COMMON ISSUES OF QUALITY CONTROL OF SURFACE MARINE DATA

DRAFT, 7 March 2008, Chairs: ETMC, DMCG, SAMOS, GOSUD, et al.

Joint World Meteorological Organization (WMO)/Intergovernmental Oceanographic Commission (IOC)
Technical Commission for Oceanography and Marine Meteorology (JCOMM)
Data Management Programme Area (DMPA)

[Note: Based on Doc 5.3 rev1 Appendix for DMCG-III, Ostend, Belgium, 26-28 March 2008: http://www.jcomm.info/index.php?option=com_oe&task=viewDocumentRecord&docID=1838].

Contents

1. Introduction

There are three large programs currently operating that collect surface marine data (meteorological and oceanographic) from Voluntary Observing Ships (VOS) and Research Vessels (R/Vs). There are some overlaps in the variables collected, but as yet there has been only initial discussions to bring some consistency to how these observations are handled. In the long term, it should be the goal that a particular observation of a marine variable, such as sea surface temperature (SST), should be handled in a consistent manner taking into proper consideration the observation techniques irrespective of the particular program making the observation. The purpose of this document is to initiate the discussion on standardizing quality control (QC[1]) of marine variables within JCOMM (2007). This limited discussion will inevitably highlight related topics where some consolidation of approach will be beneficial. These will be noted but not necessarily pursued.

VOS data are reported via long-established data channels managed by JCOMM or WMO. In real-time (RT) VOS data are reported via the Global Telecommunication System (GTS) (SHIP code; FM 13); and in delayed-mode (DM) under the Marine Climatological Summaries Scheme (MCSS), usually in the International Maritime Meteorological Tape (IMMT) format. The VOS Climate (VOSClim) project, which is based on a selection of ~200 ships within the overall VOS scheme, aims to provide a high-quality subset of marine meteorological data, with extensive associated metadata, to be available in both real-time and delayed-mode to support global climate studies.

The QC of data from two additional projects, in each case flowing (largely) outside the regular VOS program, will also be considered:

- Shipboard Automated Meteorological and Oceanographic System (SAMOS) project. This project operates a Data Assembly Center (DAC) to collect computerized high-resolution (typically 1-minute average) underway meteorological and near-surface oceanographic data from R/Vs, which are transmitted to the DAC daily via an established e-mail protocol. SAMOS observations are made using instrumentation installed primarily to support shipboard science and these systems are often (but not always) independent of the instruments used by the ship's crew for vessel operations.
- Global Ocean Surface Underway Data Pilot Project (GOSUD) project. This project handles near-surface oceanographic (e.g., salinity) data taken by ships, including VOS or R/Vs, primarily using the thermosalinograph (TSG) instrument. Some data from this project are

---

1 QC, which has been described as "operational techniques and activities that are used to fulfill given requirements for quality," is sometimes distinguished from broader quality assurance (QA): "all planned and systematic actions necessary to provide adequate confidence that a product, process or service will satisfy given requirements for quality" (ISO 1994). For simplicity, this document uses "QC" to collectively describe quality management actions of both forms.

circulated over the GTS in RT (TRACKOB code; FM 62), but others are not. A related complication (but also well known attribute of much oceanographic data) is that water samples typically are analyzed in DM to calibrate the TSG salinity values. Similarly to SAMOS, these data may fall largely outside established VOS or other JCOMM observational channels, but are gathered into a system of Global Data Assembly Centers (GDACs).

Because those two projects share some important characteristics, one joint workshop has been held (GOSUD/SAMOS 2006), with another planned for 10-12 June 2008. Possible QC, data reporting, and metadata convergences have been discussed.

One noteworthy data flow complication is that ships providing data for SAMOS/GOSUD (and similar projects) can also make regular VOS (or VOSClim) "bridge" reports (in DM and/or via GTS). Encouraging these quasi-independent observations was one recommendation from GOSUD/SAMOS (2006). Another data crossover consideration is that increasing numbers of VOS are being equipped with Automated Weather Systems (AWS). AWS data can share characteristics with SAMOS or GOSUD data, including limited formalization of their data flows at present within JCOMM. For simplicity, therefore, AWS data are considered adequately represented in this document by VOS plus the R/V data.

Table 1. Primary variables that may be estimated (E) or measured (M/A) by VOS, SAMOS, and GOSUD. "M" generally indicates partial automation, e.g., such that the measurement is instrumented but further collection and encoding may be manual (e.g., SST bucket measurement). "A" indicates a more advanced level of automation, with data typically reported at high temporal frequency (e.g., 1-minute averages).

Variables	*VOS*[1]	*SAMOS*[2]	*GOSUD*[3]
observation date & time	M	A	(E/M)[4]
latitude & longitude	M	A	(E/M)[4]
ship heading; course & speed (over ground)	(E/M)[5]	A	
ship speed & course (true, over water)	E/M	A	
sea surface temperature (SST)	M	A	A
air temperature (AT)	M	A	
moisture (DPT/WBT, RH, &/or specific humidity)	M	A	
relative wind speed & direction	(E/M)[5]	A	
true wind speed & direction	E/M	A	
visibility	E	A[7]	
present & past weather	E		
sea level pressure (SLP) (& for VOS, tendency)	M	A	
wind wave (direction[6]) period, height	E/M	A[7]	
swell direction, period, height	E	A[7]	
cloud cover & height	E	A[7]	
precipitation	(E/M)[7]	A	
shortwave & longwave radiation		A	
salinity		A	A
conductivity		A	A[8]

1. Secondary variables may include: secondary cloud type and swell fields, ice accretion, sea ice concentration, etc.
2. Secondary variables may include: photosynthetically active, ultraviolet, and total radiation, radiometric SST, etc. (ref.: http://samos.coaps.fsu.edu/html/parameters.php).
3. Since GODUD collects sea temperature and salinity data through the hull pumped water, the sample is collected at some level actually below the water line of the vessel. Some other variables, such as fluorescence, pCOs, and pH, may also be collected, but because of limitations in the character code form (TRACKOB) they are not generally reported. In addition to sea surface temperature and salinity, surface current and direction may also be reported in the TRACKOB code.
4. The mechanism of attaching time and position to GOSUD observations may be installation specific (to be clarified).
5. Not generally reported by VOS except under the VOSClim Project.
6. No longer reported as part of the SHIP code since 1968.
7. Rarely reported, but automated instruments exist for visibility, waves and swell, and ceiling (cloud base) height.
8. Conductivity is measured for GOSUD, but is not reported in RT, and may not be archived in DM (to be clarified).

Table 1 provides an initial variable list to help delineate the scope of the discussion. Following a brief background description of the common characteristics of many QC procedures (sec. 2), we consider the QC applied to this selection of ship-based data streams in RT and DM (secs. 3-4), and at higher levels as applied toward the development of operational weather predictions and climate products (sec. 5). Conclusions and recommendations for action are presented in sec. 6. Annex A provides a list of acronyms and links to website resources, and Annex B discusses in more detail QC issues specific to the International Comprehensive Ocean-Atmosphere Data Set (ICOADS).

