REPORT BY THE TASK TEAM ON DATA MANAGEMENT

(Submitted by Mayra Pazos, Chairperson TT-DM, USA)

Summary and purpose of the document

This document contains the report by the chairperson of the DBCP Task Team on Data Management.

ACTION PROPOSED

The Panel will review the information contained in this report and comment and make decisions or recommendations as appropriate. See part A for the details of recommended actions.

_____

Appendices: A. Report by the Task Team on Data Management

B. Terms of Reference of the DBCP Task Team on Data Management

C. GTS Processing Monitoring graphs from CLS-France

-A- DRAFT TEXT FOR INCLUSION IN THE FINAL REPORT

6.1.1   Ms Mayra Pazos, Chairperson of the Task Team on Data Management (TT-DM) reported on the progress during the intersessional period.

6.1.2   The Panel noted that the lack of standardization in Argos data formats raised last year is still an issue for both processing in real time (GTS) and decoding and processing data in delayed mode. The TT-DM Team encourages manufacturers to promote standard formats using the DBCP-M2 concept, one single observation per message. It is recommended that in cases where additional data are needed in the message, that the message has the "standard" data in the front portion of the message and the user-specific data behind that.

6.1.3   There is a need to make NWP/Ocean model outputs available to buoy operators for (i) Checking data quality before releasing them to the GTS, and (ii) Checking data that had been removed from the GTS, to assess if it had improved over time and could be disseminated on the GTS again. Meteo-France has developed some internal tools to make these comparisons but Technical problems avoid making these tools available on the web.

6.1.2   The Panel agreed on the following recommendation:

- The team would like to remind members of the DBCP about the WMO IDs being extended to 7 digits for BUFR data streams and to make sure the rules on 7-digit WMO ids are taken into account as appropriate (see DBCP-25 report). If using Météo-France QC-Tools to check GTS data against models, 7 digits WMO numbers should be entered (**recommendation**).

6.1.3   The Panel then agreed on the following action items:

- Promote standardization of data transmission formats using DBCP-M2 concept (**action; TT-DM; ongoing**);
- Assist Pierre Blouch and Jon Turton in preparing a methodology to compare non-GTS buoy data with NWP/Ocean models, open to anyone via the web. (Continuation of DBCP-25 Action 8.8.2) (**action; TT-DM; DBCP-27**);
- Review the DM Cookbook (The Oceanographer's and Marine Meteorologist's Cookbook for Submitting Data in Real Time and In Delayed Mode) (**action; TT-DM; DBCP-27**);
- Review the SSS QC document and finalize it (**action; TT-DM; DBCP-27**);
- The team should continue, with support of the Technical Coordinator, to assess the adoption of BUFR by GTS nodes and ensure that all data that is expected is definitely received at modeling and archiving centers (especially ISDM, NODC and ECMWF) by reviewing the differences between BUFR and BUOY messages received at each center (**action; TT-DM & TC; DBCP-27**);
- Assist in reviewing the Buoy template for BUFR, which will be updated in 2010-11 to include new requirements for observations, as well as the additional metadata identified as critical (**action; TT-DM; DBCP-27**).

6.1.4   The Panel thanked Ms Pazos and members of the Task Team for their efforts. The full report of the Task Team is provided in Appendix A and will be included in the CD-ROM that will be distributed with the Session final report.

———————————

Appendices: 3

APPENDIX A

REPORT BY THE DBCP TASK TEAM ON DATA MANAGEMENT

During the intersessional period, the TT-Data Management Team promoted discussion between members, revised the proposed recommendations from last year to assess actions taken and proposed new recommendations.

1. Receive and Review reports

The **Technical Coordinator** reviewed a new product from ISDM (based on requests from B. Bradshaw, ISDM) showing data from the previous month and provided feed back as requested. http://isdm.gc.ca/isdm-gdsi/drib-bder/MonthlyKML/MonthlyKML.htm

ISDM reported that there has not been any progress beyond the exchange of data center reports between SOC (Jean Rolland) and RNODC/DB (Bruce Bradshaw) which were included as annex to the 2009 report. During the last DMCG meeting (April 8-9 2010, Ostend, Belgium) it was recommended that members from this group contact ISDM and SOC to coordinate respective organization's activities.

Meteo-France believes the coordination between SOC and ISDM is a positive evolution and wishes to continue the effort to build a mirror system in which ISDM and Meteo-France would work in the same way, with the same collected data, the same archive, in order to have a full back-up system, with the exception of the data distribution from their site. ISDM runs a portal to make the data available which Meteo-France would not be able to do, at least for now.

