

**INTERGOVERNMENTAL OCEANOGRAPHIC COMMISSION
(of UNESCO)**

**Nineteenth Session of the IOC Committee on International Oceanographic Data
and Information Exchange (IODE-XIX)
Trieste, Italy, 12-16 March 2007**

Analysis of the QC/QA survey

By Ricardo Rojas, Vice-Chairman IODE

Analysis of QC/QA questionnaire responses

A short online survey on quality control/quality assurance procedures was carried out requesting information about procedures and details of the manuals, guidelines and special software that are used by respondents for ocean profile data. This was to assist the inter-sessional working group established through Resolution IODE-XVIII. 4 to carry out its work that including:

- (i) review existing quality control procedures and software;
- (ii) discuss quality control issues of historical, real-time, delayed-mode and modern ocean profile data;
- (iii) prepare a report on (i) and (ii) above.

Unfortunately due to time constraints of the main participants, the inter-sessional group has not been able to make much progress. Some work has progressed on (i) and (ii) above, but no report has been prepared. However the questionnaire responses provide an indication of the variety of procedures and software that are in use in the 30 countries responding to the questionnaire. The countries responding are listed in Table 1 below.

Angola	Australia	Barbados
Benin	Canada	Chile
China	Cyprus	Ecuador
Finland	Greece	Iceland
India	Ireland	I. R. Iran
Japan	Madagascar	Malaysia
Malta	Mauritania	Pakistan
Peru	Russian Federation	Senegal
Slovenia	South Africa	Spain
Turkey	Ukraine	Vietnam

Table 1: Countries responding to the online questionnaire

The countries responding are well distributed geographically, but a number of countries with large data centres did not respond, however their input is needed to aid the assessment. Some of these countries are participants in the EU funded SeaDataNet project, which has a quality control task, so input has been acquired through that route.

The online questionnaire dealt with those parameters included in the World Ocean Database (as listed in Table 2 below), what quality control manuals or guidelines are in use by data centres and what software is used to quality control the data.

Temperature	Salinity	Oxygen [O ₂]
Phosphate [HPO ₄ -2]	Silicate [Si(OH) ₄]	Nitrate [NO ₃ -] and Nitrate+Nitrite
pH	Total Chlorophyll [Chl]	Total Alkalinity [TALK]
Partial pressure of carbon dioxide [pCO ₂]	Dissolved Inorganic carbon [DIC]	Transmissivity (BAC) ¹
Pressure	Air temperature	CO ₂ warming
xCO ₂ atmosphere	Air pressure	Tritium [3H]
Helium [He]	Delta Helium-3 [.3He]	Delta Carbon-14 [.14C]
Delta Carbon-13 [.13C]	Argon [Ar]	Neon [Ne]
Chlorofluorocarbon 11 (CFC 11)	Chlorofluorocarbon 12 (CFC 12)	Chlorofluorocarbon 113 (CFC113)
Delta Oxygen-18 [.18O]		

Table 2: Parameters included in the World Ocean Database

Over 90% of respondents have temperature and salinity data, dissolved oxygen and nutrients are handled by about 50% of the respondents and some meteorological parameters (air pressure and air temperature) are handled by over 40% of respondents. Usually the same methodology or manuals are used consistently by a centre but data are not necessarily quality controlled in a consistent way across different centres.

Initial conclusions from the online survey are as follows:

1. There is no consistency across data centres responding to the questionnaire
2. Many data centres use in-house procedures and software, often no further details are given. Thus the extend of the quality control carried out is often not clear, for example, if accompanying metadata are checked, if automatic tests are used, if data are visualised, quality flags added or data values changed, etc.
3. However, there are some areas of a more consistent approach. These include use of the Manual of Quality Control Procedures for Validation of Oceanographic Data, GTSPQ QC Manual and MEDAR-MEDATLAS. The countries involved in the MEDAR-MEDATLAS project all use the same procedures. In fact, all parameters that can be presented in MEDATLAS (ASCII) format are quality controlled using QC procedures described in the MEDAR-MEDATLAS documentation.
4. Some of the parameters in the World Ocean Database are not handled by the Centres responding to the questionnaire. These include: CO₂ warming, xCO₂ atmosphere, tritium, helium, delta helium-3, delta Carbon-14, delta Carbon-13, argon, neon
5. Some manuals and software are noted. The main examples are:
 - Manual of Quality Control Procedures for Validation of Oceanographic Data, UNESCO, IOC - Manuals & Guides, 1993, Manual And Guides 26
 - GTSPQ QC Manual
 - IOC Manuals on Physical Chemical Analyse
 - WOD98 Quality Control
 - WHP 1
 - Argo Quality Control Manual (Real Time and Delayed Mode)
 - MEDAR-MEDATLAS procedures and SCOOP software
 - ICES Guidelines
 - JPOTS Manual, 1991
 - Seabird software
 - Ocean Data View