2. Common QC Characteristics

Tests for coding, reporting, and transmission errors, physical validity, and the climatological reasonableness of data can be implemented at many different stages of data collection, and in many different forms (manual, automated, etc.). In order to establish common ground for discussion across different communities, in the following we aim to clearly define some typical QC tests (not necessarily common to all or even any of the ship-based data programmes presently), and also list some important related statistical considerations and procedures:

- FL (Field Legality): syntax-type errors in data or metadata fields (arising e.g., from coding, reporting, and transmission errors).
- UP (Univariate Plausibility): gross physical range checks, e.g., SST<−5°C or AT>58°C.
- UT (Univariate Tracking): consistency checks for plausible rates of change/persistence within a time series.
- MP (Multivariate Plausibility): two or more data elements failing a physical relationship test, e.g., AT < DPT temperature, or wind direction and speed values/codes inconsistent.
- PT (Platform Tracking): spatial and temporal track checks.
- PL (Platform "Landlocked"): checks for position erroneously reported on land areas.
- PN (Platform Neighboring): comparisons with data from nearby platforms (e.g., Kent and Berry 2005), or with co-located satellite (e.g., O'Carroll et al. 2006) or model output data.
- CU (Climatological Univariate): e.g., Wolter 1997.
- CM (Climatological Multivariate): e.g., wind bivariate statistical checks or joint wind/pressure climatological checks (theoretical options, but implementation may be problematic).

Important related procedural and statistical considerations:

- The merits in different situations of correcting (modifying), versus flagging, data.
- Random and systematic errors, versus bias.
- Probabilities (related to statistical trimming problems) of rejecting good data or accepting false data (e.g., WMO 1993, 2006, Wolter 1997).

Associated or potential procedures:

- Delayed-mode instrument calibrations (e.g., for GOSUD TSG).
- Data "preconditioning," including checks for the legitimacy of platform and ID types.
- Duplicate elimination (dupelim) (e.g., Slutz et al. 1985, Supp. K).
- Uses of ancillary platform or instrumental metadata to help validate data (e.g., Kent et al. 2006).
- Historical data bias corrections.
- "Complex" QC associated with data assimilation (e.g., Ingleby and Lorenc 1993), objective data analysis (e.g., Eischeid et al. 1995), etc.

3. Real-time QC

The temporal divisions between real-time, near–real-time, and delayed-mode processing can be somewhat arbitrary and difficult to establish. Here we will simply define RT (or near–real-time) QC as that applied shipboard. All subsequent QC that is applied after the data are transmitted off the ship or downloaded to another site (generally on shore) will be covered under secs. 4-5.

3.1 VOS

Contemporary VOS (and VOSClim) data prepared shipboard and transmitted over GTS in RT (or near–real-time) are subject to a variety of QC procedures, depending on national (including commercial shipping) practices. The highest quality is probably assured through the use of electronic logbooks, such as TurboWin, OBSJMA, and SEAS. These electronic systems can assist the manual observer with compiling and encoding the observations, preparing properly formatted messages for transmission of the data over GTS, and storage on-board (or delivered to shore in DM) of IMMT reports, or other data forms that can be subsequently compiled into the IMMT format. For example, in addition to ensuring that the data are properly encoded into GTS reports, TurboWin assigns the IMMT QC flags (Table 2). So as to achieve better observational consistency and enhance data quality, systematic inter-comparisons of the electronic logbook systems have recently been recommended (ETMC-II, SOT-IV).

Table 2. Defined settings in the IMMT-3 format[1] available for the "Indicator of test procedures" and for the individual "Data quality indicators."

Flag	*Indicator of test procedures (position: 82)*	*Flag*	*Data quality indicators for individual elements (Q_1-Q_{29}) (positions: 113-132 and 153-159)*
0	No quality control (QC)	0	no QC has been performed in this element
1	Manual QC only	1	QC has been performed; element appears to be correct
2	Automated QC only /MQC (no time-sequence checks)	2	QC has been performed; element appears to be inconsistent with other elements
3	Automated QC only (inc. time sequence checks)	3	QC has been performed; element appears to be doubtful
4	Manual and automated QC (superficial; no automated time-sequence checks)	4	QC has been performed; element appears to be erroneous
5	Manual and automated QC (superficial; including time-sequence checks)	5	The value has been changed as a result of QC
6	Manual and automated QC (intensive, including automated time-sequence checks)	6-8	Reserve
7-8	Not used		
9	National system of QC (information to be furnished to WMO)	9	The value of the element [is] missing

1. Ref.: http://goos.kishou.go.jp/ws/ETMC/code_task/cmm/JCOMM2Rec9-ann1.pdf. The actual terminology used in the IMMT-3 format documentation is "QC indicator" (for the Indicator of test procedures) and "QC indicator for [data element name]" (for the Data quality indicators). We suggest that consideration be given to the possibility of clarifying that documentation along the lines used in this table as part of any future IMMT updates.

3.2 SAMOS

Real-time QC is not generally applied to automated observations collected on research vessels. Gross errors caused by system or instrument failures may be remedied on the vessel, but data are rarely flagged to mark such occurrences. The data are transmitted in an "as is" form and QC occurs at the shore-side data centers. There is a broad interest among research vessel operators to have access to real-time QC algorithms.

3.3 GOSUD

Similarly to SAMOS, real-time QC operations may be applied on a ship-by-ship basis. In the future, the activities of the QARTOD (2006) project (see also Annex B) may also become more relevant to the real-time QC of oceanographic data. The recent IODE/JCOMM Forum on Oceanographic Data Management and Exchange Standards discussed QC procedures for surface temperature and salinity observations. The procedures were specific to TSG instrumentation and so do not cover other techniques and characteristics of collecting such data. For that reason, no immediate recommendation to consider procedures for a standard are to be taken, though the GOSUD project was encouraged to update procedures as appropriate taking into consideration other published practices.

4. Delayed-mode QC

In this section we discuss QC that is applied after the ship reports have been transmitted or collected from the ship to shore, for example performed by individual countries contributing to the VOS scheme, at JCOMM Global Collecting Centers (GCCs), and at project or global data assembly centers (DAC or GDACs).

4.1 VOS

Under the JCOMM Marine Climatological Summaries Scheme (MCSS), contemporary VOS (and VOSClim) data reported on logbooks (paper or electronic) are to be compiled by recruiting countries into the IMMT format, and subjected by those Contributing Members (CMs) to the Minimum Quality Control Standard (MQCS); software (MQC) is also available for this purpose. The rationale behind MQCS is that data errors can be most easily rectified nearest the data source by CMs, which extends to potential feedback to Port Meteorological Officers (PMO), and so that the quality of VOS data from individual ships can generally be improved. The intent behind this initial phase of the scheme seems useful, but we note that not all countries have the resources to apply or fully apply MQCS, including some countries that maintain their separate QC systems (e.g., USA).

