Procedure for placing hourly super-observations from the Shipboard Automated Meteorological and Oceanographic System (SAMOS) Initiative into ICOADS

Shawn R. Smith and Jocelyn Elya

Center for Ocean-Atmospheric Prediction Studies
The Florida State University
Tallahassee, FL 32306-2741

4 September 2015

This data product and report is funded, in part, by the Climate Observation Division, Climate Program Office, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, under grant NA11OAR4320199 via the Northern Gulf of Mexico Cooperative Institute administered by the Mississippi State University. Additional support to develop the IMMA1 format was provided by CPO under grant number NA11OAR4310169.

Table of Contents

1. Summary

This document outlines the procedure used to construct hourly super-observations (superobs) from the research vessel (RV) observations collected and quality evaluated by the Shipboard Automated Meteorological and Oceanographic System (SAMOS) data assembly center (DAC). The DAC at the Florida State University (FSU) intends that these superobs will be included within the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) as part of a program to archive and distribute high-quality RV marine meteorological data. The superob procedure is based on the typical SAMOS records that include navigational (ship's position, course, speed, and heading), meteorological (winds, air temperature, pressure, moisture, and radiation), and near-surface oceanographic (sea temperature, conductivity, and salinity) parameters recorded at one-minute intervals. All procedures for both data and metadata inclusion into the ICOADS have been discussed and agreed to by members of the ICOADS project. These procedures are similar to Smith et al. (2006) which describe the data translation of World Ocean Circulation Experiment (WOCE) RV Data Quality-Evaluated by FSU/COAPS for ICOADS Release 2.5 (R2.5).

Superobs (averages of up to 11 observations spanning the 10 minutes leading up to the top of an hour) are constructed by the DAC and provided to ICOADS in the International Maritime Meteorological Archive version 1 (IMMA1) format. The superobs are constructed using only "valid" one-minute observations based on the FSU data quality evaluation (DQE) procedures (http://www.coaps.fsu.edu/RVSMDC/html/qc.shtml). Valid observations exclude missing/special values, have valid time and geographic position, and have been determined by DQE to be of good quality. We also include climatological outliers (±4 standard deviations (sdev) from a climatological value) in the superobs as these values often represent extreme, but realistic, observations. The decision to use only valid observations, as opposed to storing some form of quality flag for each value in the IMMA1, is the best compromise under the limitations of the IMMA1 format.

The SAMOS IMMA1 records will include the following six components (Table 1.1): the *Core*, ICOADS attachment (*Icoads*), IMMT-5/FM 13 attachment (*Immt*), ship metadata attachment (*Meta-vos*), Near-surface oceanographic data attachment (*Nocn*), and a supplemental data attachment (*Suppl*). All superobs will be included in the *Suppl* and an objective decision making process will be used to select the "best" of the *Suppl* superobs (when multiple values exist for a given ICOADS element) for inclusion into the *Core* (see section 5). This selection process is necessary since most RVs provide multiple measurements for each parameter (e.g., two air temperatures, three wind measurements) while the IMMA1 *Core* only allows for one value for each parameter. The superobs will be constructed using an objective procedure that takes into account the FSU DQE that is completed on the one-minute observations. Users interested in the details of the FSU quality flags will have to work with the original one-minute SAMOS data files. These files are archived at National Centers for Environmental Information (NCEI)-Silver Spring (Smith et al. 2009; http://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:COAPS-SAMOS) and will be available from the National Center for Atmospheric Research (NCAR; TBD). Mapping to an original one-minute SAMOS file is based on vessel ID, calendar date, and file versions and all required components are included in the *Suppl* for each IMMA1 record.

Table 1.1: Components of an IMMA1 record that will be included for hourly superobs created from SAMOS research vessel data.

IMMA1 Component	Abbrev.	Translation Table	Reference in IMMA1 document	Length (char.)
Core	*Core*	Table 6.1	Table C0	108
ICOADS attm	*Icoads*	Table 7.1	Table C1	65
IMMT-5/FM 13 attm	*Immt*	Table 8.1	Table C5	94
Ship metadata attm	*Meta-vos*	Table 9.1	Table C7	58
Near-surface oceano. attm	*Nocn*	Table 10.1	Table C8	102*
Supplemental data attm	*Suppl*	Table 4.1	Table C99	variable

* Or excluded if it contains no extant data elements (attm data are entirely missing).

2. Outline of Procedure

1. Read SAMOS one-minute research vessel records from daily network Common Data Form (netCDF) files and print valid records observed during superob time frames to an intermediate file.
2. Calculate hourly superobs and print them to an intermediate file.
    a. Calculate averages for up to 11 observations spanning the 10 minutes leading up to the top of an hour.
    b. Determine standard deviation, number of good values, and number of statistical flags for each mean.
    c. Extract appropriate metadata: instrument heights, vessel ID, original units, etc.
3. Read hourly summaries from intermediate file and construct *Suppl* record
4. Determine 'best' value for IMMA1 elements according to objective procedure
5. Write IMMA1 records (*Core+Icoads+Immt+Meta-vos+Nocn+Suppl*) to output files
    a. One file for each ship and month
    b. Create summary file (output from {rwimma})

3. Calculation of Hourly Averages

Through discussions with our partners in the ICOADS project, the decision was made to calculate hourly superobs using data from the 11 possible one-minute observations available for the 10-minute interval leading up to the top of an hour. For example, the 1200 UTC superob is derived from one-minute observations starting at 1150 UTC and ending with the value at 1200 UTC. The need to create a superob that is representative of hourly VOS reports led to the choice of using the ten minute period prior to the hour. Although national practices may vary, manually reported VOS observations are typically observed within the 10 minutes prior to the reporting hour.

The choice of archiving hourly superobs, as opposed to values every three or six hours, was made to provide a good representation of the diurnal variability that is captured by the original one-minute observations. Simply providing superobs every three or six hours would not resolve some of the temporal features captured by the original data (Figure 3.1). Of course, some finer temporal variability is lost, but interested users can always look at the original SAMOS files.

