**Procedure for placing hourly super-observations from
FSU Data Assembly Center's (DAC) Research Vessel data files into ICOADS**

Shawn R. Smith, Barrett Olafson, and James Lamm

Center for Ocean-Atmospheric Prediction Studies
The Florida State University
Tallahassee, FL 32306-2840

RVSMDC report 2006-01

9 April 2006

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 RV data assembly center (DAC). The RV DAC at the Florida State University 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 FSU RV 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.

Superobs (averages spanning the 10 minutes leading up to the top of an hour) will be constructed by the RV DAC and provided to ICOADS in the International Marine Meteorological Archive (IMMA) format. The superobs for will be 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 are non-missing, have valid time and geographic position, and have been determined by DQE to be of good quality or to be +/- 4 standard deviations from a climatological value. We include climatological outliers 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 IMMA, is difficult, but is the best compromise under the limitations of the IMMA format.

The FSU RV IMMA records will include the core (C0), ICOADS attachment (C1), ship metadata attachment (C4), and a supplemental data attachment (C6). All superobs will be included in the C6 and an objective decision making process will be used to select the "best" of the C6 superobs for inclusion into the C0 (core). This selection process is necessary since most RVs provide multiple measurements for each parameter (e.g., two air temperatures, three wind measurements) while the IMMA core only allows for one value for each parameter. The superobs will be constructed using a objective procedure that takes into account the FSU DQE that is completed on the one-minute observations. Both the codes for constructing the superobs (see appendix A) and the decision making for selecting values for the core (see section 5.1) will be provided to ICOADS. Users interested in the details of the FSU quality flags will have to work with the original one-minute observation files. These files and the corresponding data quality reports will be provided to the ICOADS project for archival and future distribution (note – mapping to original one-minute file is straightforward as the netCDF file names will contain the vessel ID and the date of observations – both of which are in the IMMA core).

2. Outline of procedure

1. Calculate hourly averages from DAC one-minute research vessel netCDF files.
   a. Determine standard deviations, number of good values, number of statistical flags for each mean
   b. Extract appropriate metadata: instrument heights, vessel ID, convers_units
2. Read hourly summaries and construct C6 record
3. Determine 'best' value for IMMA core values according to objective procedure
4. Write IMMA records (core+C1+C4+C6) to output files
   a. Create one file per cruise
   b. Create summary file for each cruise (procedure developed by ICOADS group)

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 during the 10 minutes 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 1). Of course, some finer temporal variability is lost, but interested users can always look at the original netCDF files.
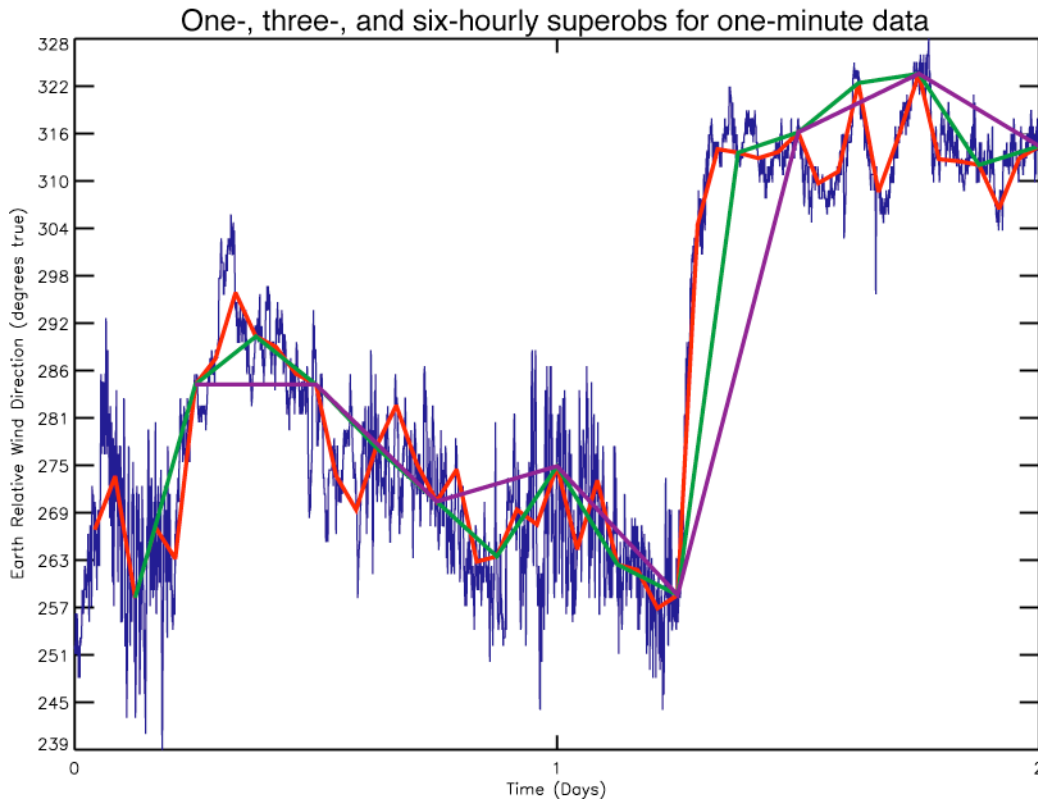

