A Guide to Submitting CTD/Hydrographic/Tracer Data and Associated Documentation to the CLIVAR and Carbon Hydrographic Data Office

James H. Swift
UCSD Scripps Institution of Oceanography

updating and replacing WOCE Hydrographic Program
Reference Manual 90-1 (Joyce, T. and C. Corry, 1994)

1. OVERVIEW

A fundamental goal of groups making ocean profile measurements is reporting of their data. This guide is aimed to assist chief scientists, measurement team leaders, technical teams, and data groups when they report data and documentation to the CLIVAR and Carbon Hydrographic Data Office (CCHDO) at the UCSD Scripps Institution of Oceanography[1]. The CCHDO is an internationally-sanctioned Data Assembly Center (DAC) dealing with what oceanographers term "hydrographic data", meaning vertical profile data from CTD, CTD/rosette, and bottle casts covering physical and chemical parameters.

The CCHDO makes all its documentation and public data available via the internet site <http://cchdo.ucsd.edu>. Also on this site is supplementary information, such as data histories for each cruise, contact information for data originators, electronic versions of manuals and guides, maps showing cruise tracks, and listings of other similar data in the same region.

Summary of Recommendations

A. Send a cruise plan to the CCHDO. (Optional.)

Before a CTD/hydrographic cruise takes place, it may be useful for the scientists planning the cruise to contact the CCHDO (via email to cchdo@ucsd.edu). Contact will help to ensure the CCHDO is ready to receive the data after the cruise.

B. Apply and document appropriate sampling and analytic protocols at sea. (Essential.)

The core of reference-quality CTD/hydrographic/tracer work lies in the application and documentation of appropriate methodology. [Analytic methods are not discussed in this data submission guide.]

Follow and document data quality control procedures **(Essential)**:

- samples should be accurately identified in time and space;
- CTD profiles and water samples require unambiguous identification;
- parameters should be co-examined for reasonableness; and use of data quality codes is recommended (Assignment of data quality codes is discussed in **Appendix D**).

[1] In 2003 the WOCE Hydrographic Program Office at the UCSD Scripps Institution of Oceanography also became known as the CLIVAR and Carbon Hydrographic Data Office. In this document "WHPO" refers to the office or its activities prior to 2003 and "CCHDO" refers to the same office in the post-WOCE era.

C. Transmit the data and documentation to the CCHDO. (Essential.)

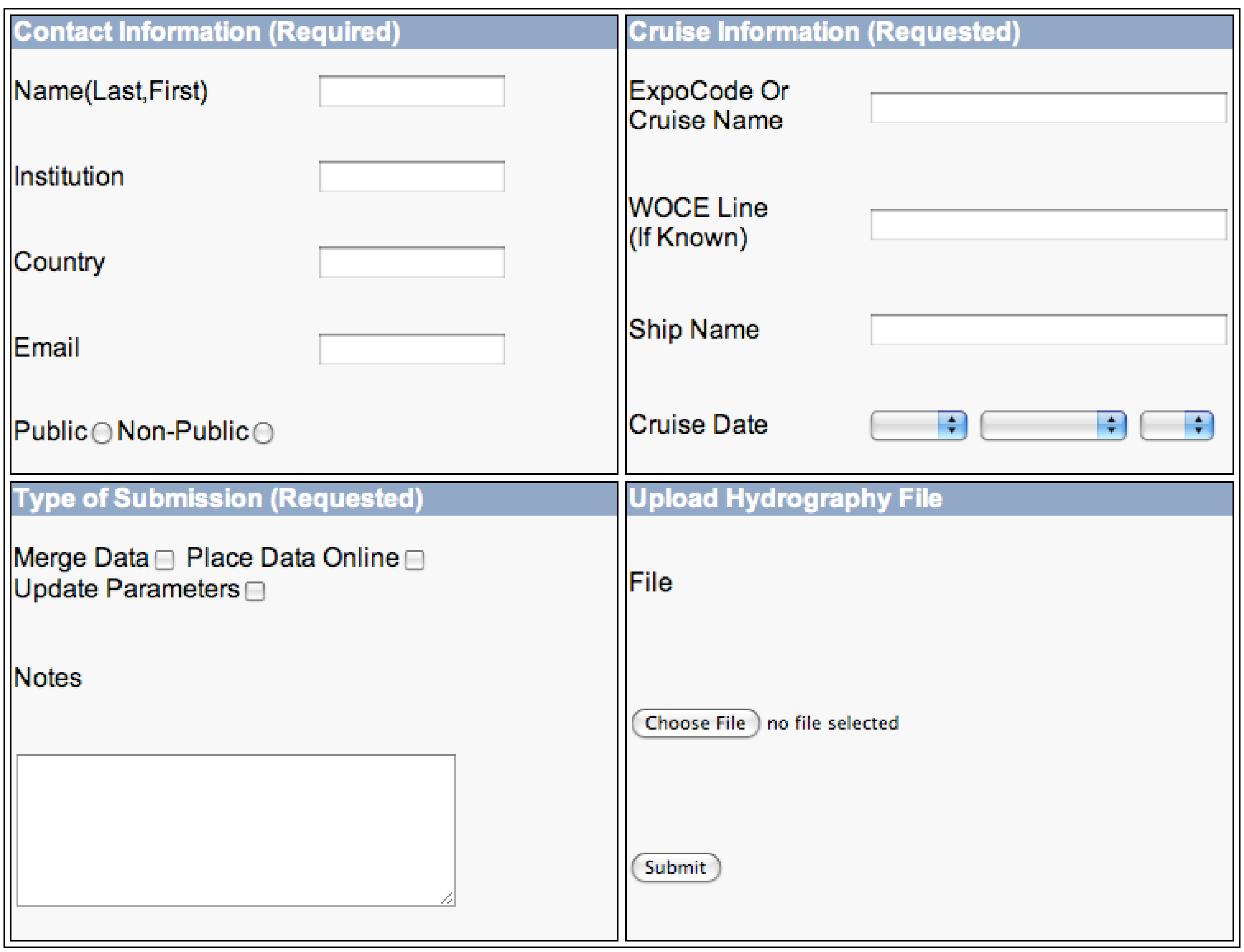
Although the CCHDO prefers that data originators submit their data in WHP-Exchange formats (see **Appendix C**), or in original WOCE formats, the CCHDO will accept data in any unambiguous format.

The minimum ancillary information which **must** be included when data are submitted to the CCHDO:

a. name, institution, country, and email address of person submitting the data;
b. clear indication regarding whether or not the data are public; and
c. parameter names and units need to be clearly defined in data files and documentation.

It is strongly recommended that all reports and data be submitted to the CCHDO by electronic file transfers using the interactive form found at < http://cchdo.ucsd.edu>. The only exceptions should be submissions from investigators who do not have fast internet connections. In those cases data can be mailed to the CCHDO on standard media (e.g., data CDs or DVDs at the time of writing).

Below is a copy of the short interactive form found by clicking on the "Submit Data" link on the CCHDO web site at < http://cchdo.ucsd.edu/submit>.



Contact Information (Required)

Name(Last,First)

Institution

Country

Email

Public Non-Public

Cruise Information (Requested)

ExpoCode Or Cruise Name

WOCE Line (If Known)

Ship Name

Cruise Date

Type of Submission (Requested)

Merge Data Place Data Online

Update Parameters

Notes

Upload Hydrography File

File

Choose File no file selected

Submit

Scientists may submit their data to the CCHDO at any time after a cruise. Bottle data parameters from different measurement groups may be merged by the scientific team or can later be merged by the CCHDO. The CCHDO maintains file histories, carries out file version control, and updates data in an orderly manner. Similarly, the individual documentation contributions from different measurement groups can be combined by the scientific team or by the CCHDO.

D. Share the data with the oceanographic community (Very Strongly Recommended)

Programs such as CLIVAR and the International Ocean Carbon Coordination Project (IOCCP) recommend that data be submitted to the CCHDO within two years after the cruise. This data submission timeline, used during WOCE, helps greatly to see that data are available for model calibration and validation, carbon system studies, heat and freshwater storage and flux studies, deep and shallow water mass and ventilation studies, and for calibration of autonomous sensors. Moreover, it makes the data available for comprehensive data quality examinations combining multiple cruises. It may take longer than two years after a cruise to prepare data from some parameters on Repeat Hydrography cruises. The CCHDO will be responsible to receive these and other data updates, and to produce revised data files reflecting the newly-submitted results.

Some national Repeat Hydrography programs define a shorter timeline for data submission. These should help to inspire other programs toward rapid release of CTD, hydrographic, ocean carbon, and tracer data from cruises for the Repeat Hydrographic Program.

The oceanographic community is best served when data are made public as soon as feasible after a cruise. Some measurement programs - or the agencies or governments which fund them - require public distribution within a specified time interval after data collection. The CCHDO does not require submission by a particular date. It is a service organization. Nor does the CCHDO make data public without the approval of the data providers. (If necessary, the CCHDO can provide password-only access to data so that they are usable only by persons approved by the data providers, who control distribution of the access passwords.)

It is well known that post-cruise analysis rarely changes most shipboard-analyzed data to a scientifically significant extent. The core rationale for the interval between data collection and public release is, in most cases, the protection of the proprietary rights of the data originators. The desire to protect ocean profile data is entirely reasonable, considering the level of intellectual effort and time involved in proposing, planning, collecting, calibrating, and processing these data. But nearly all data providers also recognize the value of putting their data to use.

First, science is advanced when data are made public in a timely manner (and it is rare that other scientists have research interests that significantly overlap those of the data originators). In fact, early community collaboration on data analyses has lead many data originators to expand and improve their own scientific analyses.