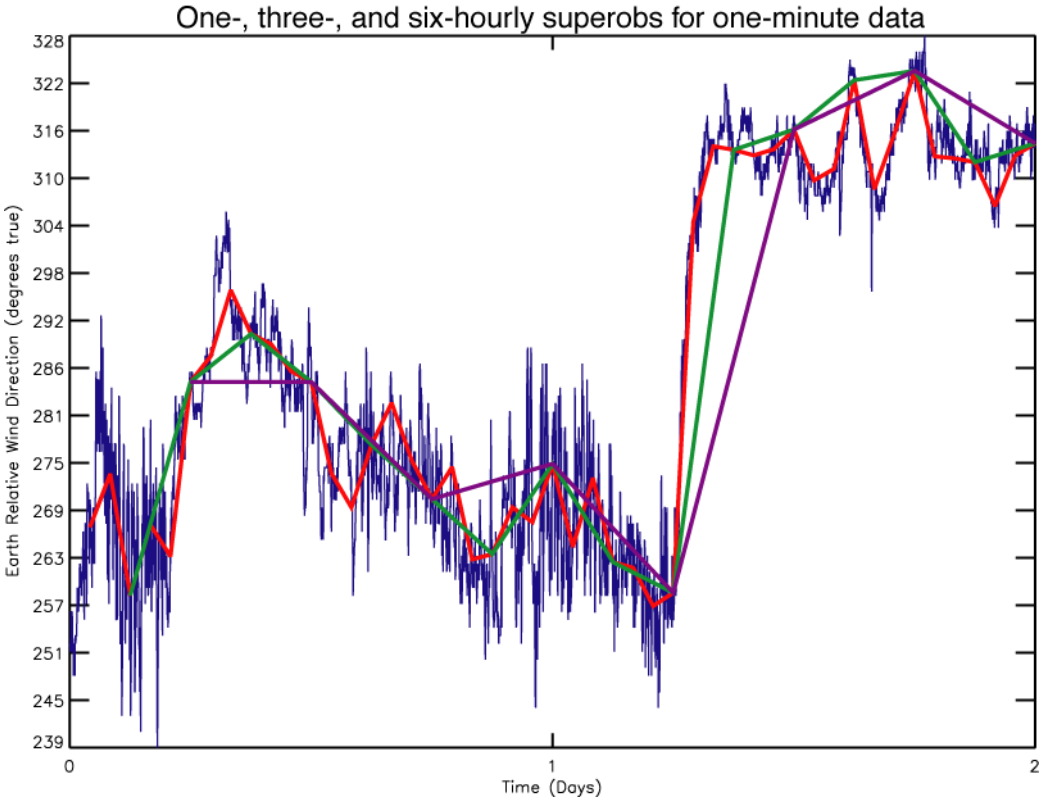

Figure 3.1: Comparison between one-minute sampled true wind direction (blue) for two days versus ten-minute averages calculated every one (red), three (green), and six (purple) hours.

Statistical tests were conducted to confirm that the 10-minute averages are representative of conditions observed by the vessel. These tests included creating averages covering 10, 20, and 30 minutes at the top of each hour from the one-minute observations and comparing the results.

Little variation in the average values existed (Figure 3.2), which supported using the 10-minute averaging period.

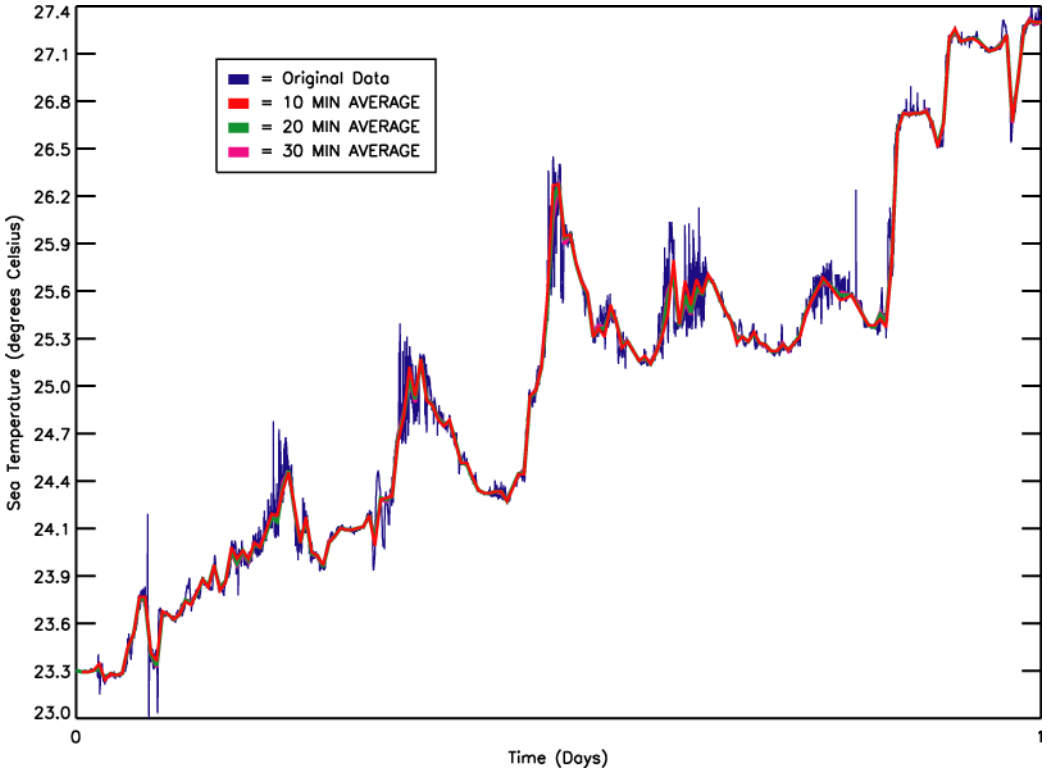

Figure 3.2: Comparison between one-minute sampled sea temperature for one day (blue) versus 10 (red), twenty (green), and thirty (magenta) minute averages constructed from the one-minute observations.

Superobs are calculated for available parameters within each daily SAMOS netCDF data file. Typical parameters will include: latitude; longitude; vessel course, heading, and speed; atmospheric pressure; vessel-relative and true wind direction and speed; sea temperature; salinity; air temperature; and humidity (one or more of wet-bulb, dewpoint, relative humidity, specific humidity). New for SAMOS, with respect to the WOCE data previously included in ICOADS R2.5, are superobs for shortwave, longwave, and photosynthetic radiation and vessel speed relative to the water, when available. Although available in some of the original SAMOS datasets, the current procedure does not create superobs or IMMA1 records for precipitation.

The averaging procedure for one-minute values must take into account missing/special values and the FSU DAC quality control flags (see the appendix for definitions). For each 11-minute averaging period only "valid" records will be included in the average. Valid records must have:

1. a time, latitude, longitude, and at least one other parameter listed in Table 4.2
2. observations with a non-missing (-9999.) and non-special (-8888.) value
3. observations marked with a good data value flag (FSU flags = A, G, I, N, O, or Z)

The number (*nn*) of valid observations used in each the 11-minute average will be provided. Furthermore, the averaging procedure will count the number of 'G' flagged valid data values

(*NG*). The 'G' flagged values represent observations that lie outside a ±4 sdev bounds from the da Silva et al. (1994) monthly climatology. These G-flagged values often represent realistic extreme values and must be included in the averages. This statistical test is only applied to wind speed, air and sea temperature, pressure, and relative humidity (see Table 4.2). The *nn* and *NG* are stored for each average in the IMMA1 *Suppl* attachment and, when necessary, are used to determine the "best" value to be placed into appropriate IMMA1 elements (see Section 5).

