# NOAA Support for the CLIVAR and Carbon Hydrographic Data Office at UCSD/SIO

T. Bruce Appelgate
UCSD Scripps Institution of Oceanography, La Jolla, CA

## Table of Contents



## 1. Project Summary

The CLIVAR and Carbon Hydrographic Data Office (CCHDO), brings together, verifies, and corrects content and format errors in hydrographic and tracer data used in large scale ocean carbon, global change, water mass, and circulation studies [CLIVAR = Climate Variability and Predictability, a component of the World Climate Research Programme]. It assembles the data with relevant documentation, and carefully prepares them for dissemination and archive. In addition they work to promote appropriate methodology, applicable community standards, communications, and data compatibility. The CCHDO brings data together to a common content and readability standard, thereby greatly reducing the difficulties research and education data users encounter. Documentation associated with the data are collected, reorganized to a common standard (where possible), and preserved with the data. The CCHDO makes it possible for all data users to cope with the temporal-, content-, and format-related file diversity these different originators engender. The CCHDO supports CLIVAR and ocean carbon science programs, and is a critical component of a global observing system for the physical climate/CO2 system.

The largest share of CCHDO support is provided by NSF. CCHDO activities of special interest to NOAA, and for which NOAA support is applied, include:

1. Provide data from specific cruises that are of special interest to NOAA.
2. Continue working to improve relationship with NOAA's National Oceanographic Data Center (NODC) via:
   a. Continued data assembly of cruise hydrographic data and metadata, particularly those from the GO-SHIP program.
   b. Working with NODC to improve efficiency of transfer of data and to make CCHDO data more "archive ready".
   c. Improved CCHDO participation in discussions towards enhancing integration of specific related Data Assembly Centers to reduce ambiguity and redundancy in data archiving.
3. Broaden and simplify the accessibility of CCHDO data sets.

4. Support CCHDO Data Manager travel to data meetings of special interest to NOAA.

## 2. Scientific and Observing System Accomplishments

A broad international community makes use of CCHDO data, with use most intense from US researchers and students. For example, during the period July 1, 2011 - July 1, 2012 these are the visit statistics for the CCHDO web site **http://cchdo.ucsd.edu** (counting is done by IP address):

10,548 visits,
80 countries with visitors who stayed on the site longer than one minute,
4402 unique visitors,
1277 unique returning visitors, and
3125 unique new visitors.

[To eliminate in-house use, the above does not count visits originating from within UCSD. This does, however, also eliminate counting of use of the CCHDO by SIO researchers and students, which is significant.]

Data submissions and subsequent CCHDO web site data updates continue to be frequent. For example, during the period July 1, 2011 - July 1, 2012 these are data submission statistics:

| | **Files** | **Cruises** |
|---|---|---|
| SUBMITTED Files (new and updated) | 108 | 61 |
| WEBSITE UPDATES | | |
| Bottle | 150* | 40 |
| CTD | 72 | 55 |
| SUM | 18 | 18 |

(*11 of these files are large-volume sample files)

Over that same period the CCHDO added 2174 actual new pages of documentation to the website. (This figure does not reflect a number of reports that went online in 2 or more places because they covered more than one cruise.) These pages represent 60 unique cruises.

During this time period the data histories have been updated 325 times. Most of those updates reflect either data submitted to the CCHDO, or website updates (i.e. data going online in either "final" format or "as received" format [in the queue]).

In terms of data submissions, participation from most US data originators (the PIs who are in charge of one or more parameters) has been excellent, with only two US PIs (not NOAA funded) presently overdue with data deliveries.

Regarding international participation, there are seven nations engaged in full-depth repeats of WHP basin-scale transects which include ocean carbon parameters: Australia, Canada, France, Germany, Japan, the UK, and the US. (The CCHDO also data with less broad parameter content from other nations.) Of these, Australian, Japanese, and US investigators continue to routinely provide their CTD and bottle data to the CCHDO on their nation's schedule, and now data

receipts from Canada from the Bedford Institute of Oceanography have become routine on the schedule they set. Some UK investigators by-pass their national data center to provide CTD/hydrographic data to the CCHDO on the schedule they set (sometimes preliminary shipboard data are sent immediately after a cruise). Through the CDIAC and Dr. Robert Key (Princeton) the CCHDO receives some data from other countries.

With respect to international coordination we note that it would help significantly if there were an active, funded international scientific oversight and planning body such as GO-SHIP, which is a widely known and appreciated effort within the community. We note that what would help the most in terms of data flow and availability would be for international GO-SHIP to gain official status within the WCRP, similar to that enjoyed by Argo. This way, the GO-SHIP participants from all nations would be more or less obligated to provide their data to the data center (the CCHDO for CTD/rosette data), and would be supported by their national funding and oversight bodies to do so.

[UNESCO lost significant funding recently. This reduced manpower in the IOC and the support that the IOC can provide to nascent GO-SHIP efforts. Partly to help deal with this, in August 2012, at the request of the GO-SHIP chair and others, the CCHDO agreed to host the GO-SHIP web site. The CCHDO will streamline and improve the data submission aspect of GO-SHIP, and provide the GO-SHIP steering committee the ability to make timely updates. The IOC will soon forward web files and data, and the CCHDO should have the GO-SHIP site up and running in the UCSD domain (but retaining the *.org name) by late 2012. The CCHDO is also working with the GO-SHIP steering committee to reconcile the CCHDO data holdings with GO-SHIP cruise records. Through this process the CCHDO will then have a detailed list of additional cruise data to chase, obtain, reformat, and add to the CCHDO site.]