Second, experience has shown time and time again that data users are extraordinarily effective in vetting data for all manner of problems. When these problems are recognized while the cruise is relatively fresh in the minds of the data originators, and while their records are close at hand, it is vastly easier to resolve the matters than it is years later.

Third, relatively rapid data release may be an agency or overall program requirement.

Fourth, although there are no CCHDO "requirements" regarding data delivery schedules, when data are delivered to the CCHDO they will be handled by the CCHDO, where a series of file and data checks are carried out. These checks are useful in locating certain types of data problems, so there is a clear advantage to rapid delivery of new ocean profile data to the CCHDO.

The CCHDO requires those who send data to the CCHDO to indicate whether or not the data are to be made public. The CCHDO will post for public use only those data for which it has received permission to do so.

Investigators who use public data from the CCHDO are urged to cite the data originators in their publications, and in the first several years after availability, contact the data originators prior to presenting or publishing results. This contact is not only a professional courtesy to the data originators: it is also an opportunity for the data originators to alert data users of any pending or unusual developments with the data and their analyses.

In the case of proprietary data, the data user should, of course, not publish any paper during that period based predominantly on those proprietary data, should offer to coauthor results with the originating investigator, and should not redistribute the data.

2. TYPES OF DATA TO BE SUBMITTED TO THE CCHDO

The CLIVAR and Carbon Hydrographic Data Office (CCHDO) deals with what oceanographers term "hydrographic data", meaning vertical profile data from CTD, CTD/rosette, and bottle casts covering physical and chemical parameters. Many hydrographic data arise from studies related to water mass and ocean circulation, such as the WOCE Hydrographic Program in the 1990s and the International Ocean Carbon Coordination Project in the 21st century. But any vertical profile data from CTD instruments and bottles may be submitted. Also, there is no restriction on the region or areal coverage of cruise data sent to the CCHDO, including the Arctic regions.

Data providers should take care to ensure that their data are reported in community standard units, because it is important when merging or using data from different sources that units be consistent. But it is much more important that the units be stated exactly correctly. *Data originators should always double-check that the actual units for their reported data are indeed the units listed.*

The CCHDO's principal interest is the suite of parameters listed in **Appendix B** of this document. However, if there are other parameters in a data file, they will be retained by the CCHDO. For example, if a vertical profile data file contains chlorophyll data, the CCHDO will preserve those chlorophyll data in the data file.

Not all data from an expedition are to be submitted to the CCHDO. For example, data from hull-mounted and/or lowered ADCP measurements made during a cruise would go to a separate data center. Only a brief description of the measurements, and contact information for the investigator or data group responsible for the data, in the documentation file sent to the CCHDO.

A. CTD Data

A directory of _ct1.csv data files, or a zipped _ct1.zip directory of _ct1.csv data files (described in **Appendix C).** CTD data should be accompanied by a quality flag for each measured parameter. The quality flag shows the data originator's evaluation of data quality. Quality flags for CTD data are optional but recommended. (CCHDO and IGOSS CTD quality flags are described in **Appendix D**.)

It is expected that CTD data will be reported to the CCHDO in a uniform pressure series in order of increasing pressure with a pressure resolution of 2 dbar. (Sometimes 1 dbar records are submitted, which is also satisfactory.)

The CCHDO will accept CTD data in other formats, but this may result in delayed on-line posting based on the time it takes to do file conversion, and on the queue of CTD data files awaiting conversion.

B. Bottle Data

A self-contained bottle data file, preferably a _hy1.csv file (i.e., in WHP-Exchange format), contains the results from all the small volume water sample measurements made during the cruise, along with header information on each row of the file, so that the complete record for each water sample, when combined with the headers, is contained within one row in the file. The format for the water sample data is described in **Appendix C**.

Columns and a quality flag for every parameter measured on the cruise should be included in the _hy1.csv file even though some of the data values may be submitted at a later date. This alerts the CCHDO and data users that additional parameters are expected. The CCHDO will merge the additional parameters into the _hy1.csv file when they have been received. Quality flags for the bottle and each measured parameter are optional but recommended. (CCHDO and IGOSS sample bottle and bottle parameter quality flags are described in **Appendix D**.)

Any and all bottle data parameters may be included in the _hy1.csv file. This provides all the available bottle data in a single file. It is not required that cruise leaders carry out all of the merging of bottle data from different data originators. For a typical example, after the original bottle data file, containing the data from the shipboard analyses, is submitted to the CCHDO, cruise leaders are not required to merge subsequent updates containing shore-analyzed data or corrections to the shipboard data. So long as the new or updated bottle data parameters are submitted to the CCHDO, and are accompanied by the cruise, station, cast, and sample number (or bottle number) for each row of data, the CCHDO can merge the various data to make the complete bottle data file.

The CCHDO will accept bottle data in other formats, but with delay in on-line posting due to queue for file conversion.

C. Station Summary (_.sum file) [optional]

The WOCE Hydrographic Program created a type of station summary file, known as a .sum file, which was to be submitted for each cruise. The intent - to gather together the most critical station header information into a compact ASCII file - was laudable. But the problems with .sum files were legion during WOCE. The problems centered on: (1) the definition of the .sum file was not sufficiently precise that a single .sum file reader could be written which would successfully read all the WOCE .sum files, (2) there were insufficient guidelines regarding dealing with missing information, and (3) it was not possible for a data center to unambiguously repair defective .sum files to the stated standard.

Therefore, the WOCE-era 'requirement' for a .sum file has been abandoned by the CCHDO. The 'WHP-Exchange" formats (see **Appendix C**) do away with the .sum file entirely.

Data providers may submit .sum files with their data. The CCHDO therefore provides the original .sum file description at <http://cchdo.ucsd.edu/manuals/pdf/90_1/chap4.pdf>. Station summary files (.sum files) submitted to the CCHDO will be posted on the CCHDO web site. The CCHDO has the capability to make a rudimentary .sum file from the information in the WHP-Exchange CTD and bottle data files.

Note that .sum files permit the data originators to supply a position, time, and the uncorrected depth for the BEginning, BOttom, and ENd of every CTD/hydrographic cast. That allows an estimate of the drift of the ship and whether or not the ship drifted across significant topography during the cast. When a single time and position are provided for a cast (as is done in WHP-Exchange data files) the BOttom time and position are normally used, but if this is not available, investigators should follow the procedures of their laboratory, and state what this procedure is in their documentation accompanying the data.

D. Cruise Report (.doc file) or other documentation

To a large degree it is the documentation of data collection and quality control for each parameter that gives the data lasting reference value.

The documentation file is sometimes referred to as the cruise report or the .doc file. The outline for a cruise report (.doc file) is provided in **Appendix A**. Cruise and data documentation are also discussed further in Section 3.

3. DOCUMENTATION

Documentation lies near the heart of reference-quality data. A cruise report (or data report), conveys to eventual data users all the information needed to understand the context and content of the reported data. Published, easily obtainable documents can be referenced in the cruise report, but data originators are strongly urged to err on the side of completeness: the documentation should be written to be useful for literally many decades. (It should be noted that this is only difficult to do the first time; with electronic word processing it is easy to generate a new cruise/data report from a successful earlier effort.) We have tried to avoid overuse of the word 'required' in this guide, but comprehensive data documentation is considered by data users and data archives to be truly essential. The reward for this extra effort is an oceanographic legacy of data which will be valuable for centuries.

The CCHDO strongly prefers that all cruise documentation and data reports be in English. If this is simply too great a burden for the data providers, then a report can be sent in another language but it will be distributed as received, and will not be translated by the CCHDO. (Sorry, but the CCHDO does not have the funds or staff to translate reports.)

Data originators are asked to take special care preparing electronic versions of figures and of reports containing figures for submission to the CCHDO. Use of common graphics and word processor file formats makes it much easier for the CCHDO to handle and post documents accurately. For reports and figures which are intended to be placed on the CCHDO web site unaltered, the most reliable standards (at the time of this guide’s preparation) are the "pdf" and "png" formats, for which free reader applications are available for nearly all computer operating systems in common use.

Paper copies of reports and data will be accepted by the CCHDO if that is all that is available. Please note, however, that the CCHDO will often simply make an electronic version by scanning the paper report, which may result in a loss in quality and readability.

A. Pre-Cruise Information (Optional.)

The sooner the CCHDO knows about your cruise the more it can do to help you get your work out to the oceanographic community, but the CCHDO generally has no way of knowing about a cruise unless informed by the cruise participants and leaders. A pre-cruise plan is a good way to get started. Once alerted to your cruise, the CCHDO can provide suggestions regarding documentation and data file formats which will save you time and effort once your cruise is underway.

A useful pre-cruise plan would include:

- an overview of the cruise describing your scientific goals,
- a list of the Principal Investigators (PIs),
- a list of measurement groups, measurement team leaders (and institutions), measurements planned, and who will be responsible for them.
- contact information for the expected data providers,
- a preliminary cruise track,
- estimates of the total number (and type) of hydrographic stations,
- a logistical summary.

If permission is given by the author(s) of the cruise plan, the CCHDO can post it in on line.

B. Initial Cruise Report (Optional.)

As with the pre-cruise plan, this step is optional, and in no way replaces the much more important documentation which will accompany your data. The main purpose of an initial cruise report is to inform the CCHDO of the post-cruise status of the sea program. This can be as simple as a brief email, or a more comprehensive discussion of the measurements taken, problems encountered, and the status of the data. If you send the CCHDO a list of stations (station numbers, positions, and station dates), we can produce a cruise track and post it on the CCHDO website. This will alert the community to the data which will eventually be posted at the CCHDO.

C. Cruise Report - from the Chief Scientist and measurement groups (Essential.)

Documentation should be submitted with data whenever possible. A cruise report (sometimes called a data report) is a detailed report of the cruise and the data released from the cruise. A comprehensive document typically prepared by the Chief Scientist with the help of the various data originators, the cruise report is best done at sea while the teams are together and memories are fresh.