As part of the averaging procedure, a sample sdev (Bhattacharyya and Johnson 1977) is calculated for the *nn* values used to construct the mean. The sdev is included to provide a simple measure of the uncertainty in the mean. It is calculated for all means (when $nn > 1$) except the paired values that represent a vector quantity **(**HD and a unit vector, CR and SS, RD and RS, and WD and WS**)**. When available, the sdev is used as part of the decision making process to select values for the *Core* (see Section 5).

4. Content of *Suppl* Attachment for SAMOS Research Vessel Data

The *Suppl* attachment (Table 4.1) includes the hourly superobs for each individual navigational, meteorological, and oceanographic parameter included within the original FSU research vessel files. A *Suppl* will be created for any hourly superob for which a time, latitude, longitude, and one other SAMOS variable (Table 4.2) exists. The *Suppl* will be of variable length, dependent upon the number of parameters included in the original data files. SI units are used for all values in the supplement (consistent with SAMOS units convention).

Table 4.1: Translation specification for elements in SAMOS supplemental (*Suppl*) attachment. Source SAMOS variables or attributes are in **bold**.

No.	Len.	Abbr.	Element description	How filled for SAMOS IMMA1	Comments
1	2	*ATTI*	attm ID	Set = 99	
2	2	*ATTL*	attm length	Set = 0	Record terminated by line feed
3	1	*ATTE*	attm encoding		ATTE is blank for all SAMOS IMMA1 records
4	2	*II*	ID indicator	Set = 1	Replicated from *Core*
5	9	*ID*	identification/call sign	Extracted from **ID** global attribute	Replicated from *Core*
6	1	*TI*	time indicator	Set = 2	Replicated from *Core*
7	10	*ISOT*	ISO time of record		ISO-8601 YYYYMMDDHH
8	3	*Sver*	Originating SAMOS file version	Parsed from original netCDF file name	
9	2	*Sodr*	Originating SAMOS file order number	Parsed from original netCDF file name	
	3	*dsv*	SAMOS-IMMA dataset version number	Set by DAC. dsv = 2 for ICOADS Release 3.0.	
For each additional variable (including latitude and longitude) the following information is provided					
10	2	*VID*	SAMOS variable identifier	See Table 4.2	Alphabetic identifier signifying measured parameter
11	1	*p*	The number of independent groups for *VID* to follow (valid range = 1-9)		e.g., “1” indicates there is only one *VID* sensor, and “2” indicates two different values exist from two independent sensors)
For each *VID*, the following information is provided *p* groups					
12	Var[1]	*data*	The average value, including the sign (when necessary)	See Table 4.2	

13	5	*sdev*[2]	Sample standard deviation for *data*		
14	2	*nn*	Number of valid points in *data* average		
15	3	*ounits*	Original units of *data*	Set based on **original_units** (see Table 4.3)	Conveys the units provided to DAC prior to conversion to SAMOS SI convention.
16	2	*prec*	Precision of *data*	Set based on **data_precision** (see Table 4.7)	Precisions range from 10.0 to 0.0001
17	3	*hhh*	Observation height (tenths of m) above mean sea level for *data*	Extracted from **height** attribute for given variable	Depth in ocean is negative. Blanks denote a missing height or an element where height is not reported
18	2	*NG*	Number of statistical outliers within the *nn* valid points		Only non-zero for elements compared to a climatology (see Table 4.2)
18	1	*type*	Observation type	Extracted from **observation_type** attribute for given variable (see Table 4.4)	
19	2	*TScat*	Sea temperature sensor category	Not populated for ICOADS Release 3.0	Retained in *Suppl* format for future use
20	1	*SLPi*	Sea level pressure indicator	See Table 4.5	
21	1	*RADi*	Radiation direction indicator	Set based on **rad_direction** attribute (see Table 4.6)	

[1] Lengths are variable. See Table 4.2 for data field formats. Except for LA and LO, all the values are stored as integers with 2 decimal place precision (e.g., multiply data by 0.01 to return actual value). For LA and LO, the format includes 4 decimal places (0.0001 multiplier).

[2] All *sdev*s will be stored to two decimal precision with a field length of 5. *sdev* is left missing (blank) when *nn=1*.

Table 4.2: Identifiers to be used by DAC within *Suppl* attachment. Also shown are the original variable identifiers from the SAMOS netCDF files, the generic parameter names, the abbreviation for the IMMA element where the value may be stored, the format that the data value would take when multiplied by 0.01 scale factor (0.0001 for LA/LO), the length of the *data* field in the *Suppl* data group, and whether or not statistical quality control is applied to that element. The '—' denotes parameters that appear only in the *Suppl*.

Suppl VID	SAMOS Variable	Measured Parameter	ICOADS Element	Format After Applying Scale Factor[1]	*Suppl data* String Length	Stat QC applied
LA	**lat**	Latitude (+N)	*LAT*	-xx.xxxx	7	
LO	**lon**	Longitude (+E)	*LON*	xxx.xxxx	7	
SS	**PL_SPD**	Ship speed over ground	*SOG*	xx.xx	4	
CR	**PL_CRS**	Ship course over ground	*COG*	xxx.xx	5	
HD	**PL_HD**	Ship heading	*HDG*	xxx.xx	5	
PW	**PL_SOW**	Ship component speed over water	—	-xx.xx	5	
RD	**PL_WDIR**	Ship-relative wind direction	*RWD*	xxx.xx	5	
RS	**PL_WSPD**	Ship-relative wind speed	*RWS*	xx.xx	4	
WD	**DIR**	True wind direction	*D*	xxx.xx	5	
WS	**SPD**	True wind speed	*W*	xx.xx	4	*
PA	**P**	Atmospheric pressure	*SLP*	xxxx.xx	6	*
SP	—	Atmospheric pressure adjusted to sea level by DAC (when sufficient metadata exist). See Section 5.	*SLP*	xxxx.xx	6	*
TS	**TS**	Sea temperature	*SST, OTV*	-xx.xx	5	*
PS	**SSPS**	Practical Salinity	*OSV*	xx.xx	4	
TA	**T**	Air temperature	*AT*	-xx.xx	5	*
TW	**TW**	Wet-bulb temperature	*WBT*	-xx.xx	5	
TD	**TD**	Dewpoint temperature	*DPT*	-xx.xx	5	
RH	**RH**	Relative humidity	*RH*	xxx.xx	5	*
SW	**RAD_SW**	Shortwave atmospheric radiation	—	xxxx.xx	6	
LW	**RAD_LW**	Longwave atmospheric radiation	—	xxx.xx	5	
RP	**RAD_PAR**	Photosynthetically active radiation	—	xxx.xx	5	

[1]Formats do not imply precision of data values. Refer to *prec* in the *Suppl* attachment and Table 4.7 for data precision information.