- Excel
- Grapher
- Surfer

The data centres responding to the questionnaire also hold many other parameters in addition to those included in the World Ocean Database, including sea level, current speed and direction, wave statistics, bathymetry, fluorescence, chlorophyll a, chlorophyll b, total-phosphorus, ammonium, phaeophytin, total-nitrogen, total organic carbon, saturated hydrocarbon, chemical oxygen demand, biochemical oxygen demand, hydrogen sulphide, mercury, total-mercury, lead, cadmium, arsenic, contaminants, polychlorinated biphenyl, suspended solid, magnesium, calcium, wind speed and direction and biological data (marine macro-invertebrates, fish species, cetaceans). No further information was requested or provided for these parameters.

Input from SeaDataNet project

Within the EU-funded SeaDataNet project there is a task as part of the standards development “Common Data Management Protocol for dissemination to all NODCs”. This includes:

- Data quality checking methodology
- Quality flag scale protocol

A number of quality control procedures and documents from NODCs (e.g. France, Greece, Italy, Norway, Spain, Sweden, UK) and major international projects (e.g. MEDATLAS, WOCE, GTSP, GOSUD, Argo, etc.) have been examined. These mainly deal with temperature and salinity (and in some cases nutrients). Other documents exist for quality control for sea level (ESEAS and GLOSS) and met-ocean data – current meter data, waves, meteorological data (EU SIMORC project, WOCE Current Meter DAC). In addition, the ICES Working Group on Marine Data Management (WGMDM) guidelines are in use in a number of data centres. There is also the QARTOD effort in the USA, dealing with real time quality control.

SeaDataNet notes the importance of the information (metadata) that must be kept alongside the data, for example:

- Where the data were collected: location (preferably as latitude and longitude) and depth/height
- When the data were collected (date and time in UTC or clearly specified local time zone)
- How the data were collected (e.g. sampling methods, instrument types, analytical techniques)
- How you refer to the data (e.g. station numbers, cast numbers)
- Who collected the data, including name and institution of the data originator(s) and the principal investigator
- What has been done to the data (e.g. details of processing and calibrations applied, algorithms used to compute derived parameters)
- Watch points for other users of the data (e.g. problems encountered and comments on data quality)

The information above is taken from the BODC web-site (guidelines for data submission), but could equally have come from the WGMDM guidelines. Obviously this information is needed to make sense of the data and it needs to be checked for completeness and correctness.

SeaDataNet have examined the series of automatic tests which are used by the Argo programme. These are based on those used by GTSP, and similar tests are used in the Global Ocean Surface Underway Data (GOSUD) project. Although these have often been used as real-time (or near-real-time) checks, they are also useful in delayed mode. Other tests are also

noted in the GTSP documents (IOC Manuals and Guides No. 22). Some of these checks must be carried out irrespective of the data type (e.g. testing for impossible position), whereas others (e.g. ship speed) would obviously be of no use for data collected from a fixed buoy or mooring.

SeaDataNet have also reviewed a number of data quality flagging schemes in use within the marine community. In general this shows small variations on a theme for oceanographic data (usually based on the GTSP quality flags), but more complicated and detailed schemes do also exist (e.g. the WOCE surface meteorology DAC has a quite extensive set of flags, often quite specific to that type of data). After discussion within the SeaDataNet Technical Task Team, an agreed list of quality flags will be recommended for use for SeaDataNet partners.

The above activities will result in a QC manual, which is currently under development in SeaDataNet, which will be shared with the whole IODE community once it has been completed.

Input from Director, WDC for Oceanography, Silver Spring

The Director of WDC for Oceanography (Silver Spring) and Leader of the GODAR project, Mr Sydney Levitus, recommends compiling a bibliography of all quality control publications and making these available on *via* web site (e.g. OceanPortal, OceanTeacher) where people can obtain such documents is a high priority task. Some input to this is available from the questionnaire and through SeaDataNet. Requesting such documentation from other marine institutions is invaluable and should be a high priority.