The cruise/data report documents the measurements made and methods used to make them, describes the contents of the data files (headers, parameters, units, quality flags, etc.), and provides complete contact information for each data provider. In most cases individual data originators send documentation for their measurements to the chief scientist for inclusion in the cruise/data report, but in other cases the data originators send their individual reports directly to the CCHDO along with their data. Either system is acceptable, the principal point being that the documentation is generated and included with the data.

The CCHDO may choose to reformat a cruise report, or add additional materials received from other data originators, data quality experts, or data users. The original materials are retained and archived by the CCHDO.

Because uniform, consistent cruise-related data documentation is needed from a diverse community, **Appendix A** contains the CCHDO's recommendations for the structure and contents of a cruise report. This standard format is currently in use by the CCHDO. Documents submitted in this format are easier to digest by readers. It is more important, however, to submit complete, accurate documentation than to rigidly follow this template.

Note that each measurement group - for example, CTD, nutrients, CFCs, helium, etc. - provides documentation in sufficient detail to establish how their data were created and to help assess the data quality. References to the analytical methods used should be supplied, and variations from these techniques described. Techniques should be described in detail if no published reference exists.

It is extremely helpful if each measurement group's report includes an assessment of the uncertainty of the measurements and notes of any problems peculiar to the data gathered during the cruise. For example, the CTD documentation provided in the cruise report should ideally address the following issues: List the instruments used during the cruise, any unusual problems, for example, aborted casts, sensor fouling, etc., or special procedures employed during the cruise that would affect data quality, such as cell cleaning, sensor replacements, etc., and specify station numbers affected by these events if applicable. The pre- and post-cruise laboratory calibration information should be included together with the coefficients used to fit CTD pressure and temperature. The conductivity and oxygen calibration coefficients and equations, and the station groups used when fitting to water sample data should also be included. Reference the standard processing procedures employed, or describe variations from these methods and identify the stations on which nonstandard methods were employed.

4. FORMATS

In order that a diverse community can exchange data easily (and to enable the CCHDO to ensure that the data can be examined and distributed by a minimum number of staff), it is preferred that all investigators make a reasonable effort to use standard formats, consistent naming conventions, and common units for data reporting.

Easy to write and read, rigorously-defined exchange file formats were developed near the end of WOCE for WHP-like CTD and water sample data. They are the so-called "WHP-Exchange" formats. In addition to their value for exchanging data, the formats ensure that WHP-like data are reported correctly, have provisions for quality flags for each measured parameter, and provide other information needed in a data file to help assure its maximum long-term usefulness. These data formats are described in **Appendix C**.

> Specifically, the CCHDO requests that when possible all CTD and bottle data be submitted in the WHP-Exchange formats (see **Appendix C**).

Post-cruise data updates for individual parameters should be sent as ASCII tables with columns for cruise, station, cast, and sample or bottle number in addition to the columns for the submitted data parameters and their accompanying data quality codes.

For each expedition, the data submitted to the CCHDO should thus include the following files:

Suffix for file type	Description of file
do.pdf do.txt	Cruise Report (e.g., cruise and data documentation in MS Word, pdf, or ASCII).
_hy1.csv	Bottle data file (if any water sampling took place and is being reported).
_ct1.csv or _ct1.zip	CTD data file(s).
.SUM (optional)	Station/cast summary file in original WOCE format (see <http://cchdo.ucsd.edu/manuals/pdf/90_1/chap4.pdf> for format description).

An original WOCE-format .SEA file is also acceptable instead of a _hy1.csv file, but must be accompanied by a WOCE-format .SUM file. (See

<http://cchdo.ucsd.edu/manuals/pdf/90_1/chap4.pdf> for format description.)

An original WOCE-format .CTD file is also acceptable instead of _ct1.csv or _ct1.zip files, but must be accompanied by a WOCE-format .SUM file. (See

<http://cchdo.ucsd.edu/manuals/pdf/90_1/chap4.pdf> for format description.)

One advantage of submitting a .SUM file is that it can be submitted without the bottle or CTD data, i.e. when those data are still proprietary/non-public. Then Program Managers, data users, and other colleagues will know the spatial coverage of the cruise, and be alerted to look for the bottle and/or CTD data when those later become public.

5. CRUISE/STATION/CAST/SAMPLE NUMBERING SCHEME

Each CTD/hydrographic profile, and each water sample from a rosette water sample bottle, must be uniquely identified in data records. Among other purposes, this helps to assure that when the data are combined with whose from other expeditions, the data remain correctly identified.

Also, for water sample data, it is essential that the intended sample depth **never** be used to index water samples. There are many good reasons for not using sample depth to index water samples, such as discovering later (sometimes months or years later) that the rosette bottle did not close at (or sometimes even near) the intended sample depth. We may dislike overuse of the word "require" but a unique identifier is indeed **required**.

The following **cruise/station/cast/sample** numbering scheme is proven to provide unique, traceable identifiers (see **Appendix C** for additional information):

First, each **cruise** in the data set is given a unique identifier, for example the EXPOCODE (also see **Appendix C**). The EXPOCODE should appear in every file.

The convention the CCHDO uses to create the EXPOCODE is:

```
ExpoCode Syntax:    NODCShipCodeYearMonthDay
Example:
   Ship Name:          Roger Revelle
   Cruise start date:  March 29, 2009
   EXPOCODE:           33RR20090329

ExpoCodes created before January 2006 will keep their established name,
but will also be discoverable by the newer style ExpoCode.

Definitions:

NODC Ship Code - Ship Codes can be found at the NODC's website
<http://www.nodc.noaa.gov/General/NODC-Archive/platformlist.txt>.

Day - Official ship's log date of departure.

If ship has an emergency return to port then continues the cruise, use
original date whether or not stations were occupied before the emergency.

If departure date is unknown, date of first station can be used.  Beware
of shakedown station complications.

Note all relevant information related to this alternate day in the
documentation.
```

Second, each **station** and **cast** on a cruise must be given a unique (alphanumeric) identifier.

Repeat and time series cruises may use the same station number on sequential cruises and that presents no difficulties as long as each cruise is given a unique EXPOCODE.

The *third* identifier is a cast number. Each over-the-side operation at a station should be given a separate cast number. If a station is reoccupied later during the same cruise, and the same station number is used, the cast numbers should increment upward in some sensible manner. In no case should a data file contain at different points the same pair of STNNBR and CASTNO on the same cruise. STNNBR and CASTNO appear in all .SUM, _hy1.csv, and _ct1.csv files submitted to the WHPO and must be used consistently in each file in which they appear. All files in the CCHDO data warehouse are tied to the string EXPOCODE, STNNBR, and CASTNO.

The *fourth* identifier is the sample number (SAMPNO) [or bottle number (BTLNBR)]. Many groups use the sample number to identify the rosette position of the bottle from which water samples are drawn while on deck. Other groups are using schemes where unique and consecutive sample numbers are printed on sticky labels before the cruise. For each bottle sampled, an identifying number is assigned and identical multiple labels are printed beforehand to allow this same number to be attached to each and every subsample drawn from the bottle and, thus, track the sample collection bottles as well. Other schemes are possible. Whatever scheme is used it is critical that either (1) the same numbering scheme be used by all participants on the cruise, or (2) the bottle data file contain both identifiers. Great confusion arises if one group uses one sample numbering scheme and another group uses a different one on the same cruise, but with the bottle data file containing only one of the two schemes. (This has happened several times in CCHDO records.) Water samples are analyzed at various times and places, sometimes years after the cruise, as is the case with AMS 14C and helium/tritium samples, for example. Inconsistent identification of the bottle from which the sample was drawn can make it difficult, if not impossible, to merge the various measurements into the database when the data are finally assembled either by the chief scientist or the CCHDO.

In summary, each bottle sampled should be identified by a combination of EXPOCODE, STNNBR, CASTNO, and SAMPNO. This combination must be capable of uniquely identifying every water sample collected from that bottle.

For WHP-Exchange format bottle data files (_hy1.csv) the bottle number (BTLNBR) is also needed in order to uniquely identify the particular device used to collect the water sample. The bottle number is defined as: (1) the permanent, unique serial number (alphanumeric) stamped or engraved on the barrel of the bottle from which the water samples are drawn or, alternatively, (2) as a unique alphanumeric identifier assigned to only that device over the duration of the expedition.

6. DATA QUALITY EVALUATION AND DATA QUALITY FLAGS

The terms "quality code", "quality flag", and "quality byte" have the same meaning in this document.

Data originators should provide a quality flag for each measured value.

For ocean climate and global change studies great emphasis is placed on documenting the quality of the data, so that the data have a long service life. In order to monitor, evaluate, maintain, and later access the data it is advisable to keep individual records for the quality of each measurement. To help accomplish this goal, quality flags are used. The flags describe the quality of the measurement they relate to in a machine-readable fashion.

The use of quality flags for data provides valuable information to data users. Most important is a degree of assurance: when the quality flag is "good", the data have passed some degree of scrutiny as to their suitability for use. A quality code of "questionable" or "bad" provides the user the option of retaining or ignoring a data value, yet the questionable or bad data are retained in the data file, where they can be subject to further inspection. (In some cases such data are later found to be good, or can be reinterpreted in some useful way.) Finally, a clever quality code scheme can provide additional information about missing or expected data.

Quality flags arise from quality control procedures. Such procedures are at the heart of the generation of reference-quality data, but they are not discussed in this data submission guide. [The author's guide, "Reference-Quality Water Sample Data, Notes on Acquisition, Record Keeping, and Evaluation", available from <http://cchdo.ucsd.edu>, contains useful information.]