Table 4.3: Encoding for original units (**original_units**) provided to SAMOS by participating vessels for measured parameters in Table 4.2.

SAMOS **original_units** String	Units Code (from SAMOS database), *ounits* in *Suppl*
bar	58
calories centimeter-2 minute-1	59
celsius	60
centimeter	61
dd/mm/yy UTC	144
degrees	62
degrees (+E)	63
degrees (+N)	64
degrees (+S)	65
degrees (+W)	66
degrees (+W/-E)	120
degrees (-W/+E)	67
degrees (clockwise from bow)	68
degrees (clockwise from true north)	69
degrees (clockwise towards bow)	70
degrees (clockwise towards true north)	71
fahrenheit	72
feet	73
gram kilogram-1	74
hectopascal	75
hh:mm:ss UTC	143
hhmmss UTC	76
inch	78
inch of mercury	77
kelvin	79
kilogram kilogram-1	80
kilometer hour-1	124
kilowatt meter-2	81
knot	82
langley	83
meter	84

meter second-1	85
microeinstein centimeter-2 second-1	138
microeinstein meter-2 second-1	139
microsiemens centimeter-1	140
microwatt centimeter-2	125
millibar	86
millimeter	87
millimeter hour-1	118
millimeter minute-1	88
millimeter of mercury	89
millimho centimeter-1	142
millisiemens centimeter-1	141
minutes since 1-1-1980 00:00 UTC	90
oktas	91
pascal	92
percent	93
PSU	131
siemens meter-1	130
tenths	94
watts meter-2	95
WMO code table	96
YYYYJJJ UTC	127
YYYYJJJhhmmss UTC	128
YYYYMMDD UTC	97
YYYYMMDDhhmmss UTC	126

Table 4.4: Encoding for **observation_type** provided to SAMOS by participating vessels for measured parameters in Table 4.2.

SAMOS **observation_type** String	*type* in *Suppl*	IMMA1 *WBTI* code for **TW observation_type**
unknown	0	—
measured	1	0
calculated	2	1

Table 4.5: Encoding for sea level pressure indicator *SLPi.*

SLPi in *Suppl*	Description
0	SAMOS **mslp_indicator** = “unknown” or missing
1	SAMOS **mslp_indicator** = “adjusted to sea level”
2	SAMOS **mslp_indicator** = “at sensor height”
3	Pressure adjusted to sea level by DAC

Table 4.6: Mapping radiation direction (**rad_direction**) provided to SAMOS by participating vessels for **RAD_SW**, **RAD_LW**, and **RAD_PAR** variables to *RADi* in *Suppl*.

SAMOS **rad_direction**	*RAD*i in *Suppl*
“downwelling”	1
“upwelling”	2
“unknown” or missing	0

Table 4.7: Mapping data precision (**data_precision**) provided to SAMOS by participating vessels to *prec* in *Suppl* and IMMA1 wind direction indicator *DI*.

SAMOS **data_precision** String	*prec* in *Suppl*	IMMA1 *DI* code for **DIR data_precision**
10	1	0
1.0	2	5
1.	2	5
1	3	5
0.5	5	6
0.3	7	6
0.2	8	6
0.1	9	6
0.01	10	6
.01	10	6
0.002	13	6
0.001	14	6
0.0001	16	6
0.000051	17	6
0.00001	18	6
.00001	18	6
0.000001	20	6
0.0000001	21	6

Within the *Suppl*, the format for a single data group would be: *VIDpdatasdevnnounitsprechhhNGtypeTScatSLPiRADi*. When multiple sensors exist, the *VIDp* will not be duplicated, just the *datasdevnnounitsprechhhNGtypeTScatSLPiRADi* part of the group

The following example

SS1 552 11 82 2 01 00TA2 1977 511 60 140 01 00 1980 011 60 202 40 00

includes a data group for a single measured (*type* = 1) value of ship speed over the ground (5.52 m s^{-1}) derived from 11 one-minute observations that were converted from original units of knots (*ounits* code = 82) and a data precision of 1.0 (*prec* = 2). For SS, *NG*, *SLPi*, and *RADi* are all set to 0 as they do not apply to SS. The second data group includes two air temperatures, 19.77˚C

and 19.80˚C, with *sdev* = 0.05˚C and 0.00˚C, respectively. Both air temperature averages included 11 values (*nn* = 11), had original units of celsius (*ounits* code = 60), and unknown data precision (*prec* set to blank). The first value was measured *(type* = 1) at a height of 14.0 m and had no statistical outliers in the average (*NG*=0), while the second value had an unknown observation type (*type* =0), was reported at 20.2 m, and had four statistical outliers in the average (*NG* = 4). *SLPi* and *RADi* are set to 0 as they do not apply to TA.

Following the IMMA uniform conventions for representation of numerical and missing data, other rules that apply when writing the *Suppl* include:

1. Each IMMA record is terminated with a line feed. This allows for variable line lengths between subsequent records. Variable lengths are used for *data* values (see Table 4.2).
2. Fill zeros will be used to the right of the decimal for each data value (this simplifies the write and read code). Refer to *prec* in the *Suppl* attachment and Table 4.7 for data precision information.
3. Unused spaces in all fields to the left of a decimal will be filled with blanks.
4. Data values in *Suppl* are scaled in order to be represented without decimal points (see Table 4.2), which results in the following consequences:
    a. A data value of 0.00 is written to a *data* element of length 4 as “ 0”.
    b. Data values less than 1 and greater than -1 lose the leading zero. For example, a data value of -0.12 is written to a *data* element of length 5 as “ -12”.
5. In the FSU *Suppl*, longitudes are always reported as positive east values (0.0000 to 359.9999˚E)
6. A *Suppl* record will only exist if a valid time, latitude, longitude, and at least one other parameter exist for the YYYYMMDDHH associated with the record.
7. A SAMOS IMMA1 record has a maximum length of 2048 characters.

