WORLD METEOROLOGICAL ORGANIZATION

INTERGOVERNMENTAL OCEANOGRAPHIC COMMISSION (OF UNESCO)

JOINT WMO / IOC TECHNICAL COMMISSION FOR OCEANOGRAPHY AND MARINE METEOROLOGY (JCOMM)

SOT-V/Doc. I-3.1.4
(8.4.2009)

SHIP OBSERVATIONS TEAM

ITEM I-3.1.4

FIFTH SESSION

GENEVA, SWITZERLAND, 18-22 MAY 2009

Original: ENGLISH

PLATFORM/INSTRUMENT METADATA REQUIREMENTS AND META-T

(Submitted by Derrick Snowden, Chairperson, META-T, Hester Viola, and Mathieu Belbéoch, JCOMMOPS)

Summary and purpose of the document

This document provides information on requirements for platform/instrument metadata, and activities of the Water Temperature Instrument/Platform Metadata Pilot Project (META-T). It is reporting on progress and proposing the way forward for active participation of the SOT in related activities.

ACTION PROPOSED

The Team 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. The Water Temperature Instrument/Platform Metadata Pilot Project (META-T)

References: A. SOT-V Doc-I-4, Section I-4.6 and Appendix G (report by the Task Team on Coding)

- A - DRAFT TEXT FOR INCLUSION IN THE FINAL REPORT

I-3.1.4.1 Mr Derrick Snowden, Chairperson, META-T presented a review of the META-T Pilot Project, including a brief reminder of the original goals of META-T and the purpose for its inception.

I-3.1.4.2 Initially the pilot project work plan has focused on collecting information from various user groups detailing the type of metadata that should be collected for each JCOMM data stream. Work that is more recent has been focused on comparing these lists of requirements with the actual JCOMM data streams to identify the gaps in the current processes. The presentation was focused on the gap analysis as it pertains to real time VOS and XBT data.

I-3.1.4.3 The Team noted that the gap analysis showed numerous metadata elements that are not currently collected and provides a list of goals for the operational community to pursue.

I-3.1.4.4 In addition to identifying the information that needs to be collected, Mr Snowden presented some options for how it might be managed as part of an end-to-end data system.

I-3.1.4.5 The Team considered the META-T suggestions below in light of the implied changes needed to the VOS and SOOP data management and data collection procedures. Mr Snowden invited the Team to work more closely with META-T to map out a realistic pathway that will address the gaps identified here but not introduce unnecessary and burdensome changes to the current SOT practices.

I-3.1.4.6 The Team

(i.) Invited the Cross cutting Task Team on Delayed Mode VOS Data (TT-DMVOS), the SOT Task Team on Pub47 metadata, and other Groups involved in SOT data management to liaise during the next intersessional period with the META-T Steering Team (***action: TT Pub47 & TT-DMVOS; SOT-VI***);

(ii.) Welcomed the participation of the following SOT members on a TSG template design group to respond to the META-T Users Survey (***action: SOT members; SOT-VI***) [*pending discussion, list of volunteers to be added in the final report*];

(iii.) Invited all interested parties, even those not part of the steering team, to go to marinemetadata.org and register as a user and participant in META-T to monitor project status (***action: SOT members; SOT-VI***);

(iv.) Invited META-T to liaise with the relevant VOS operator community to determine if the current average FM-13 message distributed on the GTS is populated completely. Additionally, obtain lists of the actual fields transmitted to shore by the three electronic logbooks (***action; META-T; SOT-VI***);

(v.) Requested electronic logbook developers to consider adding the functionality to transmit periodic Admin messages containing all known category 1 and 2 metadata (META-T website will include the list of desired fields) (***action: e-logbook developers; SOT-VI***);

(vi.) Requested SOOP Operators to consider reintroducing the collection of meteorological observations as part of a regular XBT message (***action: SOOPIP members; ongoing***);

(vii.) Invited SOT members consider joining META-T in any of the capacities listed in Appendix A (***action; SOT members: SOT-VI***).