2. Table Driven coding requirements for data buoy observations

The Panel is reminded that BUFR will be the only way to communicate on the GTS after 2012.

The Technical Coordinator:

- is working with the JCOMM Task Team on Table Driven Code Forms to progress on enhancing the Buoy template for additional requirements identified since 2005 and to include additional metadata fields (Meta-T).

- Reviewed how OceanSITES metadata could be included in the BUFR template for Buoy data.

- With respect to wave information, the TC contacted Jean Bidlot and others about concerns expressed about the definition of wave period of (including Swell and Wind waves) and Average Wave Period. The items of concern were: 022011 (Period of waves), 022012 (Period of wind waves), 022013 (Period of swell waves) which are in BUOY and SHIP BUFR templates and 022074 (Average wave period) which is in the BUOY and WAVEOB templates. Still working now on how to clarify definitions and observational requirements for waves with the JCOMM PP WET and Wave data experts.

- The BUOY template for BUFR will be updated in 2010-11 to include new requirements for observations (e.g. waves and new sensors) as well as the additional metadata identified as critical

- Dr Bill Burnett will from now on chair the JCOMM Task Team on Table Driven Code Forms, so queries about BUFR should be sent to him in the first instance.

- The **Technical Coordinator** provided inputs to WMO on additional entries for the BUFR Code tables concerning "Buoy Type" and "Location System" used in GTS messages (based on feedback from the JCOMM Task Team on Table driven Codes) and in the BUFR

template for data buoys, after consultation with the JCOMM Task Team on Table Driven Code Forms.

3. Real Time Distribution of Data

The **DAC** continues to distribute and monitor all data from AOML's drifters on the GTS. During this intersessional period the DAC insured that 1300 drifters with an average of one position fix every 1.2 hours were placed on the GTS. Consequently the DAC oversaw the distribution of more than 9.5 million quality-controlled observations on the GTS.

The **DAC** released the data of 6 SVPS drifters (SVP with salinity) on the GTS, but discovered salinity was not being reported correctly, with the help of Yann Bernard, from CLS-France, the GTS template was modified and the salinity measurements were reporting on the GTS correctly.

During this last 12 months, the **DAC** has taken a large amount of drifters' SST data off the GTS due to reports of erroneous SST data. Manufacturer of these drifters was contacted about this unusual situation.

With regards to Transition to BUFR and data flowing on the GTS in BUFR format:

The **Technical Coordinator** began to work with ISDM on reviewing the differences between BUFR and BUOY messages received to assess the adoption by GTS uplink centers.

- Some centers are sending no BUFR data at all, so the Technical Coordinator sent requests to find out what their strategy was for the transition
    o Several centers have no BUFR data sent to-date for data buoys (according to stats supplied by ISDM and contents of the JCOMMOPS database) e.g.

    RKSL Seoul, Rep of Korea – Moored buoys using Orbcomm DEMS New Delhi, India – Moored buoys using Inmarsat   CRJTD Tokyo, Japan – Arctic Buoys/POPS (using Iridium) CWAO– Montreal, Canada, some Atlantic drifters (using Iridium) CWEG – Edmonton, Canada. Some Pacific and Arctic buoys (using Argos or Iridium, but processed in Canada)

    These centers need to consider how soon BUFR can be tested and implemented ahead of the 2012 cut-off. The Technical Coordinator contacted the GTS Programme managers for each of the buoys program's using the GTS centers in question asking when their organization would be likely to begin encoding data in BUFR format. The responses will be reported to DBCP 26
- The **Technical Coordinator** began to work out how to resolve some issues and provided some feedback to ISDM, however
    o There were many and varied discrepancies that will need to be reviewed case by case.
    o ECMWF will be consulted too to assess the data received there in real time.
    o This is ongoing and will be a large task, but some recommendations should be ready by DBCP-26.

ISDM identified a recurring problem with Triton buoy data/thermistor chain data transmitting on the GTS. Of all sensors attached to the Triton moorings, the ones at 250 m depth often (>10%) report suspiciously large temperature and salinity variations that are beyond normal climatological values. Issue identified by ISDM (M Ouellet). The Technical Coordinator contacted the Program Manager (Ken Ando) and CLS Japan – issue was found to be at GTS encoding time, (checksum control wasn't applied on sensors –250m). CLS fixed the problem by making a modification to the processing for all active TRITON buoys and CLS Japan applied the same fix was made for all other Argos ids.

