

Summary of information required from individual scientists when submitting data.

When data analysis has been completed, please send a copy of your data set to Gwen Moncoiffé or Jenny Andrew, the UK SOLAS Data Management team at BODC by email (gmon@bodc.ac.uk) or on CD-ROM or DVD by post to the following address:

ATTN Drs Gwen Moncoiffé or Jenny Andrew
British Oceanographic Data Centre
Joseph Proudman Building
6 Brownlow Street
Liverpool
L3 5DA

Unless alternative time-scales have been agreed with the UK SOLAS data management team, this process is expected to take place within 6 months after the end of the fieldwork activity. Please refer to the UK SOLAS data management plan and policy available from http://www.bodc.ac.uk/projects/uk/uksolas/data_policy/ for further details.

BODC will accept most data, text and image/video file formats providing they contain or are accompanied with sufficient information to enable their use by BODC's staff.

Please ask if you are unsure whether we can deal with your file format.

- For small numerical data sets or metadata tables and sampling logs, spreadsheets are the most commonly used. If at all possible all stations or all related data should be presented in a single tabulated layout in one single sheet. Avoid submitting multiple identical Excel sheets or files for small discrete datasets as these can be very time consuming to reformat. For example do not use a separate sheet or file for each individual station sampled unless a different table layout is required (e.g. different type of analysis was made or a different sampler was used).
- For large numerical data sets such as, for example, those from continuous vertical or surface measurements, NetCDF or ASCII formats are the preferred choice (including for example CSV). In these cases, one file per station is acceptable. Unless there is clear indication of the contrary, we will assume that the format is exactly the same for the entire data series.
- For video and still camera images, we will archive the data in their original digital format or, if scanned, in a format which offers the most suitable resolution for the type of images being captured.

From all scientists who will submit data to BODC, the following standards should be followed:

- Use clear column headings.
- Provide the units for **all** parameters and numerical values.
- Always ensure that parameters are correctly and accurately defined and labelled. If used, acronyms and non-standard nomenclature should be clearly defined in plain language.

- All measured parameters should be included. If some parameters were derived then indicate which parameters are directly measured and which are derived from a combination of measurements. For derived measurements, please include the formulae used to obtain the value including full reference to published material if applicable.
- Provide quality assessment and quality control comments and information if appropriate. Indicate data points that you consider suspect. Once the data are transferred to BODC's working format, we have means to associate quality control flags alongside data values to inform data users that the value was considered suspect by the data originator. Similarly, if there were problems with the data set, please indicate this. We will include any comments in the data documentation which will accompany the data.
- Document methods and protocols used - in Word or text document plus details of any instruments used, including full specification for non-UKORS equipment, make, model, setup, manufacturer's and field calibration information, detection limits, instrument's accuracy and precision, type and pore size of filter used if applicable, instrument location and height/depth, sampling and sub-sampling procedures and frequencies, analytical methods, post-cruise data and sample processing. Information on instruments and methods should be to the same level of detail as what is required for a refereed publication.
- If any reference is made to a publication then the FULL reference should be provided including full list of authors, year of publication, title, journal, issue, page numbers.
- Provide the name and institution of the scientist who should be considered the principal "owner" of the data submitted. If not provided, then default ownership will be set to the project PI. At present, we can only link one person's name to the data in the database. However, if other people need to be acknowledged each time the data are used then please provide the full list of authors: we will include this information in the plain text data series documentation which is linked to the data.

From all scientists who will submit data to BODC, the following information noted during the cruise **is required** and should be included in all data and metadata files or in the accompanying documentation:

- Sampling mechanism (e.g. CTD rosette, air sampler, net, non-toxic sample, bucket on a rope, camera on buoy, vessel mounted camera, CLASP on buoy, vessel mounted CLASP, etc...).
- Cast/station/sample/image unique identifier(s).
- Date and time of sampling event specifying clearly whether time is GMT/UTC or local time. If local time is used then time zone must be specified. The time is very important as it is used to fix the position of the sample against the primary navigation data record. If sampling was not instantaneous then start and end time should be provided.
- Depth or height of sample or sampling instrument.

See Appendix 1 below for a summary of **additional** information required according to the type of sample/measurement collected.

Appendix 1: Examples of information required for specific type of sample collection (in addition to the general requirements listed in the text).

Data type	Information required
Water bottles from CTD rosette	CTD cast reference (as used by UKORS). Depth of sample. NISKIN bottle reference. Sampling and subsampling methods including type of filter used (if applicable).
Other hydrocast sampling (e.g. GOFLO etc.)	Hydrocast identifier . Type of water sampler and bottle used including volume. Date and accurate time of deployment. Depth of sample. Problem report (leaks, etc.). Sampling and subsampling methods as above.
Incubations/ laboratory experiments	Source of water samples with cast, station and sample identifiers (including depth of collection). Pre-incubation sample treatment (e.g. sieving: provide mesh size). Tracer used or treatment applied. Incubation/experiment identifier. Incubation start time (GMT) and duration. Subsampling times for time-course experiments. Incubation location (e.g. deck, wet lab, CT lab, in situ). Light condition: light level and regime (e.g. dark, simulated in situ with % surface light illuminating sample, etc...). Simulated depth if applicable. Experimental temperature conditions. Other experimental conditions (inhibitors, nutrient additions, time course subsampling, etc.).
Air samples	Type of sampler. Sample identifier. Sampling date and time in GMT. Date and time sampling ended or sampling duration if not instantaneous. Height of sampling with reference (e.g. sea level, ship's main deck). Information about filter used.
Stand Alone Pumps	Pump system type, make, model and settings. Deployment/sample identifier. Date and start and end time of deployment. Depth of sample. Volume of water pumped. Filter used.
Microlayer samplers	Make/model and description of sampler. Deployment method. Sampling position if deployed away from mother ship. Sampling date and time. Mesh size or any other variable characteristics. Depth or depth range of individual samples if known. Sample collection/recovery technique.
Other samples/ instruments	Make, model and description of instrument. Instruments location on the ship, height/depth of measurements. Calibration procedures (pre-, during or post-cruise). Definition and units of any data logged. Any processing carried out on the data channels before sending to BODC. Any further processing required – please discuss with BODC as we may be able to assist.



As a general rule, as much information as possible should be provided in the cruise report about individual sampling strategies and methods, indicating whether the work was integrated with other cruise participants or not, the number of stations sampled and the number of samples collected, details of what the samples were or will be used for including: measurements and processing made on board, description of post-cruise processing steps, and composition of the final dataset with details on measured and derived variables and products. References to published work should be provided in full.

End of Document. Last updated: 22 March 2007.

Gwen Moncoiffé, UK-SOLAS Data Management Project Co-ordinator.
email: gmon@bodc.ac.uk
Direct Line: 0151 795 4880
Fax: 0151 795 4912