Note that for water sampling the rosette bottle itself has its own unique quality flag (defined in **Appendix D**) to indicate problems associated with the sampling device that were noted either at the time the samples were drawn, or found later during quality control. The combination of BTLNBR and its quality flag allows 'problem' bottles to be identified and tracked. For example, individual bottles may leak chronically, may contaminate the samples being taken (e.g., chlorofluorocarbons), or may consistently mistrip or otherwise malfunction. Such bottles are usually discovered early in a cruise and replaced on the rosette for that cruise. However, the defective bottles may subsequently be reused on following legs with different personnel, but using the same equipment, if a replacement bottle is needed. Hence, users are urged to maintain a tracking capability transferable to other cruises. If no manufacturer's serial number is available for the sampling device, identifying numbers may be written on labels attached to the bottle.

Reference

Joyce, T. and C. Corry (eds), Requirements for WOCE Hydrographic Programme Data Reporting, WHPO Publication 90-1 Revision 2, WOCE Report 67/91, Woods Hole, Mass., USA, May 1994.

Appendix A. OUTLINE OF A CRUISE REPORT

> See <http://cchdo.ucsd.edu/data/onetime/atlantic/a24/a24do.pdf> for an example of a well-completed cruise report. It was assembled and reformatted by the CCHDO from separate contributions from the measurement teams.

A. Cruise narrative

1. *Highlights*

a. *Cruise designation (cruise name)* (e.g., "AIS01" for Amery Ice Shelf 1) [or, for a repeated WOCE-era section, the *section designation(s)*, for example A11, IR04, etc. Include all sections covered on cruise].
b. *EXPOCODE* (the CCHDO EXPOCODE convention is discussed on p. 11). Each leg has a unique EXPOCODE).]
c. *Chief scientist* for each leg, including postal and electronic addresses.
d. *Ship name*.
e. *Ports of call.* Port(s) where cruise begins and ends plus any stops during the cruise.
f. *Cruise dates* (Official ship's log date of departure & return, port to port, for each leg).

2. *Cruise Summary Information*

a. *Written description of the survey's geographic boundaries.*
b. *Cruise track showing each station with different symbols used to indicate station type.*
b. *Total number of stations occupied for each section, broken down by type of station and parameters sampled at each station.*
c. *Detailed list of each and every parameter measured on the cruise.*
d. *Floats and drifters deployed (type, identification number, location, and time).*
e. *Moorings deployed or recovered (type, identification, location, and time).*

3. *List of Principal Investigators for All Measurements*

a. *Name (please spell out).*
b. *Measurement responsibility.*
c. *Institution or affiliation (abbreviations should be defined).*
d. *Postal and electronic addresses.*

4. *Scientific Program and Methods*

a. *Narrative.*
b. *Interlaboratory comparisons made (if any) or comparisons with previous cruise data.*
c. *(Optional) Vertical sections along the ship's track showing the bottle depth distributions, and plots of property vs. property relationships.*

5. *Major Problems and Goals Not Achieved*

6. *Other Incidents of Note*

7. ***List of Cruise Participants***
   a. *Name (please spell out).*
   b. *Responsibility on cruise.*
   c. *Institution or affiliation (abbreviations should be defined).*
   d. *postal and electronic addresses.*

B. Underway Measurements

1. Navigation and bathymetry
2. Acoustic Doppler Current Profiler (ADCP)
3. Thermosalinograph and underway dissolved oxygen, fluorometer, etc.
4. XBT and XCTD
5. Meteorological observations
6. Atmospheric chemistry

C. Hydrographic Measurements — Descriptions, Techniques, and Calibrations

A discussion of each measurement type, including the following sub-headings as they apply:

1. Measurement name(s)
   section author's name(s) (date, or revision date)

   a. Description of equipment and technique or *published* reference.
   b. Sampling and data processing techniques followed or *published* reference for these techniques.
   c. Calibration data, including dates and laboratory where calibrations were done.
   d. Error estimates and noise sources, including:
      • effect of noise on samples and
      • comparisons with historical data or test stations.
   e. Laboratory and sample temperatures where required.
   f. Replicate analyses (tables).
   g. Standards used (for example, standard sea water batch number and ampoule number of standard sea water for each station).
   h. Reagents: purity and concentrations of stock solutions, where applicable.
   i. Values for blanks, where applicable (blank values should be subtracted from the data).
   j. Atmospheric values for tracers, where applicable.

D. Acknowledgments

Funding sources, contract numbers, contributors, etc.

E. References

F. Appendices

Deck logs, water sample quality assessment notes, etc.

Appendix B. Measurement Units and Data Quality Goals

1. Definitions of accuracy, reproducibility, precision, and limits of detection.

In the interests of uniformity the CCHDO uses the following definitions for accuracy, reproducibility, precision, and limits of detection:

***Accuracy*:** Accuracy is defined as the extent to which a given measurement agrees with an accepted standard value for that measurement.

Unfortunately, for some commonly-measured parameters there are no international community standards yet available. A secondary estimate of accuracy is consideration of the standard deviation of inter-laboratory reproducibility. This is computed from the results of a collaborative study of an analytical method involving operators in different laboratories using different apparatus for analysis of the same sample.

***Reproducibility*:** As used here, reproducibility is the total intra-laboratory standard deviation of a series of measurements. This parameter is the maximum intra-laboratory standard deviation to be expected from the performance of a method, at least on different days and preferably with different calibration curves or reagents, e.g., repeat cruises. It includes between-run as well as within-run variations. Thus, the reproducibility is always larger than the precision but it is not a measure of the accuracy of the measurement, although the two are often mistakenly used interchangeably.

***Precision*:** Precision is defined as the extent to which a given set of measurements of the same sample agree with their mean. Thus, precision is commonly taken to be the standard deviation estimated from sets of duplicate measurements made under conditions of repeatability, that is, independent test results obtained with the same method on identical test material, in the same laboratory by the same operator using the same equipment in short intervals of time. There are, however, many definitions of 'precision' in use. It is important to state how a precision figure is generated.

***Limit of Detection*:** A measure of the concentration of the substance being analyzed that is significantly different from the blank or background signal.

2. Units and community data quality goals

Present-day units and community data quality goals for reference-quality CTD and water sample data are discussed below. The suggested data quality goals for most parameters are appropriate for shipboard measurements in low-gradient portions of the water column.

T (electronic): Oceanographic units are degrees Celsius on the International Temperature Scale, 1990 (ITS-90). [Note, however, that as of this writing the algorithms for calculation of potential temperature, density, and salinity require temperature stated on the older International Practical Temperature Scale, 1968 (IPTS-68). ITS-90/IPTS-68 conversion to adequate accuracy for oceanographic work over ocean temperature ranges is a simple multiplicative factor:

$$T90 = T68 * 0.99976$$
$$T68 = T90 * 1.00024$$

Data originators should thus note that <u>it is very important that the temperature scale used be reported as part of the temperature units</u>.] Data are typically reported to the nearest 0.0001 °C. The temperature sensors provided with the most common CTD instrument in use for reference-quality measurements (at the time of writing), the Sea-Bird 911*plus*, appears to be capable of inter-cruise accuracy near 0.002°C and intra-cruise (e.g., cast-to-cast) precision smaller than 0.001°C. Drift and sudden offsets (the latter typically related to some form of shock) can and do occur. These can be monitored by *in situ* use of dual sensors and/or a SeaBird SBE-35 slow-response reference thermometer, or equivalent reference device or scheme.

T (discrete): [See above for units and scales.] Oceanographic use of deep-sea reversing thermometers (DSRTs) has all but ceased, but there remain a few occasions where they may be usefully employed. Much depends on the construction and range of the thermometer. High resolution, low-range mercury DSRTs are available and with careful calibration and reading may be capable of 0.004–0.005 °C accuracy and 0.002 °C precision. Data are typically reported to the nearest 0.001 °C for high-precision instruments and the nearest 0.01 °C for standard oceanographic DSRTs. Digital DSRTs do not require long soaking times and can serve as a means for calibration and performance checks. Carefully documented and monitored use of multiple CTD sensors have all but eliminated the standard use of DSRTs. For reference-quality measurements temperature is always measured with a CTD.

P: Oceanographic units are decibars. Data are typically reported to the nearest 0.1 decibar. Although pressure can be determined with unprotected DSRTs, for reference-quality measurements pressure is measured with a CTD. With the Sea-Bird 911*plus*, inter-cruise accuracy near 2 decibar (dbar) and intra-cruise precision near 0.5 dbar is feasible, dependent on processing. Of particular concern is the possibility of offsets due to shock, including both mechanical and thermal shock.

S (electronic): Salinity is officially unitless, though units can be stated 'PSS-78' or 'psu'. Data are typically reported to the nearest 0.0001. Inter-cruise accuracy of ≈0.002 psu and intra-cruise precision of ≈0.001 are feasible, the former depending upon the frequency and technique of calibration, and also upon data processing methodology. Conductivity is measured, but salinity is always reported for climate and global change related ocean profile measurements. Despite improvements in conductivity sensor stability, to obtain reference-quality CTD salinities it remains necessary to correct measured conductivities to conductivities derived from discrete water samples collected at the surface, the maximum cast depth, and at low-gradient points throughout the water column.

S (discrete): [See above for units and scales.] Inter-cruise accuracy of 0.002 is possible with Autosal™ salinometers and concomitant attention to methodology, e.g., close monitoring with IAPSO Standard Sea Water. In waters with low spatial and temporal variability intra-cruise accuracy with respect to one particular batch of Standard Sea Water can be achieved at 0.001 PSS-78. Intra-cruise Autosal™ precision is better than 0.001 PSS-78, but great care and experience are needed to achieve these limits on a routine basis. For example, laboratories with air temperature stability of ±1°C are necessary for optimum Autosal™ performance. To avoid changes which may occur in Standard Seawater, the use of the most recent batches is recommended. It is necessary to include in cast documentation the batch number of the IAPSO Standard Seawater.