The information in the *Suppl* also provides the ability to trace an individual IMMA1 record back to the source SAMOS netCDF file. The file naming convention for a SAMOS file is:

CALLSIGN_YYYYMMDDvVVVOO.nc

Where individual parts of the file name are outlined as follows:

	Description	*Suppl* Element
CALLSIGN	Vessel Call Sign (4-7 Alphanumeric characters)	*ID*
YYYY	Four-digit year of the first record in the file	*ISOT*
MM	Two-digit month of the first record in the file	*ISOT*
DD	Two-digit date of the first record in the file	*ISOT*
VVV	SAMOS version number (typically 200 or 300)	*Sver*
OO	Two-digit number describing the order in which files are received.	*Sodr*

5. Selection of "Best" Values for Inclusion in IMMA1 Elements

The selection process will follow a set of rules that will apply to each SAMOS variable to be assigned to the IMMA1 record. This process will apply to fill the *Core* elements *LAT*, *LON, D, W, SLP, AT, WBT, DPT,* and *SST*; *Immt* elements *HDG, COG, SOG, RWD, RWS*, and *RH*; and *Nocn* elements *OTV* and *OSV*. Since cloud, weather, wave, and swell data are not available for any research vessels participating in SAMOS these fields will not be populated in the *Core*. The rules include:

1. A superob to be included in IMMA1 elements must be derived from a minimum of 5 one-minute observations. A superob created from less than 5 values may not be representative of the physical atmospheric conditions over the 11 minutes averaged at the top of the hour.
2. When two or more superobs are available that have passed rule 1 for one IMMA1 field, the superob with the minimum sample sdev will be stored in the *Core*, *Immt*, or *Nocn* element.
3. If two or more superobs exist for one IMMA1 field with equal minimum sample sdev values, the first superob will be stored in the *Core*, *Immt*, or *Nocn* element.
4. If none of the superobs available for a *Core, Immt,* or *Nocn* element pass rules 1 - 3, the element will not be filled (set to blanks).

Additional rules apply for filling *SLP* in the *Core*. A pressure **P** superob is assigned to *SLP* if it passes rules 1 - 3 and is known to be "adjusted to sea level" (see Table 4.5). If **P** is valid and observed at a known "sensor height" (see Table 4.5) and a concurrent valid air temperature **T** observation exists, the DAC adjusts one-minute **P** values to sea level using the equation below and stores the adjusted values in a new parameter **SP. SP** superobs are then calculated and stored in the *Suppl*. If an **SP** superob passes rules 1 - 3, it is assigned to *SLP*.

$$SP = P_z e^{\left(\frac{gz}{R_a T_{air}}\right)}$$

SP is the pressure at sea level (hPa or mb), P_z is the pressure **P** from the SAMOS file (hPa or mb) at a known instrument height (z in meters extracted from **height** attribute), g = 9.81 m s^{-2}, R_a = 287.05 J kg^{-1} K^{-1}, e = 2.7182818, and T_{air} is the air temperature in K (T_{air} = **T** + 273.15) simultaneous in time with the P_z.

6. Filling the IMMA1 *Core* Record

The values in the IMMA1 *Core* are selected from the *Suppl* record using an automated and objective process (Section 5). The IMMA1 fields typically filled in the *Core* are highlighted in Table 6.1 along with the source of the information from the original netCDF files. Each *Core* value must conform to a set of rules (below) before they appear in the *Core*. A time, latitude, and longitude will exist in the *Core* for each record that has data in the *Suppl*.

Table 6.1: Translation specifications for elements in IMMA1 Core (*Core*). Source SAMOS variables or attributes are in **bold**.

No.	Abbr.	Element description	How filled for SAMOS IMMA1	Comments
1	*YR*	year UTC	Extracted from **time** using invtime[2]	
2	*MO*	month UTC	Extracted from **time** using invtime[2]	
3	*DY*	day UTC	Extracted from **time** using invtime[2]	
4	*HR*	hour UTC	Extracted from **time** using invtime[2]	
5	*LAT*	latitude	Extracted from **lat**	
6	*LON*	longitude[1]	Extracted from **lon**	Retain ICOADS convention of 0.00-359.99
7	*IM*	IMMA version	Set = 1	
8	*ATTC*	attm count	Set = 4 or 5	
9	*TI*	time indicator	Set = 2	2 = Hour plus minutes (consistent with WOCE IMMA in R2.5)
10	*LI*	latitude/long. indic.	Set = 5	5 = High resolution data
11	*DS*	ship course	—	
12	*VS*	ship speed	—	
13	*NID*	national source indic.	—	
14	*II*	*ID* indicator	Set = 1	1 = ship call sign
15	*ID*	identification/call sign	Extracted from **ID** global attribute	
16	*C1*	country code	—	
17	*DI*	wind direction indic.	Set based on **data_precision** for **DIRx**[1] (see Table 4.7)	
18	*D*	wind direction (true)	Best value from **DIRx**[1]	Range of values is 0-360 degrees. When wind is calm, D=0 and W=0. We do not use D=361 convention for calm winds.

19	*WI*	wind speed indicator	Set based on **original_units** for **SPDx**[1] (see Table 6.2)	All SAMOS winds are anemometer based and high-resolution, so decided setting WI = 1 or 4 based on **original_units** was more important that noting high-resolution (WI=8).
20	*W*	wind speed	Best value from **SPDx**[1]	
21	*VI*	*VV* indic.	—	
22	*VV*	visibility	—	
23	*WW*	present weather	—	
24	*W1*	past weather	—	
25	*SLP*	sea level pressure	Best value from **Px**[1] (if adjusted to sea level) or **SPx**[1] (see Section 5)	SLP only populated when *SLPi* = 1 or 3 in *Suppl* (see Table 4.5 and Section 5)
26	*A*	characteristic of *PPP*	—	
27	*PPP*	amt. pressure tend.	—	
28	*IT*	indic. for temperatures	Set = 9	Since AT, WBT, DPT, and SST precisions and units vary greatly in SAMOS, IT=9 (other) is the best choice. User is referred to **original_units** attribute in *Suppl* and SAMOS data may be assumed to primarily be high-resolution (tenths to thousandths)
29	*AT*	air temperature	Best value from **Tx**[1]	
30	*WBTI*	*WBT* indic.	Set based on **observation_type** for **TWx**[1] (see Table 4.4)	
31	*WBT*	wet-bulb temperature	Best value from **TWx**[1]	
32	*DPTI*	*DPT* indic.	—	
33	*DPT*	dew-point temperature	Best value from **TDx**[1]	
34	*SI*	*SST* meas. method	Populated based on database entries for **TSx**[1] (see Table 6.3)	
35	*SST*	sea surface temp.	Best value from **TSx**[1]	
36	*N*	total cloud amount	—	
37	*NH*	lower cloud amount	—	
38	*CL*	low cloud type	—	
39	*HI*	*H* indic.	—	
40	*H*	cloud height	—	

41	*CM*	middle cloud type	—	
42	*CH*	high cloud type	—	
43	*WD*	wave direction	—	
44	*WP*	wave period	—	
45	*WH*	wave height	—	
46	*SD*	swell direction	—	
47	*SP*	swell period	—	
48	*SH*	swell height	—	

[1]Where x=blank, 2, …, 8, 9; for research vessels providing values from redundant sensors.
[2]See http://samos.coaps.fsu.edu/html/tools_invtime.php.

