



# **BODC requirements for on-board CTD and underway data processing for UK SOLAS cruises.**

Version 1.5 – updated prior to D326

Author: Gwen Moncoiffé, BODC

Last updated 04JAN2008: clarification of O<sub>2</sub> concentration requirements during DERIVE plus change to scan range setting following comments from Dave Teare (NMF-SS) during D325.

Previous versions:

Version 1.3 "SOLAS\_Cruise\_Data\_Management\_UKORS\_NOV2006";

Version 1.2 "SOLAS\_Cruise\_Data\_Management\_UKORS\_OCT2006", 02-Nov-2006;

Version 1.1 "SOLAS\_Cruise\_Data\_Management\_UKORS\_OCT2006";

Version 1.0 "Cruise\_Data\_Management\_PO332.doc", 23-Jan-2006.

Intended audience: UK SOLAS PSOs, NMF cruise managers and technical personnel.

## **1. Introduction**

This document outlines BODC's minimum requirements for on-board data processing operation and post-cruise data submission by UKORS (now NMF).

Although it is specifically written for the UK SOLAS cruises, most are standard BODC requirements. Some however, such as those for on-board CTD data processing, are more prescriptive due to the fact that BODC is responsible for producing the final, quality controlled version of these data.

This document is designed to be practical and should provide a helpful reference for NMF technical staff and cruise PSOs, before and during the cruise, or when preparing the data for submission to BODC at the end of the cruise. Many of the notes and reminders are based on personal experience, and on BODC's expertise in quality control and banking of CTD and underway data originating from a wide range of projects, cruises and ships.

Please note that this document is only intended for CTD and underway data. Data files and information from other shipborne instrumentation under the responsibility of National Marine Facilities staff, such as ADCP, wave recorders, etc., should be submitted to BODC for archiving according to the usual procedure.

This document is a work in progress, and suggestions for improvement/clarification are welcome. Please forward any comments and questions to Gwen Moncoiffé at [gmon@bodc.ac.uk](mailto:gmon@bodc.ac.uk) or direct phone line 0151 7954880.

## **2. CTD Processing**

The bulk of SeaBird CTD data processing will be carried out on board by NMF technical staff using SeaBird processing software. The processing will need to produce ASCII versions of the data at full 24Hz, and at reduced 2Hz resolution. These must be sent to BODC immediately after the cruise along with:

- all the original raw binary data files (.DAT, .CON, .BL, and .HDR files);



- the ROS and BTL files;
- **all** associated documentation and log files.

Once received, the data are logged and archived on BODC's archiving system in their original format. For UK SOLAS, the 2Hz version of the data is then selected for further processing. The work involved will consist of:

1. reformatting the data received to BODC's internal format;
2. screening the data visually and apply quality control flags when necessary;
3. calibrating the salinity and fluorometer channels against water samples;
4. producing final calibrated, quality controlled and calibrated downcast profiles binned to a standard 1 db.

In order to do this, BODC has a number of requirements which must be adhered to. The completeness and accuracy of the data and information sent to us after the cruise are **critical** in helping us bank and distribute good quality data in a quick and efficient way.

## 2.1. CON files

Special attention needs to be given to the quality and accuracy of the CON files and in particular:

- **ALL sensors** logged by the SeaBird software should be selected for output. Channels, such as secondary conductivity or temperature, nephelometer/LBSS, PAR, etc., should not be left out of the on-board data processing. Failure to do this has often resulted in BODC having to reprocess an entire cruise worth of raw CTD files in order to access data from these sensors;
- The CON files should be checked at the beginning of the cruise to ensure that, for each sensor logged by the SeaBird software, the following is true:
  - the sensor's serial numbers are correct and correspond to the sensors installed on the CTD package;
  - the manufacturer's calibration coefficients are correct and up-to-date;
  - the transmissometer and light meter calibration coefficients are correctly derived. In past cruises, calibration coefficient errors have often been identified as the main reason for unrealistic attenuation values and also for PAR values too low by a factor of 10.
- If a sensor is changed or moved during the cruise then the CON file needs to be updated and NMF personnel need to ensure that the CNV version of the DAT file provided to BODC has been generated using the correct CON file;
- All changes in package configuration, including sensor's replacement, swap over, etc., should be listed in a report along with the date, CTD cast to which the change applies to, and reason for changing the sensor; the report should be included in the cruise report;
- The CON file used for converting a particular cast from its binary DAT version to its ASCII CNV version should always bear the same name as the DAT and CNV files (e.g. CTD001.DAT, CTD001.CON, CTD001.CNV);
- If a NMEA connection is present then NMEA input checkbox should be selected.