O2 (electronic): Preferred oceanographic units are μMkg^{-1}, although mll^{-1} remain in use (see below). Data are typically reported to the nearest 0.01 µMkg-1. Experience with the Sea-Bird SBE-43 pumped dissolved oxygen sensor suggest that with care in correction against discrete water samples collected at the surface, the maximum cast depth, and at low-gradient points throughout the water column, 'excellent' (the degree of 'excellence' is undefined here) fits to the dissolved oxygen values from the water samples are often feasible. [Note: A sizeable contingent within the physical oceanographic community continues to prefer dissolved oxygen data expressed in the traditional units, mll^{-1}. In those units, dissolved oxygen data should be reported to the nearest 0.001 mll-1. Unit translations, if necessary, can be carried out by a data center.]

O2 (discrete): [See above for units and scales, and for note regarding traditional units.] No international laboratory standards yet exist. Inter-cruise reproducibility of <1% (full scale) and intra-cruise precision ≈0.1% is recommended for reference-quality measurements. Some laboratories have found it possible to achieve reproducibility <0.5% over multiple cruises carried out by that laboratory, referenced to a single batch of potassium iodate standard.

NO3: Oceanographic units are μMkg^{-1}. Data are typically reported to the nearest 0.01 μMkg^{-1}. No international laboratory standards yet exist. Approximately 1% inter-cruise reproducibility (full scale) and 0.2% intra-cruise precision is recommended for reference-quality measurements. NO3 is typically first determined as NO3+NO2, with NO2 determined separately and then subtracted out. <u>It is therefore vital to note if, in fact, NO2 has been subtracted out.</u> If NO2 has not been explicitly subtracted out, and if the methodology did not determine NO3 alone, then this parameter should always be listed as 'NO3+NO2' in data lists and documentation. Many older nutrient data are expressed in volume units (e.g., μMl^{-1}). The numerical difference of nutrients expressed in µMkg-1 differs by only ≈3% from nutrients expressed in μMl^{-1}. Hence <u>it is also essential that nitrate units be explicitly verified as part of data reporting and usage</u>.

NO2: No international laboratory standards yet exist. Concentrations are low over most ocean regions. Approximately 0.04 μMkg^{-1} inter-cruise reproducibility and 0.02 μMkg^{-1} intra-cruise precision are recommended for reference-quality measurements. Note that the numerical difference of nutrients expressed in common weight units differs by only ≈3% from nutrients expressed in common volume units. Hence <u>it is essential that nitrite units be explicitly verified as part of data reporting and usage</u>.

PO4: No international laboratory standards yet exist. Approximately 1.5% inter-cruise reproducibility (full scale) and 0.4% intra-cruise precision are recommended for reference-quality measurements. Note that the numerical difference of nutrients expressed in common weight units differs by only ≈3% from nutrients expressed in common volume units. Hence <u>it is essential that phosphate units be explicitly verified as part of data reporting and usage</u>.

SiO3: No international laboratory standards yet exist. Approximately 1% inter-cruise reproducibility (full scale) and 0.2% intra-cruise precision are recommended for reference-quality measurements. The numerical range of silicate over the water column varies geographically by approximately one order of magnitude. This affects data quality guidelines expressed in concentration terms. Also note that the numerical difference of nutrients expressed in common weight units differs by only ≈3% from nutrients expressed in common volume units. Hence <u>it is essential that silicate units be explicitly verified as part of data reporting and usage</u>.

3H: reproducibility ≈1%; precision 0.5% with a detection limit of 0.05 tritium unit (TU) in the upper ocean of the northern hemisphere and 0.005 TU elsewhere.

d3He: reproducibility/precision 1.5 per mille in isotopic ratio; absolute total He of 0.5% with less stringent requirements for use as a tracer (e.g., He plume near East Pacific Rise).

CFCs: approximately 1–2% reproducibility and 1% precision, blanks at 0.005 pmol/kg with best technique.

D14C: reproducibility and precision 2 to 4 per mille via beta-counting on 200-liter samples; 5–10 per mille with Accelerator Mass Spectrometer (AMS) on 500 ml samples.

d18O: Measured mostly in high latitude oceans; these should be measured with a reproducibility of 0.02 per mille.

Appendix C. Description of WHP-Exchange Format for CTD/Hydrographic Data

An Improved Exchange Format for Hydrographic Data

September 2001; updated May 2006 & April 2008

James H. Swift & Stephen C. Diggs
CLIVAR and Carbon Hydrographic Data Office (CCHDO)
UCSD Scripps Institution of Oceanography
cchdo@ucsd.edu

Summary

Exchange formats for the CCHDO CTD and bottle data are described. The WHP-exchange formats provide simplified exchange and improved readability of hydrographic data. WHP-exchange data files carry the essential information from CTD and water sample profiles in rigorously-described comma-delimited (csv) ASCII formats designed to ease data exchange and simplify data import.

1. Overview of WHP-exchange file formats

The WHP-exchange bottle and CTD data formats include these features:

- ASCII, spreadsheet-like
- comma-delimited values (csv)
- no special meaning to blank/empty spaces
- station information in every line in the file (bottle) or in the top lines in each file (CTD)
- only one missing data value defined for all parameters
- missing data value format defined in the format for each parameter
- WHP quality flag, when provided, associated directly with its parameter
- positions in decimal degrees
- dates in YYYYMMDD format.

There are three types of WHP-exchange format files, each with a unique **8-character** suffix:

data type	8-character suffix	description
CTD data	_ct1.csv	one CTD profile in WHP-exchange format
	_ct1.zip	zipped directory holding one or more _ct1.csv WHP-exchange CTD profiles
bottle data	_hy1.csv	data from one or more bottle profiles in WHP-exchange format

2. Format description for WHP-exchange bottle data (8-character suffix _hy1.csv)

[Note: To better understand this section please refer to one of the WHP-Exchange bottle data files available from the CCHDO. The file "a24_hy1.csv" from the CCHDO from

<http://cchdo.ucsd.edu/data/onetime/atlantic/a24/a24_hy1.csv>

accompanies the cruise report referenced in this document. It is a good example. It is recommended that the reader examine "a24_hy1.csv" both in a text editor application - in order to see all characters - and also in a spreadsheet application - in order to view overall layout.]

The overall layout of a _hy1.csv bottle data file is described in Table 1.

The first line ("line" = "row") of a WHP-exchange format file is a single word which describes the file type, in this case "BOTTLE", followed by a comma and a date/time stamp.

The format next provides for 0-N optional information lines, each beginning with a "#" character, near the beginning of a _hy1.csv file. The CCHDO uses "#" lines to hold file history and data citation information referring to the data originators.

A description of the station information columns of a _hy1.csv file is in Table 2.

A description of the remaining data columns and preferred parameter names in a _hy1.csv files is in Table 3.

A line with "END_DATA" signals the end of the data lines.

After that line, a bottle data file may hold other file-specific documentation. The primary documentation for WHP data will, however, remain in the ".doc" file (or zipped directory).

General rules for WHP-exchange_hy1.csv data files:

Each line must end with a carriage return or end-of-line.

With the exception of (1) the file type line, (2) lines starting with a "#" character, or (2) including and following a line which reads "END_DATA", each line in a _hy1.csv file must have exactly the same number of commas as do all other lines in that file.

The number and names of the parameters in a _hy1.csv file is not specifically addressed, except that for WHP data certain parameters are noted as REQUIRED. For example, it is not necessary that a bottle data file contain columns for CFC measurements when there are no CFC data.

The order of the header and bottle data parameters in a _hy1.csv file is preferred to be similar to that shown in the example "a24_hy1.csv", especially for the first 13 columns, but is not strictly required. Although the _hy1.csv files should be as consistent as feasible in this regard, data users are urged to use "read" statements that are sensitive to parameter names rather than position of the parameter in the data files. Here is the order used in " a24_hy1.csv":

> EXPOCODE, SECT_ID, STNNBR, CASTNO, SAMPNO, BTLNBR, BTLNBR_FLAG_W, DATE, TIME, LATITUDE, LONGITUDE, DEPTH, CTDPRS, CTDTMP, CTDSAL, CTDSAL_FLAG_W, SALNTY, SALNTY_FLAG_W, CTDOXY, CTDOXY_FLAG_W, OXYGEN, OXYGEN_FLAG_W, SILCAT, SILCAT_FLAG_W, NITRAT, NITRAT_FLAG_W, NITRIT, NITRIT_FLAG_W, PHSPHT, PHSPHT_FLAG_W, CFC-11, CFC-11_FLAG_W, CFC-12, CFC-12_FLAG_W, TRITUM, TRITUM_FLAG_W, HELIUM, HELIUM_FLAG_W, DELHE3, DELHE3_FLAG_W, TCARBN, TCARBN_FLAG_W, PCO2, PCO2_FLAG_W, ALKALI, ALKALI_FLAG_W, PH, PH_FLAG_W, PCO2TMP, CTDRAW, HELIER, DELHER, THETA, TRITER

All parameters defined as alphanumeric (e.g. ,"A14") and integer (e.g., "I4") will be shown in the full defined width and will be right-justified, meaning that entries shorter than the defined width will be padded with meaningless spaces to the left of the first character (for example, EXPOCODEs are usually shorter than the defined maximum of 14 alphanumeric characters).

The bottle data parameter names should follow those listed in Table 3 when feasible. Data providers are urged to use caution, however, and list their actual parameter name rather than a WHP parameter name whenever there is any question on this matter.