Once compiled (and ideally MQC'd) by the recruiting countries, the IMMT data are forwarded to the JCOMM Global Collecting Centres (GCCs; Germany and UK), subjected or re-subjected to MQCS, and thence re-distributed back to eight Responsible Members (RMs) for archival and distribution. The whole MCSS system (established c. 1963) also prescribes the publication of climatological (decadal) tabular/graphical summaries based on the VOS data, and is slated for modernization by JCOMM and its Expert Team on Marine Climatology (ETMC) under two newly proposed task teams (TT-DMVOS and TT-MOCS).

4.2 SAMOS

For each participating ship, a set of 1-minute observations recorded for the previous day arrive at the DAC soon after 0000 UTC, and undergo automated QC evaluation (based on Smith et al. 1996, Smith and Legler 1997), with output flag possibilities as listed in Table 3. Additional (unpublished) automated statistical routines are employed to identify suspect observations. In addition, a trained Data Quality Evaluator (DQE) reviews the data and QC results and responds directly to vessels at sea when problems are identified. All quality-evaluated data are freely available to the user community and are distributed to national archive centers. At present, none of these data are transmitted on the GTS (because of timeliness and logistical issues).

Table 3. QC flag possibilities available from the SAMOS procedure.[1] Note: In past DAC projects, some data arrived in DM with pre-existing QC. This is why some flags indicate QC completed outside of the DAC. Those flags (Q, R) have not been used during the SAMOS project.

Flag	*Meaning*	*Flag*	*Meaning*
A	Original data had unknown units. The units shown were determined using a climatology or some other method.	N	Signifies that the data were collected while the vessel was in port. Typically these data, though realistic, are significantly different from open ocean conditions.
B	Original data were out of physically realistic range bounds outlined.	O	Original units differ from those listed in the original_units variable attribute. See quality control report for details.
C	Time data are not sequential or date/time not valid.	P	Position of platform or its movement are uncertain. Data should be used with caution.
D	Data failed $T \le T_w \le T_d$ test. In the free atmosphere, the value of the temperature is always greater than or equal to the wet-bulb temperature, which in turn is always greater than or equal to the dew point temperature.	Q	Questionable - data arrived at DAC already flagged as questionable/uncertain.
E	Data failed resultant wind re-computation check. When the data set includes the platform's heading, course, and speed along with the platform relative wind speed and direction, a program re-computes the earth relative wind speed and direction and compares the computed values to the reported earth relative wind speed and direction. A failed test occurs when the wind direction difference is > 20° or the wind speed difference is >2.5 m/s.	R	Replaced with an interpolated value. Done prior to arrival at the DAC. Flag is used to note condition. Method of interpolation is often poorly documented.
F	Platform velocity unrealistic. Determined by analyzing latitude and longitude positions as well as reported platform speed data.	S	Spike in the data. Usually one or two sequential data values (sometimes up to 4 values) that are drastically out of the current data trend. Spikes occur for many reasons including power surges, typos, data logging problems, lightning strikes, etc.
G	Data are greater than 4 standard deviations from the COADS climatological means (da Silva et al. 1994). The test is only applied to pressure, temperature, sea temperature, relative humidity, and wind speed data.	T	Time duplicate.
H	Discontinuity found in data.	U	Data failed statistical threshold test in comparison to temporal neighbors. This flag is output by automated Spike

			and Stair-step Indicator (SASSI) procedure developed by the DAC.
I	Interesting feature found in data. More specific information on the feature is contained in the data reports. Examples include: hurricanes passing stations, sharp seawater temperature gradients, strong convective events, etc.	V	Data spike as determined by SASSI.
J	Data are of poor quality by visual inspection, DO NOT USE.	(W)	(currently unused)
K	Data suspect/use with caution – this flag applies when the data look to have obvious errors, but no specific reason for the error could be determined.	X	Step/discontinuity in data as determined by SASSI.
L	Oceanographic platform passes over land or fixed platform moves dramatically.	Y	Suspect values between X-flagged data (from SASSI)
M	Known instrument malfunction.	Z	Data passed evaluation.

1. Ref.: http://samos.coaps.fsu.edu/html/docs/samos_quality_flag.pdf

4.3 GOSUD

GOSUD (2003) outlines a series of "real-time" QC checks (Table 4) that are proposed for application to the trajectory data at the earliest opportunity, so that the data can be assimilated by modeling centers, plus to help provide feedback to the ships and TSG operators. These specified processing includes tests for code legality, landlocked reports, impossible ship speed, global and regional temperature and salinity physical ranges, spikes, gradients, and instrument cross-comparisons.

Table 4. QC flag possibilities documented by GOSUD (2006, Reference table 2). GOSUD (2006) also documents a variety of other QC-related information including "data state" indicators (raw vs. processed data), QC "history action codes," and QC "test IDs" (impossible date/location, position on land, stuck value, etc.).

Flag	Meaning	Real-time comment
0	No QC was performed	—
1	Good data	All GOSUD real-time QC tests passed.[1]
2	Probably good data	—
3	Bad data that are potentially correctable	Test 16 failed and all other tests passed. These data are not to be used without scientific correction.
4	Bad data	Data have failed one or more of the tests, excluding Test 16.
5	Value changed	Data may be recovered after transmission error.
6	Not used	—
7	Not used	—
8	Interpolated value	Missing data may be interpolated from neighbouring data e.g. trajectory location.
9	Missing value	—

1. Delayed-mode comment: "The adjusted value is statistically consistent and a statistical error estimate is supplied."

5. Metadata sources and QC

If available, many selections of platform and instrumental metadata (e.g., ship size and type parameters, observational platform and anemometer heights, and instrument types and exposures) can be used indirectly by data centers or individual researchers to help improve data quality and enhance QC processing, typically in delayed- or research-mode. Examples include the validation of reported instrument/measurement types, against types expected based on separately maintained ship-specific metadata holdings.

5.1 VOS

For decades, WMO (1955) Publication No. 47 has provided an extensive collection of platform and instrumental data describing characteristics of the VOS fleet (and VOSClim ships). This international publication also forms a useful target and standard for comparison with other projects and components of the marine observing system, such as for buoys and other automated Ocean Data Acquisition Systems (ODAS), which thus far have not fully succeeded in systematizing the international collection and archival of such metadata (particularly applicable to historical ODAS metadata). Publication No. 47 metadata were published annually for many years, and transitioned recently to quarterly. Kent et al. (2006) reviewed the expanding and evolving contents of this important resource, and describe how, for the convenience of climate applications and individual researchers, selected Publication No. 47 metadata elements have blended into ICOADS by means of a defined “attachment” to the International Maritime Meteorological Archive (IMMA) format.

5.2 SAMOS

SAMOS developed a new metadata specification tailored to that project’s unique requirements, with input from VOSClim, JCOMM, and a variety of other programs involved with marine metadata standards. Upon recruitment, each SAMOS vessel completes a set of defined metadata forms, which is publicly accessible to through a ship profile database maintained at the DAC. Similarly to WMO Publication No. 47, the ship-specific metadata include general information about the vessel, vessel dimensions, and contacts for the original data provider. The parameter-specific metadata lists all measurements being provided by a vessel and allows the user to sub-select information on the variables, units, averaging methods, and instrumentation. Digital imagery includes photos of each vessel and instrument masts and also contains schematics for each vessel (following the approach and naming conventions of VOSClim). In early 2008, SAMOS will launch an on-line tool that will allow vessels participating in SAMOS to update their ship metadata profiles. Many of the metadata elements are also embedded in the SAMOS netCDF format, which allows distribution of a single metadata inclusive data file to users.

