. JGOFS Data Management Office (DMO) centered primarily around the receipt and quality control of field data. To facilitate the U.S. JGOFS Synthesis and Modeling Project, the DMO staff began generating 'merged data products'. The merging process is an attempt to 'put the ocean back together again' after all the samples have been analyzed by U.S. JGOFS investigators.

What is a merged data product?

A merged product is a large data set; a combination of data records that have related variables and a common collection method. The DMO has synthesized CTD profile, Niskin, trace-metal-clean and Go-Flo bottle merged data products for the four process studies. Merged data products are created by joining all U.S. JGOFS data sets in which analyses from the same sampling device have been reported for a single cruise. If sampling location and time are the same, data are placed in the same row. If variable name (reflecting units and methodology) is the same, data are placed in the same column. As an example, the merged Arabian Sea Niskin bottle data product comprises more than 110 parameters and was synthesized from 79 data objects from the six Arabian Sea process cruises listed in the table.

	R/V Thomas Thompson Cruise Number					
Object Name	043	045	049	050	053	054
bottle	•	•	•	•	•	•
bacteria	•	•	•		•	•
total CO2	•	•	•		•	•
total organic C	•	•	•	•	•	•
HPLC pigments	•	•	•	•	•	•
particulate organic C and N	•	•	•		•	•
total organic N	•	•	•	•	•	•
particulate matter conc.	•	•	•	•	•	
alcoplankton	•	•	•	•	•	•
nanoplankton	•	•		•		
microplankton	•	•		•		•
iodine speciation		•				
chl fluor		•		•	•	
trace Fe Al		•	•	•	•	
sulfide peroxide		•				
manganese		•				
calcite production		•	•		•	
low level NO3		•			•	
mixed layer	•	•	•	•	•	•
event log	•	•	•	•	•	•

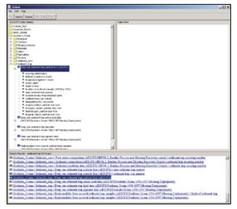


Final Data Report: Volume 1, Process Study Data

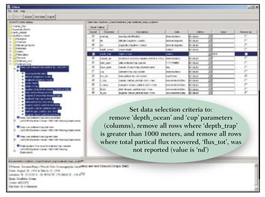
The web-based data server has provided fast, convenient access to U.S. JGOFS data and information. As part of the U.S. JGOFS legacy, the DMO has published volume 1 of the final U.S. CD-ROM data report which contains data acquired during the four U.S. JGOFS process studies, conducted in the North Atlantic, equatorial Pacific, Arabian Sea and Southern Ocean.

Volume 1 is a collection of data files and supporting documentation describing sampling methodology, and documents which provide background information and explanations of overall program design. Data report access is achieved either by web browser client or JGView, the custom-designed Java interface developed at Woods Hole Oceanographic Institution by Cyndy Chandler and Adam Shepherd.

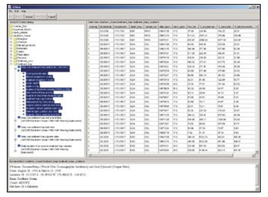
JGView: search for data of interest [Search]



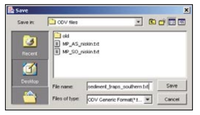
JGView: set selection criteria [Subset]



JGView: list/view selected data [View Data]

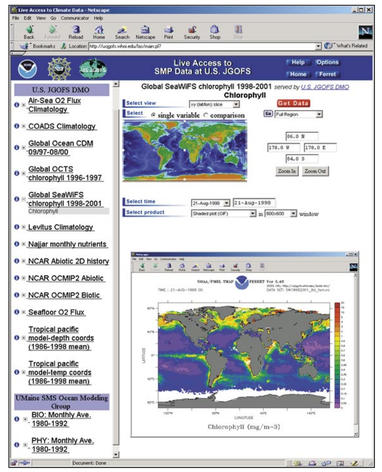


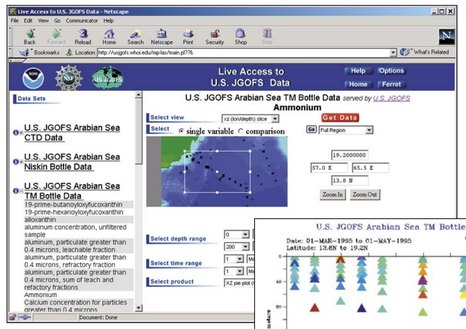
JGView: export data to a file in ODV Generic format [Export]