_____

Appendix: 1

APPENDIX A

THE WATER TEMPERATURE INSTRUMENT/PLATFORM METADATA PILOT PROJECT (META-T)

1. META-T Background and Project History

1.1 The Ship Observations Team is invited to review the META-T Project website[1] for information on the background and prior work plan of the pilot project team. As a brief review, the META-T pilot project was initiated at the urging of the Ocean Observations Panel for Climate (OOPC) with a very general aim.

> "The aim of META-T is to investigate and recommend the use of metadata to improve the quality and usefulness of ocean temperature information, particularly in real-time. The group should investigate and recommend data transmission codes and content, storage and distribution of data, for specific data streams."

1.2 In pursuing this aim, the pilot project began with a survey of the user communities involved with the JCOMM data streams. The surveys were aimed at understanding the breadth of the metadata that is relevant to a data stream and to understand when and where the metadata is available. Further, the surveys resulted in a master list of metadata needed by users from multiple communities. With the results of the surveys, the pilot project categorized the required metadata into three categories based on timeliness requirements and data distribution mechanisms. This categorization of the metadata provides the basis for the pilot project to develop a general framework describing how the metadata could be distributed to the global user community.

1.3 What follows in this document is an assessment of the current status of the pilot project and an analysis of the gaps in the metadata framework as it applies to VOS and SOOP platforms.
Throughout the review, action items will be noted that, if implemented by SOT, will advance the goal of creating a JCOMM metadata management framework. Some of the recommendations included in this document imply significant changes to the current SOT data management procedures so the team is invited to discuss whether these changes are feasible under current operational constraints.
The purpose of this summary and of the presentation is not to define an implementation plan or a technical summary of the desired systems. Rather it is to describe the general goals of the pilot project in such a way that some gaps in the data streams are identified. These gaps result from multiple operational constraints. For example, some National Met Services do not collect certain metadata, as it is not perceived to be relevant to their mission. Similarly, some operate under different timeliness constraints so that even if metadata is known it is not immediately available. Budgetary constraints prohibit or limit the amount of information transmitted by satellites and therefore introduce a delay in metadata availability and uncertainty of its accuracy. Finally, the information technology infrastructure necessary to support the management and exchange of the required metadata usually does not exist. Despite these challenges, JCOMM and OOPC initiated the pilot project based on the recognition that the current framework was inadequate for real time applications, long-term climate services applications, and for long-term preservation of data.

2. Ideal Metadata framework

2.1 What is the target or the ideal framework? Ideally, at any point after an observation reaches a user, that user should be able to determine the suite of instruments that resulted in the observation.

1: The META-T Pilot Project web site is hosted at the Marine Metadata Interoperability Initiative site at http://www.marinemetadata.org/community/teams/metat . The Pilot Project Terms of Reference are at http://www.marinemetadata.org/community/teams/metat/tor

They should be able to determine the platform on which the instrument resides and the physical configuration of the instrument on the platform. Further, the relevant contact information for the operator of the platform should be easily obtainable. They should be able to determine the calibration and maintenance history of the instruments, and whether or not, the observation underwent quality control prior to distribution. All of this information should be available using automated systems in a standardized, machine-readable format. A reasonable goal is to have the identified metadata available through networked services, using internationally accepted standards in the same way that observational data is made available through services such as OPeNDAP and Open Geospatial Consortium Web Services. The information must be made available as soon as it is known and must be kept accurate. It must be delivered using standard protocols and formats and the planning for the long-term preservation of the information must be undertaken.

2.2 For the purposes of the SOT, the operating assumption will be that we are trying to push as much metadata as is known across the GTS. To do that we need to increase the amount of metadata transmitted via satellite and make the delayed mode collection and management of metadata more robust and in some cases centralized. For the metadata that is not transmitted over the GTS, META-T is developing recommendations on how to manage and exchange that information. However, META-T is not currently considering ways to collect that metadata. In the future, a tighter integration with the operational teams concerned with collecting data such as the TT-DMVOS and TT-Pub47 will be helpful in designing a more comprehensive system (ACTION). What follows are a few suggestions about how to advance these goals.