The SAMOS project takes advantage of high-quality metadata to improve the QC of observations. For example, digital imagery of the instrument exposure provides a critical resource for DM QC. On some vessels, the exposure of the AT and RH sensors can be impacted by exhaust from the ship’s stack (more frequently an issue on RVs that conduct maneuvers for science or station work). By noting the sensor location relative to the stack, temperature and RH values can then be flagged according to the ship-relative wind direction. Over time this DM QC can be automated to flag all temperature and RH data when the ship-relative wind is from a specific set of wind angles. These ship-specific criteria for flagging observations have the potential to become RT QC for that vessel (assuming sensor locations remain constant).

5.3 GOSUD

GOSUD (2006) provides defined spaces in its netCDF format for a variety of general information about the trajectory file, the platform (e.g., WMO code and originating data center), locations and measurements from the platform (e.g., sampling method, calibration equation, time lag, and depth of water intake), and measurement cycle and detailed QC and other history information. At this stage the GOSUD project seems to lack capabilities for centralized metadata repositories on a ship-by-ship basis, such as maintained through WMO Publication No. 47, and by SAMOS.

6. Higher-level QC (HQC)

Higher-level” QC (HQC) will be defined here as that typically applied to surface marine data during the development of products, ranging from short-range weather predictions, to global atmospheric reanalyses, and other climate-scale products. HQC processing is also applied during the creation of climate databases such as ICOADS, which supplies marine data for many climate-scale applications.

6.1 QC for numerical weather prediction (NWP)

Surface marine data, and many other types of data (e.g., from satellites) are assimilated operationally into weather prediction models. Data assimilation processing typically includes a variety of QC steps, in some cases extending to manual screening and setting QC flags in the data, such as was performed by NOAA/NCEP under their “CREWSS” (formerly QUIPS) system. Similarly, the European Meteorological Network (EUMETNET) Composite Observing System (EUCOS) Surface Marine Programme provides some web-based quality tools to monitor the quality of observations from VOS both inside and outside the EUCOS area of interest (e.g., monthly statistics of comparisons with model outputs).

Related to these types of operational QC, the UK Met Office shares in WMO coordinated monitoring of the Global Observing System, by acting as lead centre for monitoring the quality of surface marine observations (ships, drifting buoys, moored buoys and other fixed marine platforms). The Met Office also holds a monitoring role in the VOS scheme and is the Real-Time Monitoring Centre (RTMC) for the VOSClim project. Specifically, upon receipt of each GTS report from a ship within that project, the RTMC appends six co-located parameters from the Met Office forecast model (SLP, relative humidity, AT, SST, and wind speed and direction). In addition the RTMC has responsibility for monitoring the quality of the VOSClim observations against defined criteria, which are set at a tighter level than for standard VOS. These quality monitoring statistics, together with lists of those ships whose observations have failed to meet the required criteria, are then relayed to the VOSClim DAC and to the national VOSClim focal points in order that any identified problems can be addressed.

6.2 QC for reanalyses and other climate-scale products

Global atmospheric reanalyses have data ingest requirements similar to NWP, typically relying on satellite and upper-air (radiosonde) data (e.g., Kalnay et al. 1996). However, newer forms of historical reanalysis are also being initiated that utilize surface observations only (e.g., Compo et al. 2005). Because they usually operate with a substantial delay, reanalyses can take advantage of more complete and higher-quality DM data and metadata sources. Since they use models similar to those for NWP, reanalyses can produce and archive large amounts of QC information about the individual input observations. At present, however, little effort has been invested in systematically associating this information back to the original observations (e.g., in ICOADS; Chang 2007)

6.3 QC for climate-scale databases

Many different ship and other surface marine data sources, both contemporary and historical, have been blended together into the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) (Worley et al. 2005, Woodruff et al. 2005). ICOADS is both a collection of climate-quality marine data (individual observations), plus year-month summaries[2]. The ICOADS individual observations have all been subjected to a complex system of processing, including the following two primary QC procedures, and with flag possibilities as summarized in Table 5:

- NCDC-QC: This processing was originally designed by NOAA/NCDC primarily around the QC characteristics of VOS data in conjunction with the publication of marine atlases (e.g., US Navy 1981), and updated for COADS Release 1 (Slutz et al. 1985, Supp. J). Weather elements are checked for legality, intercompared for consistency, and subjected to limited climatological checks.
- Trimming: Observations outside 3.5σ (standard deviations) relative to three distinct climatological periods (1854–1909, 1910–49, 1950–79) are "trimmed" and not used (Slutz et al. 1985, Supp. C). Subsequently, these Release 1-based trimming limits were found to be too restrictive (e.g., Wolter 1997) for some extreme climate events (e.g., 1982–83 El Niño), and a variety of ad hoc steps have been taken to partly address these and other problems in the trimming, including the establishment of rules for trimming data before 1854 and after 1979.

Neither of these QC procedures has been substantially updated in many years, and full renovation of the ICOADS QC remains an important outstanding problem. Annex B provides additional details about the status, and some additional known deficiencies, associated with these and other QC-related portions of ICOADS processing.

Table 5. Flag possibilities output by the two primary ICOADS QC procedures.

NCDC-QC			Trimming	
Flag	Meaning	Reason	Flag	Description
R	correct	–	1	within 2.8σ limits ($g–2.8s_1 \le a_1 \le g+2.8s_5$)
A	correctable	legality	2	less than 2.8σ lower limit ($g–3.5s_1 \le a_1 < g–2.8s_1$)
B	correctable	internal consistency	3	greater than 2.8σ upper limit ($g+2.8s_5 < a_1 \le g+3.5s_5$)
J	suspect	internal consistency	4	less than 3.5σ lower limit ($g–4.5s_1 \le a_1 < g–3.5s_1$)
K	suspect	time	5	greater than 3.5σ upper limit ($g+3.5s_5 < a_1 \le g+4.5s_5$)
L	suspect	extreme (outside 4.8σ)	6	less than 4.5σ lower limit ($a_1 < g–4.5s_1$)

2 Ten statistics (such as the mean and median) are calculated for each of 22 observed and derived variables, using 2° latitude x 2° longitude boxes back to 1800 (and 1°x1° boxes since 1960).

M	erroneous	legality	7	greater than 4.5σ upper limit ($a_1 > g+4.5s_5$)
N	erroneous	internal consistency	11	limits missing; data correct (for SST & SLP only)[1]
Q	erroneous	extreme (outside 5.8σ)	12	limits missing (ocean/coastal box)
S	missing	–	13	landlocked 2° box
			14	data unusable (for SST, AT, & SLP only)
			15	data missing or not computable

1. Drifting buoy data (from ISDM) flagged correct by external procedure.

Other examples of climate databases include the World Ocean Database (e.g., WOD05; Boyer et al. 2006) of oceanographic data. Near-surface profile temperatures from versions of the WOD database are in turn blended into ICOADS. These databases also feed into reanalysis, other climate products (e.g., blended or reconstructed SST products; Woodruff et al. 2008), and a wide variety of other scientific applications. SAMOS and other R/V data are also beginning to feed into ICOADS, in the form of sub-sampling the high temporal resolution (1-minute average) data into the IMMA format (i.e., as hourly observations).