Each data parameter listed in Table 3 - except for all flags, which are "I1" - will be listed in "F9.x" floating point format, where "x" indicates the number of decimal places. For each parameter, the CCHDO will pad with meaningless zeros data received with fewer decimal places and round data received with extra decimal places to the number of decimal places specified in Table 3.

When a quality flag is available for a parameter, that quality flag shall be placed in the column immediately to the right of the parameter. The name of a quality flag always begins with the name of the parameter with which it is associated, followed by an underscore character, followed by "FLAG", followed by an underscore, and then followed by an alphanumeric character indicating the flag type. (Also see **Appendix D**, "Parameter Quality Codes".)

The "missing value" for a data value is always defined as -999, but written in the decimal place format of the parameter in question. For example, a missing salinity would be written -999.0000 or a missing phosphate -999.00. The value -999 is out of range for all WOCE-era parameters.

Table 1. General description of _hy1.csv file layout.

1st line	File type, here BOTTLE, followed by a comma and a DATE_TIME stamp YYYYMMDDdivINSwho YYYY 4 digit year MM 2 digit month DD 2 digit day div division of Institution INS Institution name who initials of responsible person example: 20000711CCHSIOSCD
#lines	A file may include 0-N optional lines, typically at the start of a data file, but after the file type line, each beginning with a "#" character and each ending with carriage return or end-of-line. Information relevant to file change/update history of the file itself may be included here, for example.
2nd line	Column headings. A list of column headings approved and used by the CCHDO is found in Table 2. A list of parameter headings approved and used by the CCHDO is found in Table 3. Data originators are urged, however, to be careful to supply their **correct** column headings rather than to simply copy 'approved' column headings into their files.
3rd line	Units. A list of parameter units used by the CCHDO is found in Tables 2 and 3. Data originators are urged, however, to be careful to supply their **correct** units rather than to simply copy the units used by the CCHDO.
data lines	As many data lines may be included in a single file as is convenient for the user, with the proviso that the number and order of parameters, parameter order, headings, units, and commas remain absolutely consistent throughout a single file. Thus a single data file may contain data lines for as little as one bottle from one cruise to as much as many bottles from many cruises.

note	Within a _hy1.csv file it is very strongly preferred that data from each station be contiguous, it is recommended that data from each cast at a station be contiguous, and it is preferred that the data from each cast be sorted from lowest pressure to highest pressure.
END_DATA	The line after the last data line must read END_DATA, and be followed by a carriage return or end of line.
other lines	Users may include any information they wish in 0-N optional lines at the end of a data file, after the END_DATA line.

Table 2. _hy1.csv header columns

Parameter	Format	Description notes
EXPOCODE	A14	The expedition code, assigned by the CCHDO or generated by the user. A single alphanumeric word, without spaces, commas, or "/" characters (but "_" underscore characters are OK) which is unique cruise identifier code. REQUIRED.

note		The convention the CCHDO uses to create the EXPOCODE is:
		`ExpoCode Syntax:    NODCShipCodeYearMonthDay`
		`Example:`
		`   Ship Name:          Roger Revelle`
		`   Cruise start date:  March 29, 2009`
		`   EXPOCODE:           33RR20090329`

SECT	A6	If a repeat of a WOCE section, this is the WHP section identifier. Optional.
STNNBR	A6	The originator's station number. This column is used for a single alphanumeric word, without spaces, commas, or "/" characters (but "_" underscore characters are OK) which is unique station identifier. Numeric-only STATION identifiers are preferred by many data users, but provision for alphanumeric identifiers is retained to maintain compatibility with WOCE records. REQUIRED.
CASTNO	I3	The originator's cast number. This column is used for a single integer cast number. Where cast number is unknown a default value of 1 is used or written in by the CCHDO. REQUIRED.
note		No "cast type" designator is used.
SAMPNO	A7	The sample number as described in this report, Section 5. It is very strongly recommended that at least one, preferably both, of the parameters SAMPNO and BTLNBR be reported for bottle data files. Where neither SAMPNO or BTLNBR are available, the CCHDO will add a SAMPNO column containing consecutive integers for each station/cast.

Parameter	Format	Description notes
BTLNBR	A7	The bottle identification number as described in this report, Section 5. It is very strongly suggested that at least one, preferably both, of the parameters SAMPNO and BTLNBR be reported for bottle data files. It is preferred that one of these, preferably BTLNBR, include a quality flag in the column immediately to its right. This is the primary index to a water sample. [Pressure - or depth - is a measured parameter. The pressure value can change during processing, and so pressure (or sample depth) should never be used to index water sample data.]
BTLNBR_FLAG_W	I1	The parameter name of a data quality flag should be identical to the actual parameter name, followed by "FLAG" and then by a character indicating the type of quality flag, with underscores between each word. W = WHP quality flag; I = IGOSS quality flag; U = quality flag from user-defined table.
DATE	I8	Cast date in YYYYMMDD integer format. REQUIRED
TIME	I4	Cast time (UT) as HHMM. Optional. Must have all four digits. The CCHDO prefers only one TIME value per cast, usually the time the rosette was at its deepest depth (i.e. when the first bottle is closed). Users who wish to record the time each bottle closes are urged to add a second time-related column, BTL_TIME, where the closure time for each bottle can be recorded.
LATITUDE	F8.4	Latitude as SDD.dddd where "S" is sign (blank or missing is positive), DD are degrees, and dddd are decimal degrees. Sign is positive in northern hemisphere, negative in southern hemisphere. Spaces to left of leftmost digit are ignored. Data with positions not reliable to ten-thousandths of a degree should be padded with meaningless zeros. The "BO" or "bottom" position (ship position when cast is at deepest level) should be used if available, with "BE" (ship position at cast start) or "EN" (ship position at cast end) used in that priority order when "BO" position is not available. REQUIRED The CCHDO prefers only one LATITUDE per cast, usually the ship's position when the rosette was at its deepest depth (i.e. when the first bottle is closed). Users who wish to record the position each bottle closes are urged to add a second latitude-related column, BTL_LAT, where the position at time of closure for each bottle can be recorded.

Parameter	Format	Description notes
LONGITUDE	F9.4	Longitude as SDDD.dddd where "S" is sign (blank or missing is positive), DDD are degrees, and dddd are decimal degrees. Sign is positive for "east" longitude, negative for "west" longitude. Spaces to left of leftmost digit are ignored. Data with positions not reliable to ten-thousandths of a degree should be padded with meaningless zeros. The "BO" or "bottom" position (ship position when cast is at deepest level) should be used if available, with "BE" (ship position at cast start) or "EN" (ship position at cast end) used in that priority order when "BO" position is not available. REQUIRED The CCHDO prefers only one LONGITUDE per cast, usually the ship's position when the rosette was at its deepest depth (i.e. when the first bottle is closed). Users who wish to record the position each bottle closes are urged to add a second longitude-related column, BTL_LONG, where the position at time of closure for each bottle can be recorded.
DEPTH	I5	Reported depth to bottom. Preferred units are "meters" and should be specified in Line 2. In general, corrected depths are preferred to uncorrected depths. Documentation accompanying data should include notes on methodology of correction. When no depth-to-bottom is supplied by the data originator for one or more rows of data in a _hy1.csv file which contains a "DEPTH" column, -999 may be written in by the CCHDO. Optional but strongly preferred.

Table 3a. WHP-Exchange bottle file parameter names, units, and comments.

Parameter	Format	Suggested Units	Comments
CTDPRS	F9.1	decibars	corrected CTD pressure (in a _hy1.csv file the value accompanying closure of the rosette bottle); normally no data quality flag is needed when reported in the bottle file; sometimes reported as CTDP or PRES
CTDTMP	F9.4	degrees C (specify ITS-90 or IPTS-68 if known)	corrected CTD temperature (in a _hy1.csv file the value accompanying closure of the rosette bottle); normally no data quality flag is needed when reported in the bottle file; sometimes reported as CTDT or TEMP
CTDSAL	F9.4		corrected CTD salinity (in a _hy1.csv file the value accompanying closure of the rosette bottle); sometimes reported as CTDS
CTDSAL_ FLAG_a	I1	a = W for WHP quality flags; a = I for IGOSS quality flag;	The parameter name of a data flag should be identical to the actual parameter name, followed by "FLAG" and then by a character indicating the type of quality flag, with underscores between each word.
		U = quality flag from user-defined table (table to be supplied in comment lines)	[A FLAG value can follow any data value, and should follow almost every data value. FLAG is shown here only for CTDSAL for simplicity. Typically a data file will have FLAG_W values following most parameters in this table except for CTDPRS and CTDTMP.]
SALNTY	F9.4		bottle salinity; sometimes reported as SALT
CTDOXY	F9.1	μmol/kg	corrected CTD oxygen (in a _hy1.csv file the value accompanying closure of the rosette bottle; may not be available in some _hy1.csv files); sometimes reported as CTDO or CTDO2
OXYGEN	F9.1	μmol/kg	bottle oxygen (must specify actual units, not simply copy the suggested units); sometimes reported as O2 or OXY
SILCAT	F9.2	μmol/kg	silicate (values in μmol/kg units are only 3% different than values in μmol/l units; so <u>one must specify actual units reported</u>, not simply copy the suggested units); sometimes reported as SIO3
NITRAT	F9.2	μmol/kg	nitrate (values in μmol/kg units are only 3% different than values in μmol/l units; so <u>one must specify actual units reported</u>, not simply copy the suggested units); sometimes reported as NO3