3. Current Status and Work Plan

3.1 The aforementioned surveys[2], available at the META-T project page, constitute the best available information on the current needs of the operational and science communities concerned with VOS and SOOP data. The compiled needs assessment represents the best available set of requirements that can feed into the design of a metadata management framework for climate data management within JCOMM. This requirements list is not complete. For example, the Thermosalinograph community did not contribute substantially to the surveys so the META-T team was left to assemble information on its own. This will invariably introduce gaps. (ACTION: META-T chair to submit survey to representatives from the TSG community at SOT-V. TSG community to respond.)

3.2 Each metadata element from the requirements lists falls into one of the following three categories.

1. Metadata necessary for real time interpretation of the data and included in the GTS data stream (i.e. available and accurate on the order of hours after the observation),
    a. Some data is transmitted from ship to shore,
    b. Some data is available from onshore servers that are kept up to date by regular submissions of metadata, in a standard format, by ship and other platform operators,
2. Metadata necessary for analysis that is available on demand, in a standard form, and is as up-to-date as possible (i.e. available and accurate days to weeks after the observation),
3. Metadata required for long-term preservation of the data that represents the archived climate record.

3.3 The metadata lists, and the categorization lists above, were then compared with current data formats to get a sense of how the current observation and data processing procedures may need to be augmented in order to accommodate the requirement for added metadata in the real time data stream. Initially the analysis was constrained to real-time data needs (i.e. categories one and two). Consideration of the delayed mode metadata needs and the archive needs is ongoing. In the following

2: The responses from the community surveys can be found at http://www.marinemetadata.org/community/teams/metat/surveys

sections, the results of this analysis are presented separately for VOS and SOOP data streams.

3.4 VOS Status/Actions

3.4.1 The VOS survey mentioned above provides a good list of necessary requirements to ensure the long-term usefulness of VOS data as well as improving its usefulness in real time. A spreadsheet has been developed that lists all the desired fields and maps each to the various file formats used to distribute VOS data and metadata. The types of file formats used to exchange VOS information include:

1. BUFR template (To be approved)
2. FM-13 Traditional ASCII Code
3. IMMA
4. IMMT
5. Pub-47
6. SensorML (Potentially to be used for complete JCOMM platform metadata characterization. To be defined by the Meta-T.)

3.4.2 No single format includes all metadata listed in the VOS user survey so a multiple format solution must be explored. Additionally, only the BUFR template and the FM-13 TAC are strictly for use in real time. We begin with the FM-13 message as a baseline and assume that all metadata in the current FM-13 message is routinely collected and distributed in some way. It would be helpful to know from VOS operators whether this assumption is valid (ACTION: Liaise with the relevant VOS operator community to determine if the current average FM-13 message distributed on the GTS is populated completely. Additionally, obtain lists of the actual fields transmitted to shore by the three electronic logbooks.) The following list shows those metadata elements that were identified by the User Group surveys as being important for real time applications (over the GTS) but are not included in the current FM-13 message. This represents the gap in the current operational procedures and presents, clearly, the goals for VOS operators in terms of how to augment the current collection procedures.

3.4.3 Fields in the current META-T Category 1 (i.e. to be included in BUFR) and not in the current FM-13 message:

- Unique ID: Tag that is part of the GTS message that uniquely identifies the message and is retained with the data always.
- Year of the observation: A four digit year (YYYY)
- Month of the observation: A two digit month (MM)
- Time indicator flag: Gives precision/resolution of the original time recording
- Position indicator flag: Gives precision/resolution of the original position recording
- Position method flag: Method by which, position information was gathered (e.g. dead reckoning, GPS, Argos transmitter etc.)
- Ship over ground speed: Instantaneous speed taken over a period comparable to the observations i.e. ten minutes instead of averaged over the last three hours.
- Ship speed indicator flag: method of ship speed calculation
- Ship direction indicator flag: method of ship direction data
- Ship heading: instantaneous value (in addition to ship direction already reported in FM-13)
- Ship heading flag: method of determining ship heading
- Relative wind speed: instantaneous value
- Relative wind speed indicator flag: method of determining relative wind speed.
- Relative wind direction: instantaneous value
- Relative wind direction flag: method of determining relative wind direction.
- Depth of SST sampling (below maximum summer load line (include height for radiometers))