6.4 QC to create ocean climate products

Groups such as the UK National Oceanography Centre, Southampton (NOCS), and the Center for Ocean-Atmospheric Prediction Studies (COAPS), have applied QC beyond the ICOADS procedures to develop in situ based climatologies (e.g., flux fields). In some cases NOCS and COAPS have implemented bias corrections. Additional procedures have identified problems in platform IDs/type codes and consistency in the tracks of individual platforms. There is currently no mechanism to feed this DM QC back into ICOADS, etc. The topic of retaining these DM QC in ICOADS will be discussed at the CLIMAR-III Workshop in May 2008.

7. Conclusions and recommendations

The preceding sections provide an initial overview and survey of some common issues and characteristics of QC processing applied to VOS and R/V data, specifically associated with the SAMOS and GOSUD projects, with the broader aim to help initiate the process of standardizing marine and oceanographic observational QC within JCOMM. In accordance with that aim, following is a series of recommendations intended to outline a process for proceeding with a more detailed and inclusive analysis, and ultimately the proposed convergence towards some JCOMM QC standards (JCOMM 2007b, Rec. 5.2c):

- QC procedures can be very complex, in some cases with important details captured exclusively in software rather than readily accessible documentation. These initial results have necessarily only been able to scratch the surface of the common characteristics and differences between QC procedures (e.g., Table 6), including their current procedural interrelationships, e.g., as reflected in output ICOADS data (Table 7).
- Partly these results have surveyed one relatively simple but concrete characteristic of QC procedures, i.e., in terms of their currently widely divergent output flag configurations (Tables 2-5). This characteristic is related to JCOMM (2007) recommendations 5.2a and 5.2b suggesting the need to work towards better harmonization of QC flags. Any move towards standardized QC flags would however have to address the different needs of the operational (real-time) and research science (delayed-mode) data users.
- Convergence towards broader marine and oceanographic QC standards will probably be further strengthened by eventually expanding the analysis to other JCOMM ODAS data types, and the QC operations applied in RT or DM, such as oceanographic data (e.g., GTSPP 2002, Boyer et al. 2005), including Argo, and drifting and moored buoy data (e.g., NDBC 2003). However, adding the QC procedures currently associated with these and other ODAS data types will further complicate the task.
- For improved quality monitoring of SST, collaborative approaches with GHRSST-PP were suggested at ETMC-II (JCOMM 2007a, action 2.4.1.4) to use independent satellite data (e.g., O'Carroll et al. 2006) to monitor buoy and ship SST performance, define SST uncertainty estimates for individual buoys and ships, and provide guidance and assistance to the GHRSST-PP teams constructing satellite and in situ databases.

- Proposed QC standards will likely need to be knitted together with existing published WMO and emerging international quality management procedures (e.g., WMO 1992, 1993, 2003, 2006).
- So as to help enhance the identification of GTS and delayed-mode duplicates, an important QC-related improvement would be the introduction of unique report tracking number.
- Any convergence of real-time and delayed-mode QC standards must work in conjunction with improvements to data transmission formats. There is no sense in having RT QC flags that cannot be transported with the data via the GTS or other transport mechanisms.

Table 6. Comparison of selected global physical limits as applied by different QC procedures.

Variable	*SAMOS*[1]	*GOSUD*[2]	*MQCS-V*[3]	*ICOADS*[4]
Wind speed m s^{-1}	0 ≤ 40	—	0 ≤ 41.2 (80 kts)	0 ≤ 102.2
SST° C	0 ≤ 35[5]	–2.5 ≤ 45	–2.5 ≤ 37	–5.0 ≤ 40.0
AT° C	–10 ≤ 40[5]	—	–25 ≤ 40	–88.0 ≤ 58.0
SLP hPa	950 ≤ 1050	—	870 ≤ 1070	870.0 ≤ 1074.6
Salinity PSU	—	0 ≤ 60	—	—

1. Ref.: Smith et al. 1996. "Some of these out of bounds values, for example an air temperature of -15.0°C near the Antarctic Coast, are realistic and the bounds flag is removed by the DQE.
2. Ref.: GOSUD, 2003.
3. Ref.: http://goos.kishou.go.jp/ws/ETMC/code_task/cmm/JCOMM2Rec9-ann2.pdf. The SST and AT limits are not actually global physical limits, because they are applied with some dependencies on latitude (e.g., ≥45°).
4. Ref.: http://icoads.noaa.gov/e-doc/stat_trim. Wind data were considered "not computable" if some wind data existed but U and V did not result after application of Release 1, Table E2-1 (see http://icoads.noaa.gov/Release_1/suppE.html), including all cases where only wind speed resulted.
5. SAMOS uses physical limits that vary by latitude. The limits listed here are for the entire globe.

Table 7. General types of checks, per variable, progressively applied to VOS IMMT format data in RT (TurboWin), in DM (e.g., at the GCCs), and finally during two higher levels of ICOADS QC processing (NCDC-QC and Trimming) in IMMA format. Abbreviations established in sec. 2 for QC checks are used in this table (or "—" indicates the absence of any check):

- FL (Field Legality)
- UP (Univariate Plausibility)
- UT (Univariate Tracking)
- MP (Multivariate Plausibility)
- PT (Platform Tracking)
- PL (Platform "Landlocked")
- CU (Climatological Univariate)
- CM (Climatological Multivariate)

TurboWin outputs data in the IMMT format with the IMMT QC flags set, but according to a process that appears to be different than MQCS specifications. Note: "…" indicates that any TurboWin checks still need to be added to this table. Those IMMT flags may be overridden by MQCS processing. Moreover, at this time, IMMT flags settings are archived in the IMMA format but not interpreted in any way during ICOADS processing.

Variable	*TurboWin*[1]	*MQCS*[2]	*NCDC-QC*[3]	*Trimming*[4]
Platform ID (e.g., call sign)	FL	—	—	—
observation date & time	…	FL	FL	—
latitude & longitude	…	PT	UP, PL[5]	PL[5]
ship heading; course & speed (ground)	…	UP	—	—
ship speed & course (true, over water)	…	UP, MP	—	—
SST	…	UP, MP	UP, CU	CU
AT	…	UP, MP	UP, MP, CU	CU
moisture (DPT and/or WBT, or RH)	…	UP, MP	UP, MP, CU	CU
relative wind speed & direction	…	UP	—	—
true wind speed & direction	…	UP, MP	UP, MP, CU	CU
visibility	…	UP, MP	UP	—
present & past weather	...	UP, MP	UP, MP	—
SLP/tendency	…	UP, MP	UP, UT, CU	CU
wind wave (direction,) period, height	…	UP, MP	UP, MP, CU	—

swell direction, period, height	…	UP, MP	UP, MP, CU	—
cloud cover & height	…	UP, MP	MP	—
precipitation	…	UP, MP	—	—

1. Ref.: http://www.knmi.nl/turbowin/specifications/TurboWin_4_0_Quality_Control.htm
2. Ref.: http://goos.kishou.go.jp/ws/ETMC/code_task/cmm/JCOMM2Rec9-ann2.pdf
3. Ref.: Table J0-1, http://icoads.noaa.gov/Release_1/suppJ.html. A flag value "K" (see Table 5) that is available to indicate suspicious temporal consistency, e.g., based on platform track checking (Table J1-1, Ibid.) is not currently set by NCDC-QC processing except for pressure tendency.
4. Ref.: http://icoads.noaa.gov/Release_1/suppC.html. The moisture variables are trimmed based on a calculated relative humidity (RH), and wind based on testing computed vector wind components against separate limits for U and V.
5. For the two separate ICOADS QC procedures, two different landmask checks are also used: http://icoads.noaa.gov/mask.html.