## 2.2. Data conversion (DATCNV)

- The oxygen concentration channel needs to be output as ml/l or mg/l (NOT in umol per kg).
- Both downcast and upcast need to be output in the same CNV file.
- Output should be at full 24 Hz resolution.
- Other processing steps should be carried out as per NMF/BODC recommended procedures (see Appendix A).

## 2.3. Bottle files

- All channels should be selected (i.e. each sensor on the CTD package should have a data channel in the bottle file).
- Output min, max and averaged variables.
- Bottle rosette position should be used to reference the bottles (if Bottle rosette position number is different from the Niskin bottle number typically used by the scientists to reference their sample, then a mapping should be provided to BODC).
- Date and time of bottle firing should be displayed.
- The handwritten log should clearly indicate any problem with bottle firing operations: leak, misfire, firing out of order, etc.



## 2.4. Summary of BODC's minimum requirements for CTD data

Number	Description	
1.	Copy of log sheets.	
2.	Lists of ALL sensors used on CTD package and configuration of instruments, including ancillary sensors and which data channels they log to in output files.	
3.	Hard or digital copy of technical sheets and calibration certificates for each sensor used.	
4.	List of changes made during the cruise, including cast numbers to which changes apply.	
5.	Information about rosette configuration: a. number, type, and capacity of bottles, b. distance between pressure sensor and top of sampling bottles and distance between pressure sensor and bottom of bottles, c. distance from pressure sensor to reversing thermometer mount (if there is one).	
6.	Copy of the raw and processed data files directly after the cruise (please read BODC's CTD Processing requirements).	
7.	Transmissometer pathlength and air readings - preferably regular readings throughout the cruise.	
8.	Electronic copy of bottle salinity for calibration of CTD conductivity sensors referenced to CTD cast number and bottle depth (with CTD data or as soon as possible after the cruise).	
9.	Report on any instrument malfunction, problems with data logging, bottle miss-fires, changes of sensors, cleaning operations including cast numbers to which they apply.	



### 3. Underway system data/information

At this point in time, I am not aware that the new acquisition system for underway navigation, met and hydrographic data is yet fully operational on NERC ships. If it is then we will require full details of the new acquisition system and also details of the files provided to us at the time of the data submission.

In any case, it is important that NMF-SS documentation provide us with pointers to the best quality data available, and details of what has been done to the data during on-board processing.

Our minimum requirements are as follows:

Number	Description
1.	ASCII copy of data files from ship's logging system, including: <ul style="list-style-type: none"> <li>• Copy of all worked up data (best navigation, 30 second data if available)</li> <li>• Definition and units for each channel used – a clear definition of data channels is required as column headings are sometimes inconsistent or too vague.</li> </ul>
2.	Location of all met and surface hydrography sensors, including: <ul style="list-style-type: none"> <li>• anemometer position, height above sea level and vane orientation relative to the ship's bow.</li> <li>• other met instruments location and height above sea level.</li> <li>• depth (below the water surface) of inlet tube to pumped system.</li> <li>• plumbing location of surface underway sensors (TSG, fluorometer, etc...)</li> </ul>
3.	Copies of technical sheets for each sensor used with latest calibration information.
4.	Summary report on any event which could have an effect on the data: instrument malfunction, problems with data logging, changes of sensors, cleaning operations, non-toxic switch on/off operation – logged against time in GMT.
5.	Electronic copy of bench salinometer data for calibration of the TSG salinity channel (as soon as available)
6.	Transmissometer air readings.

**IMPORTANT: Best quality meteorological data is CRUCIAL to the scientific work being carried out on all UK-SOLAS cruises. It is therefore imperative that all meteorological instruments under NMF responsibility are properly maintained and remain in good working order throughout the cruises.**



#### 4. Digital data submission

- All required CTD and underway digital material should be submitted to us on DVDs or CDs **as soon as the cruise has ended**. This could be identical to the copy handed over to the PSO providing all the data we require are included. In the event of the cruise ending in a port outside the UK, we still expect to receive a copy of the data within 3 weeks of the cruise ending.
- If more convenient and providing the cruise ends in the UK, BODC can arrange to have the data collected when the ship docks. If the cruise ends outside the UK, then the DVDs and the various hardcopies of logs etc, should be posted to the following address:
  - Drs Jenny Andrew or Gwen Moncoiffé
  - British Oceanographic Data Centre
  - Joseph Proudman Building
  - 6 Brownlow Street
  - Liverpool
  - L3 5DA
  - UK
- The submission of the UK SOLAS CTD and underway data **immediately** after the cruise is specific to UK SOLAS and **is not a substitute** for the newly introduced NMF/BODC procedure of sending cruise data backup on DLT tapes to BODC after every NERC cruises. Transfer of the cruise backup to BODC's current NMF/SOC liaison officer Malcolm Hearn must still take place as soon as convenient after the cruise.
- The directory structure should be clear and the directories' names should be as self-explanatory as possible.
- A readme file should preferably be included in each sub-directory to list the directory content, briefly explain what the files are or clarify any unusual configuration. For example, if multiple versions of the same data are provided then a text file should explain what the different versions are, and the main or final version of the data should be clearly identified.
- For CTD data, test processing, if carried out, should be clearly identified and accompanied by plain text explanations.



