



Suggestions  
for Data Management within Ocean  
Timeseries Data  
April 2003

# Results of previous Meeting

- All data have to be allowed for public access.
- A concern was whether a centralized placed would always have the latest version of data.
- QC would be ok for delayed-mode data, since PI should only deliver data after careful check.
- Real-time QC for timeseries is more challenging. For salinity and CO2 and chlorophyll this will be difficult. Automatic procedures could/should just flag suspicious data..
- Start a testbed from Coriolis existing multidisciplinary data sets who will provide an example of format

# Mooring data processing context

- Two types of moorings:
  - Moorings part of a network for which :
    - there is a data center taking care of data processing
    - both in real time and delayed mode data set are handled
    - computer specialist attached at the center to provide sustainable in time data processing, archiving and distribution facilities
  - Mooring handled by a laboratory
    - data processed by the scientific team for the duration of the project
    - not much of computer specialist available
    - No funding for sustainable service (processing, distribution, archive)

# A network/mooring => a format

- Aloha: an file per parameter and depth : Temperature at 90m
  - RBR XX-105 2.7 006720
  - Host time 2000/10/23 14:47:46
  - Logger time 88/10/23 14:59:28
  - Logging start 88/06/24 00:00:00
  - Logging end 88/11/19 00:00:00
  - Sample period 00:30:00
  - Number of channels = 1, number of samples = 5838, mode : 2
  - Calibration data from file 006720.cal :
  - Drawing 1359-801, created 13sep90 for logger XX-105. #6720
  - D10%8.3f
  - 1 5 0.409376 3883.07 0.168165 -1.98301e-05 0 0 0 Degrees\_C
  - 88/06/24 00:00:00
  - 27.215 27.064 26.918 26.779 26.645 26.527 26.414 26.302 26.205 26.118 26.026 25.944 25.882 25.844 25.806 25.582  
24.092 23.034 22.889 22.758 22.657 22.712 22.666 22.561 22.624 22.645 22.729 22.763 22.860 22.830 22.822 22.670  
22.603 22.590 22.569 22.603 22.607 22.574 22.649 22.632 22.695 22.578 22.670 22.620 22.628 22.683 22.670 22.544
- NTAS:all parameters in one file. One line/set of measurements. No metadata.
  - %SYS#1 -- Daily Transmission Report Thu Apr 3 01:04:18 EST 2003 -- Argos data file #1718
  - %year month day yearday hour BP AT SST Uwnd Vwnd RH SW LW RN BV CO
  - 2003 4 1 91 21 1014 25.47 26.23 -6.50 -1.00 77 39 396 0 12.5 5.60
  - 2003 4 1 91 22 1015 25.43 26.24 -6.32 -1.57 76 4 403 0 12.5 5.60
  - 2003 4 1 91 23 1015 25.44 26.23 -6.76 -1.19 77 4 405 0 12.5 5.59
  - 2003 4 2 92 0 1016 25.45 26.22 -6.88 -1.12 77 4 401 0 12.5 5.59

# A network/mooring => a format

- Animate/Estoc: No ftp site but subsetting tool to get ascii files: one file for all parameters.
- `estoc_d_id,Date_Time,mn,bid,temp10,cond10,temp20,cond20,press20,temp50,cond50,temp100,cond100,press100,temp180,cond180,temp380,cond380,press380,temp650,cond650,temp920,cond920,press920,buoyt,buoyv,buoyb,add_dat,`  
1015,2002-06-11  
06:52:52,707,0,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,  
.222,.222,877.58,56874,41121,47935,2002-08-30 12:11:09,  
1017,2002-06-11  
09:42:11,708,0,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,999.999,999.999,999.999,999.999,999.999,9999.99,999.999,99,  
9.999,999.999,999.999,9999.99,57147,40794,47914,2002-08-30 12:11:11,  
1019,2002-06-11  
13:18:06,709,2,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,  
.222,.222,877.58,56911,40898,48024,2002-08-30 12:11:15,  
1021,2002-06-11  
15:20:28,709,0,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,  
.222,.222,877.58,56963,40806,47886,2002-08-30 12:11:15,  
1023,2002-06-12  
02:12:08,711,0,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,.222,.22,  
.222,.222,877.58,58549,40958,47946,2002-08-30 12:11:21,