Two additional QC-related recommendations associated with higher-level climate products and databases are (in part summarized from Woodruff et al. 2008):

- With a few exceptions most products to date have used SST in isolation from other variables (e.g., air temperature, barometric pressure, and winds), which are reported by VOS and to a more limited extent by ODAS. We anticipate that cross-validations between the different variables, and VOS and ODAS platform types, will lead toward the creation of improved climate-quality products for SST and other variables.
- A related data quality issue is to better capitalize in ICOADS on the QC feedback information obtained from NWP and reanalyses (e.g., Compo et al. 2006), as well as the creation of marine climate products (sec. 6.4). However, models and climate analyses can also be imperfect, especially when observations are sparse, and the resulting information would need to be carefully compartmentalized and documented within ICOADS.

References

Boyer, T.P., J.I. Antonov, H.E. Garcia, D.R. Johnson, R.A. Locarnini, A.V. Mishonov, M.T. Pitcher, O.K. Baranova, I.V. Smolyar, 2006: *World Ocean Database 2005*. S. Levitus, Ed., NOAA Atlas NESDIS 60, U.S. Government Printing Office, Washington, D.C., 190 pp., DVDs. [Available from: http://www.nodc.noaa.gov/OC5/WOD05/docwod05.html]

Chang, E.K.M., 2007: Assessing the Increasing Trend in Northern Hemisphere Winter Storm Track Activity Using Surface Ship Observations and a Statistical Storm Track Model. *J. Climate*, **20**, 5607–5628.

Compo, G.P., J.S. Whitaker, and P.D. Sardeshmukh, 2006: Feasibility of a 100-year reanalysis using only surface pressure data. *Bull. Amer. Met. Soc.*, **87**, 175-190.

da Silva, A., A.C. Young, and S. Levitus, 1994: *Atlas of Surface Marine Data 1994, Volume 1: Algorithms and Procedures*. Tech. Rep. 6, U.S. Department of Commerce, NOAA, NESDIS.

Eischeid, J.K., C.B. Baker, T.R. Karl, and H.F. Diaz, 1995: The Quality Control of Long-Term Climatological Data Using Objective Data Analysis. *J. Appl. Meteor.*, **34**, 2787–2795.

Eskridge, R.E., O.A. Alduchov, I.V. Chernykh, Z. Panmao, A.C. Polansky, and S.R. Doty, 1995: A Comprehensive Aerological Reference Data Set (CARDS): Rough and Systematic Errors. *Bull. Amer. Meteor. Soc.*, **76**, 1759–1775.

Gilhousen, D.B., 1987: A Field Evaluation of NDBC Moored Buoy Winds. *J. Atmos. Oceanic Technol.*, **4**, 94–104

GTSPP, 2002: Global Temperature-Salinity Pilot Project (GTSPP) Real-time Quality Control Manual. Intergovernmental Oceanographic Commission Manuals and Guides #22. [http://www.meds-sdmm.dfo-mpo.gc.ca/ALPHAPRO/gtspp/qcmans/MG22/guide22_e.htm]

GOSUD, 2003: User's Manual, Version 1.0. GOSUD Data Management, November 11th 2003, go-um-03-01. [Available from: http://www.ifremer.fr/gosud/gdac.htm]

GOSUD, 2006: GOSUD Real-time QC, Version 1.0. GOSUD Data Management, April 24th 2006, go-um-02-01. [Available from: http://www.ifremer.fr/gosud/gdac.htm]

GOSUD/SAMOS, 2006: Report of the 1st Joint GOSUD/SAMOS Workshop, 2-4 May 2006, Boulder, Colorado, USA. [http://www.coaps.fsu.edu/RVSMDC/marine_workshop3/docs/report_final.pdf]

Ingleby, N.B., and A.C. Lorenc, 1993: Bayesian quality control using multivariate normal distributions. *Quart. J.R. Meteor. Soc.*, **119**, 1195-1225.

ISO (International Organization for Standardization), 1994: *Quality Management and Quality Assurance—Vocabulary*. ISO Pub. 8402. [via: http://www.empa.ch/gaw/glossary.html]

JCOMM, 2007a: Second Session of the JCOMM Expert Team on Marine Climatology (ETMC), Geneva, Switzerland, 26-27 March 2007. JCOMM Meeting Report No. 50, 124 pp. (Web only; http://www.wmo.int/pages/prog/amp/mmop/documents/Jcomm-MR/J-MR-50_ETMC-II.pdf).

JCOMM, 2007b: JCOMM Data Management Strategy draft document (IOC/IODE-XIX/48; R. Keeley, 11