Table 6.2: Mappings from SAMOS true wind speed **original_units** to *ounits* in *Suppl* and IMMA1 wind speed indicator *WI* codes.

SAMOS **SPD original_units** String	Units Code (from SAMOS database), *ounits* in *Suppl*	IMMA1 *WI* code
meter second-1	85	1
knot	82	4

Table 6.3: *SI* code for a ship and time frame based on **TS** metadata from the SAMOS database as of July 13, 2015.

Name	Call Sign	Start Date	End Date	IMMA1 *SI* code
ROGER REVELLE	KAOU	6/6/2011	12/31/2014	12
ATLANTIS	KAQP	6/1/2005	12/31/2014	12
KNORR	KCEJ	5/9/2005	12/31/2014	3
DELAWARE II	KNBD	2/16/2009	12/31/2014	7
T.G. THOMPSON	KTDQ	1/1/2012	10/20/2013	12
T.G. THOMPSON	KTDQ	10/21/2013	12/31/2014	9
HEALY	NEPP	5/20/2007	10/12/2010	9
HEALY	NEPP	5/25/2011	12/31/2014	12
POLAR SEA	NRUO	9/27/2009	12/31/2014	12
SOUTHERN SURVEYOR	VLHJ	4/16/2008	12/31/2014	9
AURORA AUSTRALIS	VNAA	1/27/2008	12/31/2014	9
NATHANIEL B. PALMER	WBP3210	11/7/2006	12/5/2011	9
NATHANIEL B. PALMER	WBP3210	12/6/2011	12/31/2014	12
LAURENCE M. GOULD	WCX7445	4/11/2007	6/1/2014	6
LAURENCE M. GOULD	WCX7445	6/6/2014	12/31/2014	12
KILO MOANA	WDA7827	7/1/2009	12/31/2014	12
ATLANTIC EXPLORER	WDC9417	3/8/2010	3/31/2015	12
PELICAN	WDD6114	12/11/2014	12/31/2014	12
MELVILLE	WECB	6/10/2011	12/31/2014	12
NEW HORIZON	WKWB	5/18/2012	12/31/2014	12
ROBERT GORDON SPROUL	WSQ2674	4/15/2012	12/31/2014	3
HENRY B. BIGELOW	WTDF	4/4/2007	2/25/2009	12
HENRY B. BIGELOW	WTDF	2/26/2009	12/31/2014	9
OKEANOS EXPLORER	WTDH	6/17/2009	12/31/2014	12
DAVID STARR JORDAN	WTDK	3/19/2008	12/31/2014	12
PISCES	WTDL	7/26/2009	12/31/2014	12
MILLER FREEMAN	WTDM	1/17/2007	12/31/2014	12
OREGON II	WTDO	6/1/2008	12/31/2014	9
THOMAS JEFFERSON	WTEA	7/29/2012	12/31/2014	9
FAIRWEATHER	WTEB	1/21/2008	12/31/2014	12
RON BROWN	WTEC	3/21/2007	2/9/2015	12
BELL M. SHIMADA	WTED	2/22/2012	9/11/2012	9
BELL M. SHIMADA	WTED	9/12/2012	12/31/2014	12

OSCAR ELTON SETTE	WTEE	1/9/2009	12/31/2014	9
RAINIER	WTEF	1/20/2008	12/31/2014	9
MCARTHUR II	WTEJ	4/20/2010	12/31/2014	9
GORDON GUNTER	WTEO	6/6/2007	12/31/2014	3
OSCAR DYSON	WTEP	1/25/2007	12/31/2014	12
NANCY FOSTER	WTER	10/24/2004	12/31/2014	9
KA'IMIMOANA	WTEU	3/21/2007	12/31/2014	12
HI'IALAKAI	WTEY	1/21/2007	3/16/2015	12
OCEANUS	WXAQ	3/25/2008	12/31/2014	9
FALKOR	ZCYL5	3/1/2013	12/31/2014	9
TANGAROA	ZMFR	4/27/2011	12/31/2014	9

7. Filling the *Icoads* Attachment

FSU will populate only three elements in the ICOADS attachment (Table 7.1). The remainder of this attachment will be filled by the ICOADS program when these RV records are integrated into the ICOADS. The DCK and SID were assigned by the ICOADS group.

Table 7.1: Translation specifications for elements in ICOADS (*Icoads*) attachment.

No.	Abbr.	Element description	How filled for SAMOS IMMA1	Comments
1	*ATTI*	attm ID	Set = 1	
2	*ATTL*	attm length	Set = 65	
3	*BSI*	box system indicator	—	
4	*B10*	10° box number	—	
5	*B1*	1° box number	—	
6	*DCK*	deck	Set = 740	740 = Research Vessel (R/V) Data Quality-Evaluated by FSU/COAPS
7	*SID*[1]	source ID	Set = 131	131 = Research Vessel (R/V) Data Quality-Evaluated by FSU/COAPS: SAMOS
8	*PT*	platform type	Set = 5	5 is code for ships
9	DUPS	dup status	—	
10	*DUPC*	dup check	—	
11	*TC*	track check	—	
12	*PB*	pressure bias	—	
13	*WX*	wave period indicator	—	
14	*SX*	swell period indicator	—	
15	*C2*	2nd country code	—	
16-27	*SQZ-DQA*	adaptive QC flags	—	
28	*ND*	night/day flag	—	
29-34	*SF-RF*	trimming flags	—	
35-48	*ZNC-TNC*	NCDC-QC flags	—	
49	*QCE*	external (e.g. MEDS)	—	
50	*LZ*	landlocked flag	—	
51	QCZ	source exclusion flags	—	

[1]For the earlier WOCE data included in R2.5 also under DCK=740, SID=130 was set: Research Vessel (R/V) Data Quality-Evaluated by FSU/COAPS: WOCE ver.3.0

8. Filling the *Immt* Attachment

FSU will take advantage of a number of fields in the IMMT-5/FM 13 attachment to provide user access to additional parameters commonly measured by SAMOS on research vessels (Table 8.1). When multiple values for a given element exist in a SAMOS record, the rules in section 5 will apply to select the value assigned to the *Immt* element.

Table 8.1: Translation specifications for elements in IMMT-5/FM 13 (*Immt*) attachment. SAMOS uses this table to report elements available from research vessels that cannot be stored elsewhere in IMMA1 (e.g., relative humidity). Source SAMOS variables or attributes are in **bold**.