The **TC** Worked with Cubic I (CLS Japan) and JAMSTEC to get WMO IDS for two CO_2 buoys to send data to the GTS. JAMSTEC has agreed to distribute SST and Salinity of their CO_2 buoy to GTS via CLS.

The **TC** recommended that the bulletin headers used for buoy data disseminated from Toulouse and CLS America in BUFR format be updated, to adapt to changes in recent years in the GTS Manual Table, table C6. Most other data providers use the Header "IOB"aii rather than IOZX. CLS-France will implement changes at the end of July, Pierre Blouch also send a similar request to his colleagues at Meteo-France and to the Spanish Meteorological Service.

Meteo-France reports that the distribution of BUFR messages onto the GTS should start during summer 2010 after the first SVP-B drifters ordered for E-SURFMAR and having a resolution of 0.01K for SST are deployed. The resolution is only 0.1K in FM18-BUOY messages. BUFR messages had been generated for more than one year but they were not distributed on the GTS until now.

ISDM has been decoding BUFR data for two years, inconsistencies with the retransmission of data to the GTS from Iridium equipped platforms have been found and Hester, Yann, Bruce and others are looking into this. ISDM archival data source remains the GTS FM18 data stream at this time, but data gaps resulting from the small number of reports only circulated in BUFR format will be filled. Overall the comparison of the two data streams shows excellent correlations with a few exceptions.

ISDM reported that at the June IABP meeting, Environment Canada reported that Christian Haus has developed software to encode data in FM18 format from Iridium equipped buoys. Environment Canada is using the software to encode and circulate Iridium source data on the GTS. Opportunities may exist to use this system by others in the absence of other routes to the GTS for Iridium buoy data. Contact: Edward.Hudson@EC.gc.ca for more information.

CLS-France
The average GTS delivery time has been 18 min improved between June 2000 and June 2010. This improvement is explained by:
- A sixth operational Argos satellite in August 2009 (NOAA NP)
- Improvements on CLS Argos-GTS processing system (shorter time of processing, faster bulletins compilation before sending, …)
- Best management of settings on processing templates especially on the time of observation computing.

Developments are in course to provide statistics on operational GTS monitoring tool by ocean basin. Deadline = DBCP XXVI meeting.

Correction of an issue on TRITON buoys template: the checksum control wasn't applied on sensors at -250m; all settings (template an IDs) were adjusted in June 2010.

BUFR coding for SYNOP observations is in course of validation. Deadline = end of 2010.

The CLS GTS processing system was updated in January 2010 to adhere to the new World Magnetic Model: http://www.ngdc.noaa.gov/geomag/WMM/DoDWMM.shtml.

Integration of GPS positions in GTS observation for all buoys, which have a GPS with an Automatic switch on Argos location when GPS is failed.

NDBC reported that during this period, they provided 24x7x365 data analysis and quality control support for 115 NDBC moored buoy platforms, 51 coastal marine stations, 210 water level stations, 39 deep-ocean tsunameters, 55 Tropical Ocean Atmospheric moored buoys in the

equatorial Pacific, 40 oil and gas platforms in the Gulf of Mexico and 250 Integrated Ocean Observing Systems (IOOS) partner platforms (moored buoys and coastal stations). Through this effort, NDBC provided over 10 million quality-controlled observations to the Global Telecommunications System (GTS) in real-time.

South Africa Weather Service's main concern remains to be the delayed time in real time. Unfortunately, recent global economic downturn put paid to the SAWS' plans to reinstall and upgrade LUT's. Since then, CLS has revealed a plan to upgrade their IS antenna network to achieve greater coverage over the oceans. This network will extend to the Southern Hemisphere to ensure the timely delivery of buoy messages for inclusion in the synoptic analyses. They continue to report data problems as soon as they noticed them. They appreciate the monthly maps/data they get from Hester every month.

4. Delayed mode distribution and archiving of data

Regarding delayed mode distribution of data to the RNODC archiving center for drifting buoy (ISDM, Canada), the Drifter Data Assembly Center (DAC) at AOML, is preparing the next update that will cover the period from July 2007 through December 2009 and expects to have it ready for submission by end of August 2010. Quality –Controlled and interpolated drifter data through March 2010 can be downloaded from the DAC web site (www.aoml.noaa.gov/phod/dac/dacdata.php).

ISDM reported that future SVP drifter data submissions from AOML, will be processed and available at their center's web site in a more timely fashion due to improved processing systems.

Several functional updates to the **ISDM** RNODC/DB web site have been completed. Work is nearly complete on a new monthly Goggle Earth KML data product, which will include buoy tracks and a new observed data graphical products as well as improved inventory level meta-data.