December 2007).
Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.
Kent, E.C., and D.I. Berry, 2005: Quantifying random measurement errors in Voluntary Observing Ships' meteorological observations. *Int. J. Climatol. (CLIMAR-II Special Issue)*, **25**, 843-856.
Kent, E.C., S.D. Woodruff and D.I. Berry, 2006: WMO Publication No. 47 metadata and an assessment of observation heights in ICOADS. *J. Atmos. Oceanic Technol.*, **24**, 214-234.
Kent, E., S. Woodruff, N. Rayner, T. Arbetter, C. Folland, F. Koek, D. Parker, R. Reynolds, R. Saunders, V. Smolyanitsky, S. Worley, and T. Yoshida, 2007: Advances in the use of historical marine climate data (Second International Workshop on Advances in the Use of Historical Marine Climate Data). *Bull. Amer. Meteor. Soc.*, **88**, 559-564.
NDBC (National Data Buoy Center), 2003: NDBC Technical Document 03-02, Handbook of Automated Data Quality Control Checks and Procedures of the National Data Buoy Center. [http://www.ndbc.noaa.gov/handbook.pdf].
O'Carroll, A.G., Eyre, J.R., Saunders, R.W., 2006: Three-way error analysis between AATSR, AMSR-E and in situ sea surface temperature observations. Submitted to *J. Atmos. Oceanic Technol.*
QARTOD (Quality Assurance of Real-Time Oceanographic Data), 2006: QARTOD-IV, Final Report: Fourth Workshop on the QA/QC of Real-time Oceanographic Data. [Available from: http://nautilus.baruch.sc.edu/twiki/bin/view]
Slutz, R.J., S.J. Lubker, J.D. Hiscox, S.D. Woodruff, R.L. Jenne, D.H. Joseph, P.M. Steurer, and J.D. Elms, 1985: *Comprehensive Ocean-Atmosphere Data Set; Release 1.* NOAA Environmental Research Laboratories, Climate Research Program, Boulder, CO, 268 pp.
Smith, T.M. and R.W. Reynolds, 2003: Extended reconstruction of global sea surface temperatures based on COADS data (1854-1997). *J. Climate*, **16**, 1495-1510.
Smith, S.R., C. Harvey, and M. Legler, 1996: Handbook of Quality Control Procedures and Methods for Surface Meteorology Data. World Ocean Circulation Experiment, Report No. 141/96, COAPS Technical Report No. 96-1. Data Assembly Center for Surface Meteorology, Center for Ocean Atmospheric Prediction Studies, Tallahassee. [http://www.coaps.fsu.edu/woce/docs/qchbook/qchbook.htm]
Smith, S.R. and D. Legler, 1997: Addendum: Handbook of Quality Control Procedures and Methods for Surface Meteorology Data. World Ocean Circulation Experiment, Report No. 141/96, COAPS Technical Report No. 96-1a. Data Assembly Center for Surface Meteorology, Center for Ocean Atmospheric Prediction Studies, Tallahassee. [http://www.coaps.fsu.edu/woce/html/qch_addm.htm]
US Navy, 1981: *Marine Climatic Atlas of the World. Vol. IX: World-wide Means and Standard Deviations.* NAVAIR 50-1C-65, U.S. Government Printing Office, Washington, D.C., 169 pp.
WMO, 1955: *International List of Selected, Supplementary and Auxiliary Ships*. WMO-No. 47. WMO: Geneva. (Serial publication; recently annual. Editions prior to 1966 were entitled *International List of Selected and Supplementary Ships*.).
WMO, 1992: Manual on the Global Data-processing System. WMO-No. 485, Volume I – Global Aspects [Appendix II-1, Minimum Standards for Quality Control of Data for Use in the GDPS].
WMO, 1993: Guide on the Global Data-processing System. WMO-No. 305.
WMO, 2003: *Manual on the Global Observing System*. Volume 1, WMO-No. 544 [Part V, Quality Control].
WMO, 2006: Guide to Meteorological Instruments and Methods of Observations. WMO-No. 8, Preliminary 7th edition [Part III, Chapter 3, Quality Management].
Wolter, K., 1997: Trimming problems and remedies in COADS. *J. Climate*, **10**, 1980-1997.
Woodruff, S.D., R.J. Slutz, R.L. Jenne, and P.M. Steurer, 1987: A comprehensive ocean-atmosphere data set. *Bull. Amer. Meteor. Soc.*, **68**, 1239-1250.
Worley, S.J., S.D. Woodruff, R.W. Reynolds, S.J. Lubker, and N. Lott, 2005: ICOADS Release 2.1 data and products. *Int. J. Climatol. (CLIMAR-II Special Issue)*, **25**, 823-842.
Woodruff, S.D., H.F. Diaz, S.J. Worley, R.W. Reynolds, and S.J. Lubker, 2005: Early ship observational data and ICOADS. *Climatic Change*, **73**, 169-194.
Woodruff, S.D., H.F. Diaz, E.C. Kent, R.W. Reynolds, and S.J. Worley, 2008: The evolving SST record from ICOADS. In *Climate Variability and Extremes during the Past 100 Years* (S. Brönnimann, J. Luterbacher, T. Ewen, H.F. Diaz, R.S. Stolarski, and U. Neu, Eds.), Advances in Global Change Research, Vol. 33, Springer, 65-83.

Annex A: Acronyms and Website Resources

Argo:
http://www.argo.net/
AWS: Automated Weather System
CLIMAR-III: Third JCOMM Workshop on Advances in Marine Climatology
http://icoads.noaa.gov/climar3/
CM: Contributing Member (under MCSS)
COADS: Comprehensive Ocean-Atmosphere Data Set (now ICOADS)
COAPS: Center for Ocean-Atmospheric Prediction Studies (at Florida State University)
http://www.coaps.fsu.edu/
CREWSS: Collect, Review, and Edit Weather data from the Sea Surface (NOAA/NCEP)
http://www.emc.ncep.noaa.gov/mmb/data_processing/data_dumping.doc/document.htm
http://www.nco.ncep.noaa.gov/pmb/qap/
DAC: Data Assembly Center
DM: delayed-mode
DQE: data quality evaluator (for SAMOS)
ETMC: Expert Team on Marine Climatology (DMPA/JCOMM)
http://icoads.noaa.gov/etmc/
EUCOS: EUMETNET Composite Observing System
http://www.meteo.shom.fr/vos-monitoring/ (VOS observational monitoring)
http://www.eucos.net
FSU: Florida State University
GCC: Global Collecting Centre
http://www.dwd.de/gcc
http://www.metoffice.gov.uk/research/interproj/gcc/index.html
GDAC: Global Data Assembly Center
GHRSST-PP: Global High-Resolution SST Pilot Project
http://www.ghrsst-pp.org/
GOSUD: Global Ocean Surface Underway Data Pilot Project
http://www.ifremer.fr/gosud/
GTS: Global Telecommunication System
http://www.wmo.int/pages/prog/www/TEM/GTS/gts.html
GTSPP: Global Temperature-Salinity Profile Program
http://www.meds-sdmm.dfo-mpo.gc.ca/meds/Prog_Int/GTSPP/GTSPP_e.htm
http://www.nodc.noaa.gov/GTSPP/
HQC: Higher-level Quality Control
ICOADS: International Comprehensive Ocean-Atmosphere Data Set
http://icoads.noaa.gov/
IMMA: International Maritime Meteorological Archive format
http://icoads.noaa.gov/e-doc/imma/
IMMT: International Maritime Meteorological Tape format (currently IMMT-3)
http://goos.kishou.go.jp/ws/ETMC/code_task/
IOC: Intergovernmental Oceanographic Commission
http://ioc.unesco.org/iocweb/index.php
IODE: International Oceanographic Data and Information Exchange (of IOC)
http://www.iode.org/
ISDM: Integrated Science Data Management (formerly MEDS, Canada)
http://www.meds-sdmm.dfo-mpo.gc.ca/meds/Home_e.htm
JCOMM: Joint WMO/IOC Technical Commission for Oceanography and Marine Meteorology
http://www.jcomm.info/
MCSS: Marine Climatological Summaries Scheme
MQC: Minimum QC software (currently MQC-3)
http://www.wmo.ch/pages/prog/amp/mmop/mqc_soft.html
MQCS: Minimum QC Standards (currently MQCS-V)
http://goos.kishou.go.jp/ws/ETMC/code_task/
Marine Climatology Wiki:
http://www.marineclimatology.net/
NCDC: National Climatic Data Center (NOAA)
http://www.ncdc.noaa.gov/oa/ncdc.html
NCEP: National Centers for Environmental Prediction (NOAA)