## Appendix A: Recommended steps for on-board basic processing of SBE-911 CTD data.

The basic data processing procedure should include the following routines: DATCNV, BottleSum, AlignCTD, CellTM, Derive and BINA VG.

### A.1. Data conversion (DATCNV)

This step converts the raw files to ASCII and apply calibrations as appropriate through .CON file. Our recommended procedures are as follows:

- The CON file should be checked for accuracy and updated as required for ALL the sensors present on the CTD package – extra caution should be taken with the records for PAR and transmissometer to ensure that voltage conversion coefficients are correct and match the manufacturers' latest calibration and SeaBird latest recommendations.
- All sensors' output should be converted to engineering units as recommended by the manufacturer.
- Output at full 24 Hz resolution.
- Other recommended settings are as follows:

**Process scans to end of file:** yes  
**Scans to skip over:** 0  
**Output format:** ASCII  
**Convert data from:** Upcast and downcast  
**Create file types:** both data and bottle file  
**Source of scan range data:** Bottle log .BL file  
**Scan range offset:** -2.5 seconds (changed from 0 second as advised by Dave Teare, NMF-SS during D325 cruise)  
**Scan range duration:** 5 seconds

*The setting for this should be discussed with the CTD technicians and the choice made should be documented in the cruise report. On some cruises, 5 seconds was recommended for standard casts based on the fact that the CTD package is unlikely to move on again within 5 seconds of bottle firing. For mesoscale survey casts when bottles are fired "on the fly", 1.5 seconds might be more appropriate as 1.5 seconds corresponds to roughly 1m travel.*

**Merge separate header file:** No

**Select Output Variables:**

*ALL sensors logged by the SeaBird software should be selected as variables for output even if not specifically requested by the PSO. For a standard package, selected variables would include for example:*

Time elapsed in seconds,  
Pressure in db,  
Primary (and secondary if present) temperature ITS-90 in deg C,  
Primary (and secondary if present) conductivity S/m,  
Primary (and secondary if present) salinity PSU (see note below),  
Oxygen SBE 43 (if appropriate to the sensor used) in voltage  
Beam attenuation 1/m

*Plus channels for any other auxiliary sensors fitted on the package: e.g. fluorometer, LBSS, upwelling and downwelling PAR.*



If the **voltage channels** are used then a mapping between voltage channels and the sensors from which the data come from should be provided in a separate file and also included in the cruise report. If the configuration was changed during the cruise then this should also be indicated alongside the casts identifiers to which the configuration applies.

## A.2. BottleSum

a) It is important to always use **Bottle rosette position** to reference the bottles. For this reason, one needs to ensure that the .ROS files created during DATCNV are in the same directory as the raw .BL files when running Bottle Summary. If this is not the case then, by default, the SeaBird processing software will output Bottle firing sequence numbers instead of Bottle rosette position numbers.

### b) Data Setup

**Output min/max values for averaged variables:**           yes

**Select Averaged Variables:**

All variables except time elapsed (and cell counts if present) need to be selected.

**Select Derived Variables:**   Oxygen, SBE 43 in ml/l

(NOTE: oxygen MUST be in ml/l; if scientists require output in  $\mu\text{mol}$  per kg then this can be added to the output but it should not replace oxygen in ml/l).

Salinity does not need to be selected as a derived variable at this stage as it has already been selected as part of the Averaged Variables. Other derived variables (i.e. Density, etc) may be added if requested by the scientists.

c) If not automatically selected then date and time of bottle firing stamp selected for output.

## A.3. ALIGNCTD

Align CTD aligns parameter in time, relative to pressure. This ensures that calculations of salinity, dissolved oxygen concentration, and other parameters are made using measurements from the same parcel of water. For the SBE911 package, alignCTD is mainly used to align oxygen measurements relative to pressure and temperature.