Parameter	Format	Suggested Units	Comments
NO2+NO3 (shown only if separate NITRAT and NITRIT are not available)	F9.2	μmol/kg	nitrate plus nitrite (values in μmol/kg units are only 3% different than values in μmol/l units; so one must specify actual units reported, not simply copy the suggested units) [Most modern techniques for determining dissolved nitrate return a value of nitrate (NO3) plus nitrite (NO2). A separate determination is then done for nitrite and the result subtracted by the data originator to obtain nitrate. If no separate nitrite determination was carried out - or in rare cases the nitrite number was not subtracted - data providers should list the result as NO2+NO3. Because nitrite values are in most regions small compared to nitrate, most data users will not adversely affect their results by relabeling NO2+NO3 as NITRAT.]
NITRIT	F9.2	μmol/kg	nitrite (see NO2+NO3) (values in μmol/kg units are only 3% different than values in μmol/l units; so one must specify actual units reported, not simply copy the suggested units); sometimes reported as NO2
PHSPHT	F9.2	μmol/kg	phosphate (values in μmol/kg units are only 3% different than values in μmol/l units; so one must specify actual units reported, not simply copy the suggested units); sometimes reported as PO4
CFC-11	F9.3	μmol/kg	sometimes reported as CFC11 or F11 (must specify actual units, not simply copy the suggested units)
CFC-12	F9.3	pmol/kg	sometimes reported as CFC12 or F12 (must specify actual units, not simply copy the suggested units)
CFC113	F9.3	pmol/kg	sometimes reported as CFC113 or F113 (must specify actual units, not simply copy the suggested units)
CCL4	F9.3	pmol/kg	carbon tetrachloride
TRITUM	F9.3	TU	tritium (must specify actual units)
HELIUM	F9.4	nmol/kg	dissolved helium
DELHE3	F9.2	%	
DELC14	F9.1	0/00	
DELC13	F9.1	0/00	
O18O16	F9.2	per mille	$\delta^{18}O$; oxygen isotope ratio
ALKALI	F9.1	μmol/kg	total alkalinity AT (sometimes reported as ALK or TALK)
TCARBN	F9.1	μmol/kg	total carbon (sometimes reported at TIC or DIC)

Parameter	Format	Suggested Units	Comments
PCO2	F9.1	μatm	partial pressure of CO2
PCO2_TMP		degrees C	PCO2 temperature, reported if PCO2 is reported
FCO2	F9.1	μatm	fugacity of CO2
FCO2_TMP		degrees C	FCO2 temperature, reported if FCO2 is reported
PH	F9.2		pH
PH_TMP		degrees C	PH temperature, reported if PH is reported
PH_SCALE	A3	"TS" or "SWS"	total scale (TS) or seawater scale (SWS), reported if PH is reported
DOC	F9.1	μmol/kg	dissolved organic carbon
DTN	F9.1	μmol/kg	dissolved total nitrogen

Table 3b. Other bottle parameters which have been submitted to the WHPO or CCHDO

Column Heading		**Units**		**Reporting Precision**	
Parameter	Mnemonic	Scientific	Mnemonic	Range	FORTRAN Format
39Argon	AR-39	% modern	PCTMOD	0,100	F9.1
Argon	ARGON	µmol/kg	UMOL/KG	5,25	F9.2
Abundance of bacteria	BACT	cells´108/kg	CELL/KG		
Barium	BARIUM				
Methane	CH4	nmol/kg	NMOL/KG	1,20	F9.2
Chlorophyll a	CHLORA	µg/kg	UG/KG	0,9	F9.2
Carbon monoxide	COMON	µmol/kg	UMOL/KG		
137Cesium	CS-137	dpm/100 kg	DM/.1MG	0,100	F9.2
Nitrogen (dissolved organic)	DON	µumol/kg	UMOL/KG	200,900	F9.1
Iodate	IODATE	nmol/kg	NMOL/KG	200,600	F9.3
Iodide	IODIDE	nmol/kg	NMOL/KG	0,300	F9.3
85Krypton	KR-85	dpm/1000 kg	DM/MG	0,5	F9.2
Nitrous oxide	N2O	nmol/kg	NMOL/KG	1,200	F9.2
Neon	NEON	nmol/kg	NMOL/KG	0,10	F9.3
Ammonium	NH4	µmol/kg	UMOL/KG		
Particulate organic carbon	POC	µg/kg	UG/KG		
Particulate organic nitrogen	PON	µg/kg	UG/KG		
Phaeophytin	PPHYTN	µg/kg	UG/KG	0,9	F9.2
226Radium	RA-226	dpm/100 kg	DM/.1MG	3,80	F9.2
228Radium	RA-228	dpm/100 kg	DM/.1MG	–1,10	F9.2
90Strontium	SR-90	dpm/100 kg	DM/.1MG	0,100	F9.2
Aluminum	ALUMIN				
Apparent Oxygen Utilization	AOU				
concentration of arabanose after hydrolysis	ARAB				
Calcium	CALCIUM				

Column Heading		Units		Reporting Precision	
Parameter	Mnemonic	Scientific	Mnemonic	Range	FORTRAN Format
Copper	CU				
Dissolved Combined Neutral Sugars	DCNS		(a20_2003)		
Dissolved Inorganic Nitrogen	DIN				
Biogenic sulfur compounds (DMS or DMSP)	DMS				
Fluorescence (total chlorophyll & phaeopigments)	FLUOR	mg/m3	MG/CUM	0,50	
concentration of fucose after hydrolyses	FUC				
concentration of galactose after hydrolysis	GAL				
concentration of glucose after hydrolysis	GLU				
Iodine 129	I-129				
concentration of Mannose after hydrolysis	MAN				
methyl chloroform	MCHFRM				
Nickel	NI				
pigmented picoeukaryotes	PEUK		(cell /L)		
Prochlorophytes	PRO				
concentration of rhamnose after hydrolysis	RHAM				
Sulfur Hexafluoride	SF6				
Synechococcus	SYN		(cell /L)		
Total Organic Carbon	TOC				
Transmissometer	XMISS	%light transmitted	%TRANS	0,100	

Also see

<http://cchdo.ucsd.edu/parameter_descriptions>

for a current list of CCHDO parameter names.

3. Format description for WHP-exchange CTD data (8-character suffix _ct1.csv for single CTD profiles and _ct1.zip for a zipped directory containing one or more _ct1.csv files)

[Note: To better understand this section please refer to document "example_ct1.csv" available from <http://cchdo.ucsd.edu/exchange/index.htm>. It is recommended that the reader examine "example_ct1.csv" in a text editor application in order to see all characters and also in a spreadsheet application in order to view overall layout.]

The overall layout of a_ct1.csv CTD data file is described in Table 4.

The first line of a WHP-exchange format file is a single word which describes the file type, in this case "CTD", followed by a comma and a date/time stamp.

The format next provides for 0-N optional information lines, each beginning with a "#" character, near the beginning of a _ct1.csv file. The CCHDO intends to use the "#" lines to hold file history information.

Next is a line indicating the number of header lines (counting the present line and those following), usually 10 in WHP CTD data in WHP-exchange format.

Next are the remaining 9 lines (usually) of header information. These mostly match the description of the similar information in a _hy1.csv file.

Next are the remaining 9 lines (usually) of header information. These mostly match the description of the similar information in a _hy1.csv file.

A line with "END_DATA" signals the end of the data lines.

After that line, a CTD data file may hold other file-specific documentation. The primary documentation for WHP data will, however, remain in the ".doc" file (or zipped directory).

General rules for WHP-exchange_ct1.csv data files:

Each line must end with a carriage return or end-of-line.

With the exception of the file type line, lines starting with a "#" character, the 10 header lines, or including and following a line which reads "END_DATA", each line in a_ct1.csv file must have exactly the same number of commas as do all other lines in that file.

The order of the parameters in the header lines in a _ct1.csv file should follow the order listed (and in "example_ct1.csv") to make it simplest for users to import files. All _ct1.csv files prepared by the CCHDO will adhere to the header parameter line order shown in "example_ct1.csv". Still, CTD data users are urged to use "read" statements that are sensitive to parameter names rather than position of the parameter in the data files.

It is not necessary that a CTD data file contain a column for CTD oxygen probe measurements (CTDOXY) when there are no CTD oxygen probe data.

If other parameters are included in the CTD data stream, they, and their quality flags, can be included in the _ct1.csv data file, following the overall protocols.

When a quality flag is available for a CTD parameter, that quality flag shall be placed in the column immediately to the right of the parameter.

The name of a quality flag always begins with the name of the parameter with which it is associated, followed by an underscore character, followed by "FLAG", followed by an underscore, and then followed by an alphanumeric character indicating the flag type. (Also see **Appendix D**, "Parameter Quality Codes".)

The "missing value" for a data value is always defined as -999, but written in the decimal place format of the parameter in question. For example, a missing salinity would be written -999.0000.

The value -999 was chosen because it is out of range for all WHP parameters.

Each data parameter listed in Table 5 - except for all flags, which are "I1" - will be listed in "F9.x" floating point format, where "x" indicates the number of decimal places. For each parameter, the CCHDO will pad with meaningless zeros data received with fewer decimal places and round data received with extra decimal places to the number of decimal places specified in Table 5.

Table 4. General description of _ct1.csv file layout.