- One sample of netcdf file. A few metadata attached to file

# A network/mooring => a format

- Tao/Pirata: Ascii one file per mooring and parameter on FTP. WWW interface to get a file per parameter (possible to get multi-moorings files) ascii or netcdf.

Few metadata QC flags.

- Location: 2N 137E 1 Mar 2003 to 1 Apr 2003 ( 32 times, 12 depths, 1 blocks) Gen. Date Apr 3 2003
- Units: Depth (Meters), Temperature (C), -9.999 = missing
- Time: 1200 1 Mar 2003 to 1200 1 Apr 2003 (index 1 to 32, 32 lines, 12 depths)
- Index: 1 2 3 4 5 6 7 8 9 10 11 12
- Depth(M): 1.5 25 50 75 100 125 150 200 250 300 500 750 QUALITY INDICES
- DATE TIME SST TEMP TEMP TEMP TEMP TEMP TEMP TEMP TEMP TEMP TEMP TTTTTTTTTTTT
- 20030301 1200 28.850 27.410 24.480 23.970 23.020 22.420 20.530 16.900 14.300 11.330 8.430 6.300 222222222222
- 20030302 1200 28.800 26.820 24.470 24.010 23.080 22.220 20.880 17.540 14.420 11.300 8.480 6.300 222222222222
- 20030303 1200 28.650 27.710 24.480 24.030 23.230 22.520 21.430 18.410 14.130 11.480 8.520 6.280 222222222222
- 20030304 1200 29.040 27.950 24.570 23.750 23.050 22.270 20.880 17.970 13.690 11.240 8.560 6.240 222222222222

# A step by step approach

## Study a distributed approach

1. Can be simple: Just a WWW portal and a set of links to all the datasets  
⇒ Can be put in place within 2 months
2. A bit more integrated for users
  - Data processing stay at the level of National center or Laboratory (including quality control process)
  - Centralized metadata catalog
  - Centralized WWW portal providing simple subsetting and printing tools
  - Data distributed among the network (national centers/PI laboratories)
4. Can be more complicated if we want to provide interactive interfaces on the WWW with fancy manipulation and subsetting tools.

# A step by step approach

## Ease access to OTS data : an intermediate solution for 2003

⇒ *An intermediate solution between options 1&2 of previous slide for 2003*

- Set up a common generic format able to serve the core variable defined for OTS project
- Rely on a few global data centers to provide an easy centralized access to data by mirroring the national/laboratory servers
- Delayed mode data + some realtime data

⇒ Define a common format in the 6 next months (Format working group to be created)

⇒ Who will be the volunteers Global centers?

⇒ Set up these global servers with volunteers data providers



# Defining Common Data Format

- Rely on common standards already used for in-situ data:
  - Netcdf (Argo, Woce) + an version in Ascii
  - Contains metadata and data
  - Must be generic to be able to handle new parameters
  - For each parameter Data + QC flags

# To start the format discussion

- **For each data set :**
  - unique address for each platform : does it exist ?
  - Organization
  - project
  - contact person : person in charge of the mooring
  - data centre : centre which processed the data
  - date of last update of the data set
  - position
  - list of available parameters GF3/IOC code ?
  - technical parameters
  - QC

# To start the format discussion

- **For each parameter :**
  - sampling method description
  - data processing level
  - list of instruments and sensors
  - calibration equation
  - parameter unit
  - technical parameters
- **For each measurement :**
  - date and time (UTC)
  - date QC flag
  - measured parameter