In addition, he recommends requesting papers from marine data centres (data centres should be especially urged to participate, perhaps an IOC Circular Letter is an appropriate mechanism to request such cooperation), institutes, and projects (e.g., JGOFS, WOCE, etc..) regarding QC issues they face or have faced. This could be broken down into categories including:

- 1) issues regarding QC of real-time data;
- 2) issues regarding QC of the delayed-mode versions of real-time data;
- 3) issues regarding QC of ocean profile and plankton data not in the "real-time" category;
- 4) issues regarding QC of classes of oceanographic data such as "physical", "chemical", and "biological" (plankton in particular).

Mr. Levitus cites one problem faced at NODC/WDC is that metadata is frequently missing or incorrect. Data are often received that have the wrong sign of latitude and/or longitude.

The document to be produced should contain reports from any data centre, institution, or project that wishes to contribute. It could include reports from users of data, particularly from operational data (e.g. ISDN, formerly MEDS) and forecast centres (e.g. ECMWF, NCEP) regarding what problems they encounter with real-time data. The final report should summarize the documents submitted and include a list of "best" practices regarding QC. The report should focus on problems that data centres encounter, that users of data centre data and databases encounter, that operational centres encounter, etc. The purpose of the report is to improve the quality of oceanographic data and products based on these data for user communities.

However, a very interesting product, in the field of capacity building, has been initiated during 2006. Dr. Murray Brown, Chief editor for data management in OceanTeacher, has started integrating the ideas and contributions of the QC Working Group into OceanTeacher. Dr. Brown cites that, since the format of a catalog of the various QC-related documents would be very similar to the documents called "Courses" in OceanTeacher, he created a pseudo-course entitled OCEANOGRAPHIC DATA QUALITY CONTROL. This

course (number DM 209) can be found at: <http://ioc.unesco.org/oceanteacher/OceanTeacher2/CoursesHome.htm>. If you read through the "course" you'll find links to every document we already have directly or indirectly related to data quality control. They fall into 6 areas (table 3), ranging from initial planning to detailed QC procedures for specific data types. The "course" includes indications of the numbers of documents included in each category. Eventually, as QC/QA group completes the collection of primary documents -- and provides other discussions and papers – Dr. Brown can convert this pseudo-course into a real training course to be offered to Oceanteacher international students

Course DM 209	OCEANOGRAPHIC DATA QUALITY CONTROL
GOAL	[Proposed] To introduce the student data manager to the broad concept of quality control of marine data, including management practices, field measurements practices and technical procedures and methods for direct examination and quality control of datasets
LESSONS	<ol style="list-style-type: none"> 1. <u>Introduction</u> 2. <u>Science & Implementation Plans</u> 3. <u>Data Management Policies & Guidelines</u> 4. <u>Manuals & Guides for Fields Measurements</u> 5. <u>Programmatic Aspects of Quality Control</u> 6. <u>Technical Aspects of Quality Control</u> 7. <u>Integration of QC into Operational Oceanography Infrastructure</u>

Table 3. Outline of the future Oceanteacher course to be taught to trainees of IODE in QA/QC topics.

Action required from the Committee

It is clear that quality control is an important issue to IODE and needs to have further effort put in. It becomes even more important if we progress with the OceanDataPortal concept (e.g. marine data ATM) that we have commonality in our quality control procedures. The aim is for convergence to a standard way to carry out QC on a suite of variables.

The Committee is requested to:

1. Agree to continue the life of the inter-sessional working group for one further period, provided a suitable chair/leader and members can be identified, to carry out the original remit of:
 - i. reviewing existing quality control procedures and software;
 - ii. discussing quality control issues of historical, real-time, delayed-mode and modern ocean profile data;
 - iii. preparing a report on (i) and (ii) above.

A draft of the report should be submitted to the 2008 IODE Officers meeting for review and a final report submitted to IODE-XX. This should also include a bibliography of all quality control publications.

2. Agree to organising (or jointly organising) a QC Workshop, in collaboration with JCOMM (JCOMM-II recommendation 7.6.2). There is a possibility to use the opportunity of the forthcoming International Conference on Marine Data and Information Systems (IMDIS-2008) Conference in Spring 2008 for this; a QC session is already planned. IOC and the EU are jointly sponsoring this Conference.

This QC working group should work in collaboration with the JCOMM/IODE Expert Team on Data Management Practices (ETDMP).

[end]