The CCHDO continues to enjoy a mutually beneficial relationship with both the Argo and OceanSITES programs. In exchange for the CCHDO providing both programs with specifically formatted CTD data, these NOAA programs in turn assist the CCHDO in the discovery and acquisition of hydrographic data from the PIs involved in those communities. CCHDO data management support of the NSF-funded DIMES project has developed well during the past year, including establishing the format of the HRP data, CF format compliance and overall data access and distribution. The point here is that the tools, formats and methods that NSF has invested in over the years for the CCHDO have proven to be directly applicable to effective management of data for these types of process studies and field programs.

Since the time of WOCE and the WHPO (the CCHDO's predecessor), NOAA and NOAA researchers have benefitted from CCHDO (and WHPO) services, for example providing a full range of data and documentation services to NOAA cruises for the program, providing data to many NOAA scientists who routinely use CCHDO data and documentation, supporting CTD data services for the NOAA-supported Ocean-SITES program, and providing significant "front-end" data services which aid NOAA/NODC. The CCHDO's new, partial fiscal support from NOAA helps to support the above activities and to provide data from specific cruises that are of special interest to NOAA, to work more closely with NOAA on data assembly as related to NOAA data centers, to broaden and simplify the accessibility of CCHDO data sets to NOAA investigators, and to make holdings more query-able and accessible through modern data

management standards and practices (including those that NOAA and the international oceanographic community recommend).

On the internal support side of CCHDO operations, the CCHDO's two Mac Mini servers are now rack-mounted in the SIO server room, which provides smooth power, climate control, fire suppression, and a windowless locked room. CCHDO web servers (known as cchdo, ushydro, goship, dimes, joa, and seahunt) run on virtualized machines on top of the hardware, allowing for software upgrades to hosted sites with minimal down time. CCHDO disk storage moved to a RAID high capacity hard drive with daily backups.

New parameters arise from time to time in the CCHDO data files, for example from biogeochemical ancillary measurement programs on CLIVAR/CO2 cruises. The CCHDO works with data originators to track and keep meta data for new parameters which are measured on cruises in the CCHDO's holdings, and continues to update and maintain on-line information about these parameters.

The CCHDO currently maintains files for more than 1300 cruises. The underpinnings of the CCHDO's in-house (staff only) and on-line (public) operations require continual modifications and updating to maintain compliance with security and operational guidelines, always with a move toward apparent ease and power of use combined with underlying simplicity of operations and maintenance. For example, in the past year, we made moving submitted data to the on-line queue a matter of simply filling out a short form and one click (which also automatically adds a data history note), thus reducing human error and the time between data being submitted and getting on-line.

The CCHDO now has (internal) software modules that read and write data in WOCE, WHP-Exchange, netCDF, OceanSITES, MATLAB, and some "in-house" formats used by major data submittors. This software includes some unit conversions, conversion of depths to pressures, reorder of rows based on pressure, and so forth. In the past year a module has been developed to read some SeaBird CTD files directly. The new and updated modules resulted in large performance increases for reading and writing data formats.

During the next year, work well underway will provide a single design for lists of cruises. Every list of cruises on the site will then look the same. More importantly, a single, modern, more easily maintained structure will underlie all internal and on-line cruise and data information. This will include the ability to call up lists of cruises for objects such as collections, people, institutions, ships, country, etc.; the ability to provide "branding" for cruise data pages that belong to certain collections. The search engine will understand what objects are. User authentication (optional) will provide additional interactive features, such as permitting authorized guests to suggest changes to cruise metadata and to make data and information history notes. (Authentication as staff provides access to staff-only tools.) The CCHDO will add the ability to make bulk archives of CTD files, a long-requested feature that was difficult to implement until the new unified underlying internal structure was introduced during the past year.

### *2.1. Outreach and Education*

The CCHDO maintains a website http://cchdo.ucsd.edu where the data and accomplishments of the project are maintained for public/scientific view and use. CCHDO data form the core data used in the Exercises developed for the new textbook "Descriptive Physical Oceanography - An Introduction" by Talley, Emery, Pickard, and Swift (see http://joa.ucsd.edu/dpo). The CCHDO contributed to and maintains the outreach pages on the USHYDRO web site, including a virtual cruise (ship plans, photos, videos, etc.) and an example of a complete cruise from proposal through preliminary data interpretation (see http://ushydro.ucsd.edu/outreach). Additional community outreach is accomplished by attendance and poster presentations at conferences and meetings.

## 3. Publications and Reports

### *3.1. Publications by Principal Investigators*

The CCHDO is a data assembly and distribution center. The work plan for the NOAA funding for the CCHDO does not involve traditional research activities that lead to peer-reviewed publications from the CCHDO Principal Investigators. Also, the NOAA funding does not support preparation of peer-reviewed publications by the CCHDO PIs using NOAA OCO funds.

### *3.2. Other Relevant Publications*

Many peer-reviewed publications are generated by the users of CCHDO data, by PIs from a wide range of institutions and nations who use data provided by the CCHDO, including data whose assembly and distribution are funded by NOAA support of the CCHDO. There is, however, no viable mechanism to track the publications generated by the thousands of data requests during the past year.