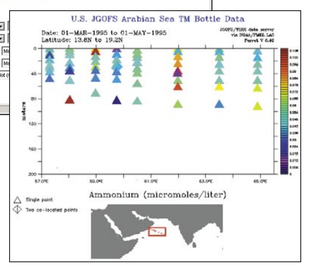


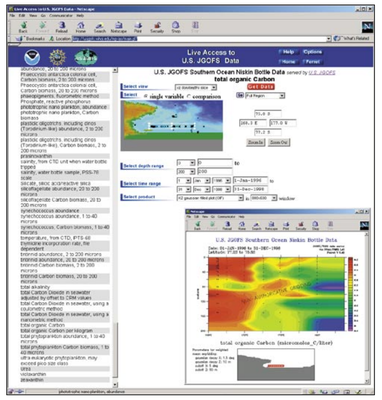
Live Access Server Data System Interface

The web interface to the U.S. JGOFS Data Server was sufficient during the field study portion of the program. During the Synthesis and Modeling Project, a new access and visualization interface was developed through collaboration with people at the University of Washington and National Oceanic and Atmospheric Administration (NOAA) Pacific Marine Environmental Laboratory (PMEL) in Seattle, Washington. The Live Access Server (LAS) interface, customized for U.S. JGOFS in-situ data sets, provides a means of navigating, visualizing, selecting, reformatting and exporting subsets of multidimensional data sets. LAS is a web interface to Ferret, the interactive computer visualization and analysis environment developed by the Thermal Modeling and Analysis Project (TMAP) at PMEL.









DMO Timeline

Information systems technology and 'latest system configuration' descriptions
US JGOFS events
US JGOFS DMO activities

1981		"640K ought to be enough for anybody." (Bill Gates)
1982		Internet formed from existing TCP/IP networks
1984	September	US NAS GOFS workshop
		DNS has 1000 registered hosts
1985		MS-DOS 3.0, 4.7 MHz Intel 386 processor Dual 5 1/4-inch floppy disk drives
1987		JGOFS launched
		First Sun SPARC system (10 MIPS)
1988		HOT and BATS initiated
		Data manager: George Heimerdinger (NODC)
		Multitasking meant 3 computers and a wheeled chair.
		Intel 80286 with a 5 inch floppy drive and time on a Sigma-7 with 9 track tape; data tranmitted via US mail.
1989	April	NABE Atlantis 119 cruise
	June	NABE Endeavor 198 cruise
		Data interface: seven function, X-based GUI for UNIX clients with TCP/IP
		World Wide Web invented at CERN, EPPL
		Intel 386 processor, 16 MHz Intel 80386, 1 MB RAM, 9 track tape drive
1990		CO_2 Survey initiated
		Windows 3.0 released
1991	February	NABE data report published
1992	February	Equatorial Pacific TT007 cruise
	November	Equatorial Pacific TT013 cruise
		16 MB Process Study data being served
1994	September	Arabian Sea TT039 cruise
		Netscape 1.0 released
		DMO created at WHOI (Manager: Christine Hammond) 104 MB Process Study data being served Data server: Sun SPARC 5, 85 MHz 1 GB disk space, CD-ROM drive
1995	December	Arabian Sea TT054 cruise
		OS/2 system, 90 MHz Pentium 64MB RAM, 880 KB floppy, 1MB HDD
1996	August	SMP Workshop (Durham, NH)
	September	AESOPS NBP 96-4 cruise
		Windows 95 system, 133 MHz 64 MB RAM, 880 KB floppy, 170 MB HDD
		Netscape Navigator 2.0 released (JavaScript support)
1997		CO_2 Survey completed
		176 MB Process Study data being served
1998	March	AESOPS NBP 98-2 cruise
	May	Data server: SGI Origin 200, dual CPU, 128 MB RAM, three 9 GB disks
1999		JGOFS Arabian Sea CTD CD-ROM published
		PowerMac G4, 400 MHz processor, 64 MB RAM, CD-ROM, 10 GB HDD
2000	January	US Timekeeper (USNO) reports year as 19100
	September	DMO Director: David M. Glover DMO Manager: Cyndy Chandler
2001	December	514 MB Process Study data being served 80% NABE acquired, 90% EqPac 95% Arabian Sea, 80% AESOPS
2002	July	1.6 million DNS registered hosts
	December	602 MB US JGOFS data being served (final data set submitted to DMO) 100% NABE acquired, 99% EqPac 98% Arabian Sea, 96% AESOPS 3.8 GB Process Study and SMP data being served
2003	April	JGOFS DMTT publishes final Data Collection
		US JGOFS CD-ROM Data Report published
	May	JGOFS Open Science Conference (Washington, D.C.)
		Windows XP, 2.4GHz Intel Xeon processor, 1 GB RAM, DVD-RW, 40 GB HDD
	July	Final SMP Workshop (Woods Hole, MA)

"Computers have enabled people to make more mistakes faster than any other invention in human history, with the possible exception of handguns and tequila." (paraphrased from Mitch Radcliffe)

Computer timeline info taken from: White,Stephen. (2001) "A Brief History of Computing" [http://www.ox.compsoc.net/~swhite/history/timeline.html]