- Difference of MSLL from MWL: as for VOSClim, gives changes in draft
- Original units of SST measurement
- Precision of SST measurement
- Sampling interval and scheme for SST measurements
- Quality Indicators: Was the data QC'd before GTS distribution? Possible to also include the results of the QC tests (i.e. specific QC flags on specific variables)

3.4.4 Some of the fields listed above are actual observations of contextual variables that vary on the same time scales as the meteorological variables and must therefore be included in the real time report transmitted to shore (e.g. Ship heading, relative wind). Others are slowly varying metadata that change on much longer time scales and are generally constant over the course of one cruise. However, they do vary and often vary more frequently than the typical quarterly Pub47 submission. An ideal solution for this type of metadata is to implement a new message type from the ship, which does not contain data, but contains values for slowly changing metadata. For example this could be a ship to shore transmission that only occurs periodically, (like the Admin message implemented by the SEAS electronic logbook). A new Admin message could be sent periodically, such as at the beginning of a cruise, or only when there is a change to the metadata. Not every VOS ship is equipped with electronic logbook software, but a significant number are (ca. 70% for the US VOSF). It is clear that an Admin message that carries all the metadata describing the configuration and inventory of the sensors on the ship would significantly increase the quantity and accuracy of the information available in real time transmissions. The actual implementation of this type of solution would require an analysis of the cost implications of an added message to the program. Other electronic logbooks should begin to develop the capability to transmit similar Admin type messages. (ACTION: Electronic logbooks develop the functionality to transmit periodic Admin messages containing all known category 1 and 2 metadata. META-T website will include the list of desired fields.)

3.4.5 For ships, not using electronic logbook software the solution involves relying more heavily on the Pub-47 or a similar delayed mode collection scheme. However, as noted by the Task Team on Pub47 Metadata, Pub-47 is not managed in a timely fashion at the WMO currently and improvements to Pub-47 distribution must be made. Nevertheless, for ships without the capability to transmit metadata to shore either due to the lack of an electronic logbook or because an electronic logbook lacks an Admin message functionality, the delayed mode collection of metadata by PMO's remains the only way to collect and submit metadata. Every effort should be made by National Agencies to enable the automation of this process.

3.5 XBT Status/Actions

3.5.1 The XBT community was not very active in responding to the initial META-T User Group surveys. The Chair of the Data Management Program area of JCOMM, Robert Keeley, prepared a summary of the metadata required, which was then reviewed by the WMO Secretariat and the SOT Technical Coordinator.

3.5.2 Subsequently, in March 2008 a group of XBT experts convened in Miami to discuss the ramifications and solutions to the XBT Fall Rate Error problem. This group convened a side meeting to discuss the content of a new BUFR template. The results of this discussion formed an initial template that was then passed to the JCOMM Data Management Program Area Task Team on Table Driven Codes for further consideration and refinement. Ms Hester Viola described recent efforts concerning the refinement and approval of this template (SOT-V Doc-I-4, Section I-4.6 and Appendix G (report by the SOT Task Team on Coding)). Using this template in lieu of a User Survey, a similar gap analysis was carried out for the XBT data stream. However, this time instead of using only the FM-63 TAC message, the actual messages transmitted from ship to shore were known for the two most widely used XBT data collection systems (Devil and SEAS).