1st line	File type, here CTD, followed by a comma and a DATE_TIME stamp YYYYMMDDdivINSwho YYYY 4 digit year MM 2 digit month DD 2 digit day div division of Institution INS Institution name who initials of responsible person example: 20000711WHPSIOSCD ORIGINAL_DEPTH_HEADER=
#lines	A file may include 0-N optional lines at the start of a data file, each beginning with a "#" character and each ending with carriage return or end-of-line. Information relevant to file change/update history may be included here, for example.
2nd line	NUMBER_HEADERS = n (n = 10 in this table and the example_ct1.csv file.)
3rd line	EXPOCODE = [expocode] (see Table 2 for definition)
4th line	SECT = [section] (see Table 2 for definition)
5th line	STNNBR = [station] (see Table 2 for definition)
6th line	CASTNO = [cast] (see Table 2 for definition)
7th line	DATE = [date] (see Table 2 for definition)
8th line	TIME = [time] (see Table 2 for definition)
9th line	LATITUDE = [latitude] (see Table 2 for definition)
10th line	LONGITUDE = [longitude] (see Table 2 for definition)
11th line	DEPTH = [bottom] (see Table 2 for definition)
next lines	Parameter headings. A list of CTD parameter headings approved and used by the CCHDO is found in Table 5. Data originators are urged, however, to be careful to supply their **correct** column headings rather than to simply copy 'approved' column headings into their files.
next lines	Units. A list of parameter units used by the CCHDO is found in Table 5. Data originators are urged, however, to be careful to supply their **correct** units rather than to simply copy the units used by the WHP.
data lines	A single _ct1.csv CTD data file will normally contain data lines for one CTD cast. Generally these will be what is called a "2 decibar" file, i.e. there will be a 2-decibar interval between data lines, and each line will lay at either even or odd whole decibars. Other pressure intervals are accepted; for example, the CCHDO has many CTDO data reported at 1-decibar pressure intervals.
END_DATA	The line after the last data line must read END_DATA, and be followed by a carriage return or end of line.
other lines	Users may include any information they wish in 0-N optional lines at the end of a data file, after the END_DATA line.

Table 5. _ct1.csv common parameter names, units, and comments.

Parameter	Format	Suggested Units	Comments
CTDPRS	F9.1	decibars	corrected CTD pressure; sometimes reported as CTDP or PRES
PARAMETER_NAME_FLAG_a	I1	W = WHP quality flag. I = IGOSS quality flag.	The parameter name of a data flag should be identical to the actual parameter name, followed by "FLAG" and then by a character indicating the type of quality flag, with underscores between each word.
		[U = quality flag from user-defined table]	[A FLAG value can follow any data value in this table. FLAG is shown here only for CTDPRS for simplicity. Typically a WHP data file will have FLAG_W values following every parameter in this table.]
CTDTMP	F9.4	degrees C (specify ITS-90 or IPTS-68 if known)	corrected CTD temperature; sometimes reported as CTDT or TEMP
CTDSAL	F9.4		corrected CTD salinity; sometimes reported as CTDS
CTDOXY	F9.1	μmol/kg	corrected CTD oxygen (must specify actual units, not simply copy the suggested units); sometimes reported as CTDO or CTDO2
Other parameters embedded in the CTD data stream, and their associated quality flags, may be included in the _ct1.csv data file following the general protocols listed here. Some examples:			
CTDNOBS	F9.0 (or I2 or I3)		number of CTD observations (scans) averaged for the CTD data reported at this pressure interval (no quality flag needed); sometimes reported as NUMOBS
TRANSM	F9.3	volts DC	transmissometer voltage; sometimes reported as XMISS or TRANS
FLUORM	F9.3	volts DC	fluorometer voltage; sometimes reported as FLUOR

Appendix D. Parameter Quality Codes (Quality Flags)

Below we discuss "WOCE" quality codes (quality flags) for sample bottles, water samples, and CTD data. We end with a discussion of IGOSS quality flags, including translation of "WOCE" to "IGOSS" quality codes.

1. Sample bottle quality codes

The bottles on a rosette water sampler can leak, the control mechanism may fail to release the lanyard, or there can be other problems with the water bottles. It is therefore recommended that each sampling bottle on a cast be accompanied by a quality code as defined in Table D.1. (The CCHDO does not, however, *require* that data providers include bottle quality codes.)

TABLE D.1: "WOCE" quality code definitions for water bottles.

Flag Value	Definition
1	Bottle information unavailable.
2	No problems noted.
3	Leaking.
4	Did not trip correctly.
5	Not reported.
(6)	(Significant discrepancy in measured values between Gerard and Niskin bottles.)
(7)	(Unknown problem.)
(8)	(Pair did not trip correctly. Note that the Niskin bottle can trip at an unplanned depth while the Gerard trips correctly and vice versa.)
9	Samples not drawn from this bottle.

Use of code 1 is generally limited to cruises where bottle information is not available. BTLNBR is sometimes set equal –9 in older data sets. Present-day cruises should use code 5 if bottle information is not reported.

Flags 6, 7, and 8 apply primarily to large volume samplers, which are not currently in use.

Note: It is critical that questionable bottles (**especially leaking bottles**) be flagged at the earliest possible time.

2. Water sample (measured parameter) quality codes

Each water sample measurement should be accompanied by a data quality code. (The CCHDO does not, however, *require* that data providers include parameter quality codes.) Water sample quality code definitions are given in Table D.2.

TABLE D.2: "WOCE" quality code definitions for water sample measurements.

Flag Value	Definition
1	Sample for this measurement was drawn from water bottle but analysis not received. *Note that if water is drawn for any measurement from a water bottle, the quality code for that parameter should be set equal to 1 initially to help ensure that all water samples are accounted for.*
2	Acceptable measurement.
3	Questionable measurement.
4	Bad measurement.
5	Not reported.
6	Mean of replicate measurements (Number of replicates should be specified in the .DOC file and the replicate data tabulated there).
7	Manual chromatographic peak measurement.
8	Irregular digital chromatographic peak integration.
9	Sample not drawn for this measurement from this bottle.

The definitions in this table apply to quality codes in a bottle data file, but not to the CTD (***CTDSAL*** or ***CTDOXY***) parameters or the bottle number (***BTLNBR***) in that file. See the separate tables for the bottle quality code and CTD quality codes.

If water is drawn for any quality-coded measurement from a bottle, the CCHDO recommends that the data team at sea set the quality code for that parameter equal 1 initially, next to the otherwise empty data column, to ensure that all water samples are accounted for later when the data are received and merged. If the parameter is not sampled on a given station, cast, or level the quality code for that parameter is instead set to 9.

All measured values should be reported, including bad values, in data files which contain quality codes. In other words, questionable or bad values due to sampling, analytical or other problems are coded appropriately, but not removed from the data file. Whenever data were expected to be measured from a water sample drawn from a bottle (quality flag = 1), but the observation is missing due to sample loss, contamination, etc., the numerical "missing value" (e.g., -999) is placed in the measurement field in the data file and the respective quality code is reset to 5.

It is not possible to define what is meant by an “acceptable” measurement (quality code = 2) for all cruises or even all measurements from the same bottle. What may be a questionable, or even bad, measurement on a one cruise may be quite acceptable on another cruise. Water from the same bottle may be quite adequate for one parameter, for example, salinity, but badly contaminated for another, for example, CFCs. Also, investigators should be certain that their quality code assignments for their water samples are consistent with the quality code for the water bottle itself.

3. CTDO data quality codes

The CTDO quality codes are defined in Table D.3. Each measured CTDO parameter may have one quality code associated with it. CTDO data quality codes are optional but recommended.

A CTDO quality code of 1, *not calibrated*, applies to salinity and oxygen measurements only when water samples are collected from the present cast, or a nearby cast, but corrections have not yet been applied to the CTD data. For pressure and temperature, a quality code of 1 would indicate final CTD calibrations have not been applied.

TABLE D.3: "WOCE" Quality code definitions for CTD data.

Flag Value	Definition
1	Not calibrated.
2	Acceptable measurement.
3	Questionable measurement.
4	Bad measurement.
5	Not reported.
6	Interpolated over a pressure interval larger than 2 dbar.
7	Despiked.
(8)	(Not used for CTD data.)
9	Not sampled.

4. IGOSS Quality Codes

It may be advantageous for some users to translate the WOCE quality codes into the more widely recognized IGOSS quality codes. The table below list the translation recommended by the CCHDO.

The WMO IGOSS observation quality codes are:

0	No quality control yet assigned to this element
1	The element appears to be correct
2	The element is probably good
3	The element is probably bad
4	The element appears erroneous
5	The element has been changed
6 to 8	Reserved for future use
9	The element is missing

A perfect translation is probably not feasible, but we suggest the following WHP-to-IGOSS (not IGOSS-to-WHP) translation rules as reasonable:

	WOCE	**IGOSS**
bottle		
	1	0
	2	1
	3	3 (see note #1)
	4	4
	5	0
	6	4
	7	4
	8	4
	9	9
water sample		
	1	0
	2	1
	3	2 (see note #2)
	4	4
	5	0
	6	2
	7	2
	8	2
	9	9
ctd		
	1	0
	2	1
	3	2 (see note #2)
	4	4
	5	0
	6	2
	7	2
	9	9

Note #1: The CCHDO, in the interest of being conservative, has chosen to translate the WOCE bottle quality code 3 into IGOSS quality code 3. A leaking water sample bottle typically results in a discrepancy or error in gas samples, such as oxygen and CFCs, but less often results in data discrepancies for salinity and nutrients. It is suggested that data users who wish to import only "good" data not import any water sample data from bottles with a WOCE code 3 or IGOSS code 3. A data user who is willing to entertain slightly greater risk might choose to import non-gas sample data (e.g., salinity and nutrients) from a WOCE code 3 or IGOSS code 3 water sample bottle, and allow import of gas sample data (e.g. oxygens and CFCs) for bottles with IGOSS Code 2. (The CCHDO is not, however, currently assigning IGOSS code 2 to water sample bottles; but future data originators or data centers may wish to use code 2.)

Note #2:	The CCHDO has noted that in general, data originators tend to be conservative and so some WHP-code-3 ("questionable") water sample parameter data may be deemed WHP-code-2 ("good") by a data user. The IGOSS code 2 ("probably good") seems to be a reasonable interpretation. The CCHDO is not currently assigning IGOSS code 3 ("probably bad") to WHP water sample data values.