http://www.ncep.noaa.gov/
NDBC: National Data Buoy Center (NOAA)
http://www.ndbc.noaa.gov/qc.shtml
NOCS: National Oceanography Centre, Southampton UK
http://www.noc.soton.ac.uk/
NWP: Numerical Weather Prediction
QA: quality assurance
QARTOD: Quality Assurance of Real-Time Oceanographic Data
http://nautilus.baruch.sc.edu/twiki/bin/view
http://cdip.ucsd.edu/qartod/waves_qc
OBSJMA: Japan Meteorological Agency (JMA) electronic logbook
QC: quality control
ODAS: Ocean Data Acquisition System
PMO: Port Meteorological Officer
http://www.bom.gov.au/jcomm/vos/pmo.html
QUIPS: QUality Improvement Performance System (NOAA/NCEP; replaced by CREWSS)
RH: relative humidity
RM: Responsible Member (under MCSS: Germany; Hong Kong, China; India; Japan; The Netherlands; Russia; United Kingdom; and USA)
RT: real-time
RTMC: Real-Time Monitoring Centre (for VOSClim project; UK Met Office)
http://www.metoffice.gov.uk/research/nwp/observations/monitoring/marine/index.html
R/V: research vessel
RVSMDC: Research Vessel Surface Meteorology Data Center
http://www.coaps.fsu.edu/RVSMDC/
SAMOS: Shipboard Automated Meteorological and Oceanographic System (SAMOS)
http://samos.coaps.fsu.edu/html/
SASSI: Spike and Stair-step Indicator (SASSI) procedure (developed by SAMOS)
SEAS: Shipboard Environmental (data) Acquisition System (NOAA electronic logbook, etc.)
http://seas.amverseas.noaa.gov/seas/
SLP: sea level pressure
SSSOS: Sea Surface Salinity Observation Service
http://www.legos.obs-mip.fr/observations/sss/
SST: sea surface temperature
TAO: Tropical Atmosphere Ocean project QC
http://www.pmel.noaa.gov/tao/proj_over/qc.html
TT-DMVOS: Task Team on Delayed-mode VOS data (ETMC/DMPA and SOT/OPA, JCOMM)
http://www.wmo.ch/pages/prog/amp/mmop/documents/data_mgmt/tt-dmvos.html
TT-MOCS: Task Team on Marine-met. and Oceanographic Summaries (ETMC/DMPA, JCOMM)
TSG: thermosalinograph
TurboWin: electronic logbook (KNMI)
http://www.knmi.nl/turbowin/
VOS: Voluntary Observing Ship scheme
http://www.bom.gov.au/jcomm/vos/
VOSClim: VOS Climate Project
http://www.ncdc.noaa.gov/oa/climate/vosclim/vosclim.html
WMO: World Meteorological Organization
http://www.wmo.ch/pages/index_en.html

Annex B: Detailed QC Issues Related to ICOADS

Following are currently the major stages of QC, and related processing, for ICOADS:

- Translation into the IMMA format. In many cases, input QC flags are interpreted as part of this processing. Presently, however, IMMT flags are carried forward into the format, but not interpreted. Also, as a critical measure to help ensure archival data quality by guarding against the possibility of translation errors or omissions, the original (input format) data are preserved in an attachment (e.g., IMMT or FM 13 messages) to each individual marine report.
- Application of the NCDC-QC (Slutz et al. 1985, Supp. J) and trimming (Slutz et al. 1985, Supp. C).
- Data preconditioning: This step is used as the framework to set some data tracking flags, and make some known data corrections (e.g., deletion of known bad data sources).
- Duplicate elimination (dupelim): since frequently slightly different sources of the same original reports are being blending together (e.g., GTS and IMMT reports from VOS).

The year-month summary products for COADS Release 1 (Woodruff et al. 1987) comprised a selection of statistics (e.g., the mean and median) computed for observed and derived variables using 2°×2° latitude–longitude boxes. Prior to product computation, observations outside 3.5σ (standard deviations) relative to the three climatological periods (1854–1909, 1910–49, 1950–79) were "trimmed" and not used (Slutz et al. 1985, Supp. C). Among early update improvements were higher-resolution summary products (1°×1°) for the period 1960 onward when the sampling density supports this higher resolution.

The Release 1 trimming limits were found to be too restrictive (Wolter 1997) for some extreme climate events (e.g., 1982–83 El Niño). As a partial fix, "enhanced" summaries (in addition to the "standard" ship-based summaries) have been created, with the trimming limits expanded to 4.5σ (but still fixed around the Release 1 climatological medians) to partially account for larger environmental variability. Another characteristic of the enhanced summaries is that they include more platform types (ships plus most other ODAS).

Similarly, the NCDC-QC (Slutz et al. 1985, Supp. J), which forms a less critical component of the overall ICOADS QC (but still has a bearing on the selection of the "best" duplicate), has not been updated significantly since Release 1. For example, part of this procedure uses outdated 5°×5° climatologies based on pre-ICOADS NOAA/NCDC data.

The two ICOADS QC procedures (NCDC-QC and trimming) have employed two interrelated landlocked checks (at 1°×1° and 2°×2° resolution, respectively; http://icoads.noaa.gov/mask.html). As discussed on that webpage, consideration for the use of landmasks for earlier historical ship data is that early ship positions (including for ports) may be greatly reduced in accuracy from values available today. Hence use of a more forgiving (e.g., 2°) landmask may still be desirable for early data.

An important problem, as ICOADS is updated with newly available data, is that the old trimming limits may be missing over regions of new data. For example, the RH values are nearly all missing for the earliest (1854-1909) trimming period, due to a general absence of Release 1 humidity data during this period. This poses a problem for blending new data into ICOADS, because we lack a defined basis for trimming newly available humidity data prior to 1910, and thus for calculating trimmed 2° monthly summaries.

Similarly, the trimming limits are less complete for SLP (and to a lesser extent for other variables) in the earliest period (e.g., in the Pacific basin), than they are for SST. This could tend to cause similar problems, but on a smaller scale, for the addition of newly available data in those regions e.g. prior to 1910 (http://icoads.noaa.gov/dsul.html).

For purposes of calculation of ICOADS monthly summaries, and for sub-selection of observational data for users, flags resulting from the NCDC-QC, trimming, and external QC are utilized (or unused but available to users) as documented here: http://icoads.noaa.gov/e-doc/stat_trim

Developmental work went into a new "Adaptive" QC procedure (based initially on Smith and Reynolds 2003) for ICOADS, and SST observations through 1997 (ICOADS Release 2.0) were flagged using it (http://icoads.noaa.gov/aqc.html). However, the procedure was not operationalized, and the results of its QC improvements have not been independently assessed versus the existing (static, trimming). Conceptual difficulties also arose when it was attempted to extend the procedure to other variables (e.g., wind).

Liz Kent reported in 1996 on data with positions errors in ICOADS, which were not detected by dupelim, at an estimated level of 2-3%. The problem may generally arise because the DM dupelim used at that time did not check for duplicates outside 1°×1° boxes (some blocks of data also appeared to be systematically displaced, such as in January 1984). In conjunction with improved QC and dupelim, track-checking could also be a very beneficial QC addition. Some individual data sources are already track checked.