3.5.3 Fields in BUFR but not in at least one of the ship to shore transmissions:

- Unique ID: Tag that is part of the GTS message that uniquely identifies the message and is retained with the data always.
- Call sign: ITU Call sign
- IMO Number
- Ship Name: Current ship name helpful to match ship identification in Lloyds registry as sometimes ship names change out of sync with Call sign.
- SOOP line number: e.g. AX10, PX05 etc
- SOOP transect number: Number identifying a single occupation by a single ship of a SOOP Line (i.e. incremented at the beginning of each occupation).
- Platform speed
- Platform direction
- Observation time: Some formats do not report seconds, others do
- Position of observations: Differing accuracies in reports, BUFR requirements (10^-3)
- Indicator for digitization: Devil BOM and Devil CSIRO only reports selected depths. The Devil software offers several options for digitization however, the method is not reported
- Total depth of water
- QC Indicator: Global indicator of depth measurements quality
- QC Indicator: Global indicator of temperature measurements quality
- Level by level QC flags for Temperature
- Height of XBT/XCTD launcher
- Program operating SOOP ship
- Launcher type
- Data acquisition software type and version
- Probe serial number
- Probe manufacturing date
- Drop number: sequential drop along the current transect.

3.5.4 Additional missing Meteorological observations:

3.5.4.1 In addition to the metadata listed above that are pertinent to the XBT system alone, the current BUFR template follows the previous FM-63 TAC message by including fields for an abbreviated meteorological observation to be included with the XBT observation. In this case, many of the same issues that the VOS community is addressing are relevant to the XBT data stream. We note that currently no XBT messages used for ship to shore transmission contain this Met data, however the META-T team felt that the inclusion of Met observations with XBT observations could be useful for multiple applications and invites the XBT operators to consider re-introducing this capability into their operational practices (ACTION: SOOP Operators and Ship-to-shore software developers.).

3.5.4.2 The solutions for augmenting the XBT metadata collection are similar to the VOS solutions. XBT data acquisition software should be improved to collect and record more metadata and to transmit as much as possible from ship to shore. For those elements that are not transmitted to shore, the data acquisition software should record the information in a standard format that can be submitted once retrieved on shore. Additionally, the concept of an Admin message summarizing the metadata describing the current state of a ship is valid for the XBT data stream as well. The Admin message represents a good compromise between augmenting the size of the data message and relying solely on manual delayed mode submission of metadata, which will invariably introduce time delays in the metadata availability.

3.5.4.3 For example, the following sequence of events would enable the collection of all available metadata and data during the course of one high-density XBT cruise:

1. Prior to the cruise, an Admin message is transmitted stating the physical configuration of the XBT launcher on the ship as well as programmatic elements of the ship and the operator,
2. Each drop is transmitted to shore in a reduced format,
3. Each drop is recorded locally in a complete format (preferably an internationally agreed-upon format),
4. At the end of the cruise another Admin message is sent duplicating the first and providing "endpoints" to the valid dates of the first Admin message,
5. A "Cruise Summary" document (essentially a more detailed Admin message) is automatically produced that contains all available metadata from the pre-cruise Admin message, as well as the pertinent metadata in the XBT data files that enable the data files to be permanently linked to the metadata description (e.g. filenames, unique IDs, drop number etc). This file would be exchanged in an internationally developed standard, such as SensorML.

4. Summary and Recommendations

4.1 The META-T Pilot Project has grappled with creating a metadata framework for all the JCOMM platforms. It was hoped that by limiting the scope of the pilot project to only platforms sensing temperature, that progress could be made. However, in retrospect, the more fundamental organizing entity is the platform type (including the international program coordinating the use of that platform type) and not the variables measured. For this reason, as they are both ship-based systems, the VOS and SOOP programs have similar issues and solutions when it comes to collecting and managing metadata.

4.2 The above discussion serves to underscore one important point:

- In order to satisfy the needs of the User Surveys, the operational data collection procedures for both VOS and SOOP will need to be modified.

4.3 META-T invites participation from the SOT members in any way that advances this goal. Below are a few specific areas where SOT expertise would prove valuable.

(i) Expertise in the content and collection of metadata for Pub-47.;the real world issues involved in its maintenance and the community of users that rely on it,
(ii) Expertise in the electronic logbooks or XBT/TSG data acquisition software,
(iii) Expertise in the installation and maintenance of the hardware on VOS ships,
(iv) Expertise in the delayed mode data streams for either VOS or SOOP.

_____