No.	Abbr.	Element description	How filled for SAMOS IMMA1	Comments
1	*ATTI*	attm ID	Set = 5	
2	*ATTL*	attm length	Set = 94	
3	*OS*	observation source	—	
4	*OP*	observation platform	—	
5	*FM*	FM code version	—	
6	*IMMV*	IMMT version	—	
7	*IX*	station/weather indic.	—	
8	*W2*	2nd past weather	—	
9	*WMI*	indic. for wave meas.	—	
10	*SD2*	dir. of second. swell	—	
11	*SP2*	per. of second. swell	—	
12	*SH2*	ht. of second. swell	—	
13	*IS*	ice accretion on ship	—	
14	*ES*	thickness of I_s	—	
15	*RS*	rate of I_s	—	
16	*IC1*	concentration of sea ice	—	
17	*IC2*	stage of development	—	
18	*IC3*	ice of land origin	—	
19	*IC4*	true bearing ice edge	—	
20	*IC5*	ice situation/trend	—	
21	*IR*	indic. for precip. data	—	
22	*RRR*	amount of precip.	—	
23	*TR*	duration of per. *RRR*	—	

24	*NU*	national use	—	
25	*QCI*	quality control indic.	—	
26-45	*QI1-20*	QC indic. for fields	—	
46	*QI21*	MQCS version	—	
47	*HDG*	ship's heading	Best value from **PL_HDx**[1]	
48	*COG*	course over ground	Best value from **PL_CRSx**[1]	
49	*SOG*	speed over ground	Best value from **PL_SPDx**[1]	Units converted from m/s to kts using *SOG* = **PL_SPD**(1.943844492 4406047516)
50	*SLL*	max.ht.>Sum. load ln.	—	
51	*SLHH*	dep. load ln.: sea lev.	—	
52	*RWD*	relative wind direction	Best value from **PL_WDIRx**[1].	Range of values is 0-360 degrees. When relative wind is calm, RWD=0 and RWS=0. We do not use RWD=361 convention for calm winds.
53	*RWS*	relative wind speed	Best value from **PL_WSPDx**[1]	
54-61	*QI22-29*	QC indic. for fields	—	
62	*RH*	relative humidity	Best value from **RHx**[1]	
63	*RHI*	relative humidity indic.	Set based on **data_precision** and **observation_type** attributes for **RHx**[1] (see Table 8.2)	
64	*AWSI*	AWS indicator	Set = 1	All SAMOS are AWS.
65	*IMONO*	IMO number	Extracted from SAMOS ship database (see Table 8.3)	

[1]Where x=blank, 2, …, 8, 9; for research vessels providing values from redundant sensors.

Table 8.2: Mapping data_precision (**data_precision**) and observation type (**observation_type**) provided to SAMOS by participating vessels for the **RH** variable to *prec* and *type* in *Suppl* and IMMA1 relative humidity indicator *RHI*.

SAMOS **RH data_precision** String	*prec* in *Suppl*	SAMOS **RH observation_type** String	*type* in *Suppl*	IMMA1 *RHI* code
0.1	9	unknown	0	0
0.1	9	measured	1	0
0.1	9	calculated	2	3
1.0	2	unknown	0	1
1.0	2	measured	1	1
1.0	2	calculated	2	4
1	3	unknown	0	1
1	3	measured	1	1
1	3	calculated	2	4

Table 8.3: Ship metadata extracted from the SAMOS database on April 15, 2015.

Name	Call Sign	IMO Number	Length in meters
KILO MOANA	WDA7827	9229037	57
T.G. THOMPSON	KTDQ	8814419	84
NEW HORIZON	WKWB	7723821	52
ENDEAVOR	WCE5063	7604300	56
KNORR	KCEJ	7738618	85
ATLANTIS	KAQP	9105798	83.2
OCEANUS	WXAQ	7603617	54
NATHANIEL B. PALMER	WBP3210	9007257	91.4
LAURENCE M. GOULD	WCX7445	9137337	70.2
POLAR SEA	NRUO	7391252	121.9
HEALY	NEPP	9083380	128
NANCY FOSTER	WTER	8993227	56.7
OSCAR DYSON	WTEP	9270335	63.8
KA'IMIMOANA	WTEU	8835231	68.2
HENRY B. BIGELOW	WTDF	9349057	63.8
RON BROWN	WTEC	9105786	56.7
HI'IALAKAI	WTEY	8835619	68.3
GORDON GUNTER	WTEO	8835255	68.3
RAINIER	WTEF	6711003	62

FAIRWEATHER	WTEB	6710920	70.4
OREGON II	WTDO	6728068	51.8
OSCAR ELTON SETTE	WTEE	8835097	70
OKEANOS EXPLORER	WTDH	8835114	68.28
PISCES	WTDL	9349071	63.6
MCARTHUR II	WTEJ	8833867	68
ATLANTIC EXPLORER	WDC9417	8120014	52.12
AURORA AUSTRALIS	VNAA	8717283	94.9
ROGER REVELLE	KAOU	9075228	84.4
MELVILLE	WECB	7738591	85
TANGAROA	ZMFR	9011571	70
BELL M. SHIMADA	WTED	9349069	63.8
ROBERT GORDON SPROUL	WSQ2674	—	38
THOMAS JEFFERSON	WTEA	8892033	63.4
FALKOR	ZCYL5	7928677	82.9
FERDINAND HASSSLER	WTEK	9478559	37.7
DAVID STARR JORDAN	WTDK	7333195	52.1
DELAWARE II	KNBD	7629946	47.2
MILLER FREEMAN	WTDM	6621636	65.5
SOUTHERN SURVEYOR	VLHJ	7113002	66.1

9. Filling the *Meta-vos* Attachment

FSU will take advantage of a number of fields in the ship metadata attachment to provide user access to additional elements commonly provided by SAMOS research vessels (Table 9.1). Although the *Meta-vos* was originally designed to include a subset of the metadata included in the WMO Publication 47 for ships participating in the Voluntary Observing Ship program, during the development of IMMA1 the ICOADS team added a metadata source (MDS) element to support use of the attachment by other data providers. This approach will simplify metadata access for users as opposed to only providing these metadata in the more complicated *Suppl* attachment.

Table 9.1: Translation specifications for elements in Ship Metadata (Meta-vos) attachment. Source SAMOS variables or attributes are in **bold**.

No.	Abbr.	Element description	How filled for SAMOS IMMA1	Comments
1	*ATTI*	attm ID	Set = 7	
2	*ATTL*	attm length	Set = 58	
3	*MDS*	metadata source	Set = 1	0=WMO Pub. 47; 1=COAPS
4	*C1M*	recruiting country	—	
5	*OPM*	type of ship (programme)	—	
6	*KOV*	kind of vessel	Set = RV	Code for research vessel (Source: WMO47IMMA_1973_2006-R2.4.pdf)
7	*COR*	country of registry	—	
8	*TOB*	type of barometer	—	
9	*TOT*	type of thermometer	—	
10	*EOT*	exposure of thermometer	—	
11	*LOT*	screen location	—	
12	*TOH*	type of hygrometer	—	
13	*EOH*	exposure of hygrometer	—	
14	*SIM*	*SST* meas. Method	—	
15	*LOV*	length of vessel	Extracted from SAMOS ship database (see Table 8.3)	
16	*DOS*	depth of *SST* meas.	Extracted from **height** attribute for the **TS** value assigned to SST in the *Core*.	DOS = -1(**height**) to account for depth being negative in SAMOS format and positive in IMMA1.