Based on past observations, the optimum oxygen alignCTD advance value will vary depending upon the package used (e.g. whether the 12-way and 24-way bottle frames are used), difference in the CTD unit orientation (i.e. vertical or horizontal), number of sensors on the frame and subsequent effect on entraining water, length of the tubing, etc. It is therefore strongly recommended that the CTD operator and data processing person on board try different lag times in order to find the best fit for the specific package.

According to Jeff Benson, the SBE recommendations (+2 to +5 seconds for pumped systems and +1 to +5 seconds for non-pumped systems) seem to work best for vertical mounted CTDs, whereas 6 to 8 seconds are more typical with horizontal mounted underwater unit.

## A.3. CellTM

Cell Thermal Mass filters conductivity cell thermal mass effects from the measured conductivity. The recommended SeaBird settings should be applied to both primary and secondary sensors.



For SBE9plus with TC duct and 3000 rpm pump the recommendation is as follows:

alpha=0.03  
1/beta=7

#### A.4. Derive

Note that if the data are noisy then WILDEDIT and FILTER (see paragraph A.5.) will need to be used prior to running DERIVE and BINAVERAGE.

Otherwise, once all the adjustments have been made to the data, oxygen concentration and salinity can be calculated and added to the ASCII file by running *DERIVE*.

Although salinity was output in the initial running of *DATCNV*, adjustments to conductivity through *ALIGN* and *CELLTM* need to be taken into account, so salinity needs to be derived from the adjusted conductivity.

#### Derived Variables to be selected are:

Salinity from both primary and secondary sensors if present;  
Oxygen concentrations in ml/l: must always include the actual oxygen concentration (currently labeled oxygen, SBE43 in ml/l). Oxygen saturation in ml/l can also be selected if requested by the scientists.

Note that if it is absolutely necessary to output oxygen concentrations in  $\mu\text{mol per kg}$  then the channel can be added to the output but it **MUST NOT** replace oxygen in ml/l.

Other derived variables (potential temperature, sigma-t, oxygen percent saturation) are not required by BODC because they will be generated from the fully quality controlled data; however, they could be selected if requested by PSO and scientists on board.

#### A.5. BINAVERAGE

BINAVERAGE needs to be run to average the data to 2Hz. This version needs to be saved under a new name (appending *\_2Hz* to the filename for example) so that the 24Hz version of the data is not overwritten.

#### Data Setup:

<b>Bin type:</b>	Time, seconds
<b>Bin size:</b>	0.5
<b>Include number of scans per bin:</b>	no
<b>Exclude scans mark bad:</b>	yes
<b>Scans to skip over:</b>	0
<b>Cast to process:</b>	upcast and downcast

The reduced version of the data should be submitted to BODC alongside the 24Hz version. BODC will archive both versions but use the 2Hz CNV version of the data for banking into its database unless a better fully quality controlled and calibrated version of the data is received from the scientists responsible for CTD data processing.



## A.6. STRIP

This step may be used to remove redundant salinity and oxygen voltage channel generated during DATCNV. In order to do this, in Data Setup>Select included variables, just unselect the first primary and secondary salinity channels from the list and the Oxygen voltage channels.

## A.5. Additional steps depending on data quality

The steps described above are the minimum recommended steps for basic on-board CTD processing. However, in case of noisy data, the SeaBird sub-routines WILDEDIT will need to be used.

BODC's recommendation is to run WILDEDIT on the pressure channel only to remove pressure spikes. Also it is worth noting that WILDEDIT should always be run before FILTER as pressure spikes can cause FILTER to smooth the data incorrectly.

The typical settings are as follows:

<b>Standard deviation for pass one:</b>	2
<b>Standard deviation for pass two:</b>	20
<b>Scans/block:</b>	100
<b>Keep data within this distance of the mean:</b>	0
<b>Exclude scans marked bad:</b>	yes

WILDEDIT might need to be run more than once, with different block sizes and number of standard deviations if the data is particularly corrupted. It is worth writing the output to a different file name at this stage to avoid overwriting the input file.

## 5. Additional processing at the discretion of PSO and collaborators

Other edits such as running WILDEDIT on other channels if these are particularly noisy or using FILTER or LOOPEDIT to produce cleaner data, could be added to the processing steps prior to using DERIVE and BINAVERAGE; however this should be left to the discretion of the PSO and of the personnel in charge of data processing on board.

For DATCNV: if binary output is selected during DATCNV then use TRANSLATE after BOTTLESUM and before DERIVE.

End of Document.

Gwen Moncoiffé, UK-SOLAS Data Management Project Co-ordinator.

email: [gmon@bodc.ac.uk](mailto:gmon@bodc.ac.uk)

Direct Line: 0151 795 4880

Fax: 0151 795 4912