Meteo- France would like to reiterate that their purpose for handling delayed mode data is mainly for quality control. Every day, Meteo-France automatically collects the surface marine observations in the Meteo-France GTS databank. In parallel, the model data at the position and time of the observations are collected. Comparisons are performed to issue black lists and monthly statistics which are made available on the Web (see the **http://www.meteo.shom.fr/qctools**/). Plots of the data and of the differences with model outputs over the past 2 weeks are also displayed there. **Important change**: after July 1, 2010, all data buoys (drifting or moored) are referring to their 7-digit WMO numbers in Meteo-France archives and for quality control tools, even if they only report in alphanumeric codes (FM13-SHIP or FM18-BUOY). These 7 digit WMO numbers follow the conversion rules described by WMO office. For example, if you are using QC-Tools to check the data of moored buoy 62163, even if it is reporting through FM13, you must enter 62**00**163 in the query form. Using the 5-digit WMO number will produce no graph.

NDBC met with NOAA's National Oceanographic Data Center (NODC) to establish new techniques to support the archive of climate observations.   This partnership will lead to the capability to preserve marine datasets that meet the requirements for climate observations established by NOAA and the Global Climate Observation System (GCOS). These standards include regularly assessing the quality of data as well as providing the details and history of local conditions, instruments, operating procedures, data processing algorithms and other factors pertinent to interpreting data (i.e. metadata). Taking advantage of new techniques, using Open Geospatial Consortium Inc. standards and Sensor Observation Services (funded by NOAA's Integrated Ocean Observing System Program), will enable a new degree of interoperability within and between NDBC and NODC. By the end of January 2011, NDBC will end the procedure of archiving data in the outdated F291 format and begin archiving all 2011 (and beyond) observations using netCDF.

OceanSITES data management capabilities continue to mature. The Data Management (DM) group held their 4th meeting in Paris in April 2010. This "face-to-face" meeting followed upon four "virtual" meetings held using WebEx technology (in December 2009 and January, February and March 2010). During this period the OceanSITES DM group accomplished the following items: First, the group confirmed the following file name convention, as finalized during one of the virtual meetings. However the group modified the definition to ensure any data file without a PARTX will not end with an underscore (i.e. the final section of the filename definition was changed from _<PARTX>.nc to <_PARTX>.nc). For more information,
visit the OceanSITES web page: www.oceansites.org

5. Format Issues

Lack of standardization in Argos data formats raised last year is still an issue for both processing in real time (GTS) and decoding and processing data in delayed mode. The **TT-DM** Team encourages manufacturers to promote standard formats. The DBCP-M2 concept, one single observation per message, standard algorithm to compute the observation time) proved its efficiency. The DBCP_M2 format is not a strict data format and derived formats exist or may be easily developed for WOTAN and salinity drifters. Resolution and range may be adapted for each sensor. An example, given by Pierre Blouch (Meteo-France), were salinity drifters produced by Pacific Gyre have their own data format which don't follow the DBCP-M2 concept, consequently salinity drifters deployed by LOCEAN in France were not released onto the GTS, due to lack of time to make the necessary changes to the GTS templates. It is recommended that in cases where additional data are needed in the message, that the message has the "standard" data in the front portion of the message and the user-specific data behind that.

Meteo-France reports that Iridium SBD formats for drifting buoys were updated in the frame of the DBCP Iridium Pilot Project. The main change was the introduction of a format identifier (one byte) at the beginning of each message type. Format #000 was allocated to SVP-B drifters fitted with a GPS, format #001 to SVP-Bs having no GPS, format #020 for salinity drifters, etc. The list of available formats as well as their descriptions is published on the Iridium-PP website and will be given in the annex of the final Iridium pilot project report.

6. Comparison to Models for non-GTS data

With respect to DBCP-25 agenda item 8.8.2 to make NWP/Ocean model outputs available to buoy operators i) to check data quality before sending to the GTS upon deployment and ii) to check data that had been removed from the GTS to assess if it had improved over time and can be released to the GTS again:

Meteo-France reports that they have developed some internal tools to compare: a) raw observations of air pressure and SST reported by SVP-Bs exclusively, b) with outputs from Meteo-France, ECMWF and Mercator models, over the past 8-10 days, with certain conditions: 1) raw data must have been received through e-mails on their development machine (Iridium SBD), 2) or through an access to CLS raw Argos data for the concerned buoys, 3) the buoys to be checked must at least report one parameter onto the GTS (even battery voltage) so that the co-located model data are automatically acquired. These tools are presently used for E-SURFMAR operations. They may be also used for iridium buoys for which Meteo-France process the data and makes GTS bulletins. Technical problems avoid making these tools available on the web. Internal tools also exist at Meteo-France to extract model data from Meteo-France model analysis (i.e. sea level pressure, SST, wind) at a given location for a given short period. Extractions might be occasionally provided to buoy operators on request.