17	*HOP*	height of visual observation platform	—	
18	*HOT*	height of *AT* sensor	Extracted from **height** attribute for the **T** value assigned to AT in the *Core*.	HOT is rounded to whole m, but stored at full resolution in *Suppl*.
19	*HOB*	height of barometer	Extracted from **height** attribute for the pressure value assigned to SLP in the *Core*.	HOB is rounded to whole m, but stored at full resolution in *Suppl*.
20	*HOA*	height of anemometer	Extracted from **height** attribute for the **DIR** value assigned to D in the *Core*.	HOA is rounded to whole m, but stored at full resolution in *Suppl*.
21	*SMF*	source metadata file	—	
22	*SME*	source meta. element	—	
23	*SMV*	source format version	—	

10. Filling the *Nocn* Attachment

Almost every vessel reporting to SAMOS is equipped with a thermosalinograph plumbed into their scientific seawater system. These sensors report conductivity and salinity to very high precision and the salinity value is stored in the *Nocn* (Table 10.1) along with a high-precision sea temperature (matching the value in the *Core*). When multiple values for a given element exist in a SAMOS record, the rules in section 5 will apply to select the value assigned to the *Nocn* element.

Table 10.1: Translation specifications for elements in Near-surface Oceanographic Data (*Nocn*) attachment. Source SAMOS variables or attributes are in **bold**.

No.	Abbr.	Element description	How filled for SAMOS IMMA1	Comments
1	*ATTI*	attm ID	Set = 8	
2	*ATTL*	attm length	Set = 102 [2U in base36]	
3	*OTV*	temperature value	Best value from **TSx**[1]	Same value as SST in *Core*, but higher precision when available.
4	*OTZ*	temperature depth	Extracted from **height** attribute for the **TS** value assigned to SST in the *Core*. To match IMMA1 convention that *OTZ* is positive, the **height** is multiplied by -1.	
5	*OSV*	salinity value	Best value from **SSPSx**[1]	
6	*OSZ*	salinity depth	Extracted from **height** attribute for the **SSPS** assigned to *OSV*. To match IMMA1 convention that *OSZ* is positive, the **height** is multiplied by -1.	
7	*OOV*	dissolved oxygen	—	
8	*OOZ*	dissolved oxygen depth	—	
9	*OPV*	phosphate value	—	
10	*OPZ*	phosphate depth	—	
11	*OSIV*	silicate value	—	
12	*OSIZ*	silicate depth	—	
13	*ONV*	nitrate value	—	
14	*ONZ*	nitrate depth	—	
15	*OPHV*	pH value	—	
16	*OPHZ*	pH depth	—	
17	*OCV*	total chlorophyll value	—	

18	*OCZ*	total chlorophyll depth	—	
19	*OAV*	alkalinity value	—	
20	*OAZ*	alkalinity depth	—	
21	*OPCV*	partial pressure of carbon dioxide value	—	
22	*OPCZ*	partial pressure of carbon dioxide depth	—	
23	*ODV*	dissolved inorganic carbon value	—	
24	*ODZ*	dissolved inorganic carbon depth	—	
25	*PUID*	provider's unique record Identification	—	

[1]Where x=blank, 2, …, 8, 9; for research vessels providing values from redundant sensors.

References

Bhattacharyya, G. K., and R. A. Johnson, 1977: *Statistical Concepts and Methods*. Wiley and Sons, 639pp.

da Silva, A. M., C. C. Young, and S. Levitus, 1994: Atlas of Surface Marine Data, Volumes 1: Algorithms and Procedures. NOAA Atlas Series, U.S. Dept. of Commerce, NOAA, NESDIS, Data, and Information Service : For sale by the U.S. G.P.O., Supt. of Docs.

Smith, S. R., B. Olafson, and J. Lamm, 2006: Procedure for placing hourly super-observations from FSU Data Assembly Center's (DAC) Research Vessel data files into ICOADS. Available from http://icoads.noaa.gov/e-doc/other/transpec/coapsrv/RVtoIMMAprocedure.pdf.

Smith, S. R., J. J. Rolph, K. Briggs, M. A. Bourassa, 2009: Quality-Controlled Underway Oceanographic and Meteorological Data from the Center for Ocean-Atmospheric Predictions Center (COAPS) - Shipboard Automated Meteorological and Oceanographic System (SAMOS). National Oceanographic Data Center, NOAA. Dataset. doi:10.7289/V5QJ7F8R

Appendix

A description of the FSU/SAMOS quality evaluation process can be found at:
http://www.coaps.fsu.edu/woce/docs/qchbook/qchbook.htm

The quality control flags are single alphabetic characters for each data value. Only those variables with a qcindex in the original one-minute data files have flag values (the qcindex is an integer pointer to the flag for a selected variable), i.e. not all meteorological variables are quality controlled. More details on the qcindex are available at:
http://samos.coaps.fsu.edu/html/docs/samos_netcdf_manual.pdf

The flag definitions are:

A: Original data had unknown units. The units shown were determined using a climatology or some other method.

B: Original data were out of physically realistic range bounds outlined.

C: Time data are not sequential or date/time not valid.

D: Data failed T>=Tw>=Td 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.

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 degrees or the wind speed difference is >2.5m/s.

F: Platform velocity unrealistic. Determined by analyzing latitude and longitude positions as well as reported platform speed data.

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.

H: Discontinuity found in data

I: Interesting feature found in data. More specific information on the feature is contained in the data reports. Examples include: hurricanes passing stations, sharp sea water temperature gradients, strong convective events, etc.

J: Data are of poor quality by visual inspection, DO NOT USE.

K: Data suspect/use with caution - this flag applies when the data look to have obvious errors, but no specific reason for the error can be determined.

L: Oceanographic platform passes over land or fixed platform moves dramatically.

M: Known instrument malfunction.

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.

O: Original units differ from those listed in the convers_units variable attribute. See quality control report for details.

P: Position of platform or its movement are uncertain. Data should be used with caution.

Q: Questionable - data arrived at DAC already flagged as questionable/uncertain.

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.

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.

T: Time duplicate

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.

V: Data spike as determined by SASSI.

X: Step/discontinuity in data as determined by SASSI

Y: Suspect values between X-flagged data (from SASSI)

Z: Data passed evaluation.