Figure 1. Comparison between one-minute sampled true wind direction for two days (blue) versus ten-minute averages calculated every one (red), three (green), and six (purple) hours.

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 2), which supported using the 10-minute averaging period.
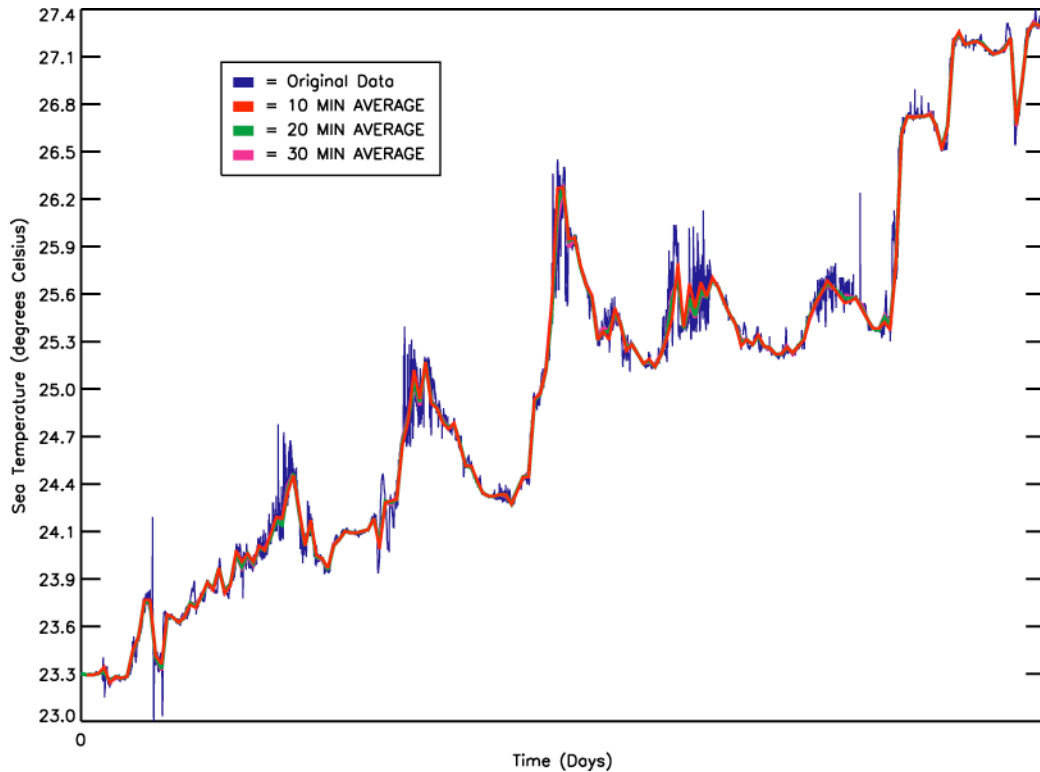
Figure 2. Comparison between one-minute sampled sea temperature for one day (blue) versus 10 (red), twenty (green), and thirty (magenta) minute averages constructed from the one-minute observations.

Superobs are calculated for all available parameters within each one-minute data set. Data sets are processed by cruise (a port-to-port leg of a research vessel cruise) since measured parameters do not tend to change within a cruise leg. Typical parameters will include: latitude; longitude; vessel course, heading, and speed; atmospheric pressure; vessel-relative and true wind direction and speed; sea temperature; air temperature; and humidity (one or more of wet-bulb, dewpoint, relative humidity, specific humidity). Although available in many of the original RV data sets, the current procedure does not create superobs or IMMA records for precipitation or shortwave and longwave radiation.