7. Review all relevant JCOMM Publications

The **TC** sent the DM Cookbook (The Oceanographer's and Marine Meteorologist's Cookbook for Submitting Data in Real Time and in Delayed Mode) to TT-DM for interest and feedback, but not much response was received.

New Iridium formats put on the PP website: http://www.jcommops.org/dbcp/iridium-pp/
and Findings (http://www.jcommops.org/dbcp/iridium-pp/findings.html).

The **TC** Reviewed and Commented on an Article by *Jou*Beh Technologies entitled "End-to-End Processing and Transmission of Environment Data to the Global Telecommunications System using the Iridium Satellite System." (for submission to the Canadian Meteorological and Oceanography Society Bulletin.

Document 37 - "GUIDE TO BUOY DATA QUALITY CONTROL TESTS TO PERFORM IN REAL TIME, BY A GTS DATA PROCESSING CENTER". The **TC** updated the document sent out last year to ensure it is generic in terms of Satellite Telecommunications systems – meaning it is not related to any one telecoms system and mentions when the systems require (or do not require) different processes. Document is available on the meeting website for review and is on http://www.jcommops.org/dbcp/doc/DBCP-37/DBCPDOC37_QC_v1_4.doc

The **TC** also added a link on the DBCP website under Data->Quality Control (Other Resources) - to the tool "iQuam - quality monitor for in situ sea surface temperature", this is an online monitoring tool developed at NOAA NESDIS/STAR the link is: http://www.star.nesdis.noaa.gov/sod/sst/iquam/

The **TC** reported regarding Action item 8.8.3 from DBCP 25 (The development of quality control processes and approaches for SSS) that inputs from various people and existing documentations were compiled and a draft document was created to provide details of Sea Surface Salinity quality control processes to JCOMMOPS for the TTDM and TT-IBP teams to review. The document is posted on the DBCP documents page.

_____

APPENDIX B

TERMS OF RERERENCE OF THE DBCP TASK TEAM ON DATA MANAGEMENT
(as adopted at DBCP-XXIV)

The DBCP Task Team on Data Management shall:

1. Receive and review reports from the Data Management Centres specializing in buoy data, i.e., (i) the Météo-France SOC / DB, and (ii) the ISDM, Canada RNODC / DB; reconcile any overlaps with emphasis on differences;

2. Liaise with the DBCP Task Team on Quality Management for compiling table driven coding requirements for data buoy observations, for all relevant applications, and submit them in a consolidated way to the DMPA Task Team on Table Driven Codes;

3. Address issues to do with real-time distribution of data, including GTS issues, timeliness and methods to improve data / flows;

4. Address issues relating to delayed-mode distribution and archiving of the data;

5. Seek input from data users on which instrumental metadata is most important and how it is best managed and coordinate these activities with the JCOMM Meta-T Project;

6. Review all relevant JCOMM Publications, to make sure they are kept up-to-date and comply with Quality Management terminology;

7. Follow-up with regard to the development of the WIGOS Pilot Project for JCOMM and make sure that the developments proposed by the Task Team are consistent with the WIGOS and WIS requirements;

8. Make recommendations to the DBCP Executive Board or the DBCP for addressing the issues above; and

9. Report to the DBCP Executive Board and the DBCP at its biennial Sessions.

Membership:

The membership is open to all Panel members. The Chairperson, appointed by the Panel, has selected the following team members:

> Ms. Mayra Pazos (TT Chairperson and GDP representative);
> Mr. Bruce Bradshaw (RNODC representative);
> Mr. Jean Rolland (SOC representative);
> Mr. Pierre Blouch (France);
> Mr. Johan Stander (SAWS);
> Dr. Bill Burnett (NDBC data manager);
> Mr. Yann Bernard (CLS data manager);
> Ms. Hester Viola (DBCP Technical Co-ordinator (*ex officio*));
> Mr. Jeff Wingenroth (Technocean Inc.), and;.
> Ms. Emily MacPherson (MetOcean).

_____

APPENDIX C

GTS PROCESSING MONITORING GRAPHS FROM CLS-FRANCE

Number of WMO platforms processed per day

Number of drifters GTS processed per day

Daily average delivery time (in min) of text bulletins on GTS

Number of text bulletins sent on GTS per day

_____