The averaging procedure for one-minute values must take into account missing/special values and the FSU DAC quality control flags (see appendix B for definitions). For each 11 minute averaging period only "valid" observations will be included in the average. Valid observations must:
1. have a time, latitude, and longitude flagged as good (FSU flags = A, I, N, O, or Z flags)
2. be a non-missing (-9999.), non-special (-8888.) value
3. have a good data value flag (FSU flags = A, G, I, O, or Z flags)

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 s.d. bounds from the da Silva (1994) monthly climatology. These G-flagged values often represent realistic extreme values and must be included in the averages. The statistical test is only applied to wind speed, air and sea temperature, pressure, and relative humidity. The *nn* and *NG* will be stored for each

5

average in the C6 IMMA supplement and will be used to determine the "best" value to be placed into the IMMA core.

As part of the averaging procedure, a sample standard deviation is calculated for the *nn* values used to construct the mean. The standard deviation is included to provide a simple measure of the uncertainty in the mean. It is calculated for all means (when *nn > 1*) except the directional values (HD, CR, RD, and WD). A directional uncertainty will be added in future versions of the FSU-IMMA conversion. When available, it is used as part of the decision making process to select values for the core.

4. Content of C6 supplement for FSU research vessel data
The C6 supplement includes the hourly superobs for each individual navigational, meteorological, and oceanographic parameter included within the original FSU research vessel files. Exclusions will only be made for those variables that FSU deems are, as a whole, of too poor a quality to include. The C6 will be of variable length, dependent upon the number of parameters included in the original data files.

The format of the C6 supplement will include:
1. ATTI=99, ATTL=0 (record terminated by line feed), ATTE=blank
2. The vessel identification indicator (2 character length)
3. The vessel identification (9 character length)
4. The time indicator (1 character length)
5. The record time in the format YYYYMMDDHH (10 character length)
6. For each additional parameter (including latitude and longitude) a group of information will include:
    a. *ID* - An alphabetic identifier signifying what data element follows, see Table 1 (2 character length)
    b. *p* - A single digit signifying the number of duplicate parameter fields to follow (e.g., 2 would signify that 2 different values exist for the parameter from 2 independent sensors)
    c. *data* - The average value, including the sign (when necessary). See Table 1 for details on data field formats as length of data will vary. 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 will be precise to 4 decimal places (0.0001 multiplier).
    d. *sdev* – The sample standard deviation that corresponds to the average data value. All standard deviations will be stored to two decimal precision with a field length of 4. Standard deviation is left missing (blank) when *nn=1*.
    e. *nn* - The number of valid points used in the average (2 digit)
    f. *cc* – The convers_units code for the value from the netCDF file (2 digits). This convers_units code is based on the original units/precision of data received by FSU and is related to the ICOADS indicator values (see Appendix 3). The *cc* will have a value for all parameters in Table 1, except RH. RH units are always % so the *cc* field will be left missing (blank).
    g. *hhh* – The height (depth for TS) of the sensor above (below) mean sea level to the nearest 0.1 meters (stored as 3 digit integer). Only exists for groups with *IDs* marked in Table 1. When height is expected in a group, but value is missing/unknown, height will be shown as three blank spaces.
    h. *NG* – The number of statistical outliers within the *nn* valid points (2 digit) Only exists for groups with *IDs* marked in Table 1.

Table 1: Identifiers to be used by DAC within C6 supplement. Also shown are the original DAC variable identifiers from the netCDF files, the parameter names, the format that the data value would take when multiplied by 0.01 scale factor (0.0001 for LA/LO), and the length of the *data* field in the C6 data group. An * in the *NG* and *hhh* columns indicate that groups with this *ID* will include this element.

C6 *ID*	DAC Variable	Measured Parameter	Format	Length	*hhh*	*NG*
LA	latitude or lat	Latitude (+N)	-xx.xxxx	7		
LO	longitude or lon	Longitude (+E)	xxx.xxxx	7		
SS	PL_SPD	Ship speed	xx.xx	4		
CR	PL_CRS	Ship course	xxx.xx	5		
HD	PL_HD	Ship heading	xxx.xx	5		
RD	PL_WDIR	Ship-relative wind direction	xxx.xx	5	*	
RS	PL_WSPD	Ship-relative wind speed	xx.xx	4	*	
WD	DIR	True wind direction	xxx.xx	5	*	
WS	SPD	True wind speed	xx.xx	4	*	*
PA	P	Atmospheric pressure	xxxx.xx	6	*	*
TS	TS	Sea temperature	-xx.xx	5	*	*
TA	T	Air temperature	-xx.xx	5	*	*
TW	TW	Wet-bulb temperature	-xx.xx	5	*	
TD	TD	Dewpoint temperature	-xx.xx	5	*	
RH	RH	Relative humidity	xxx.xx	5	*	*
QA	Q	Specific humidity	xx.xx	4	*	
PR[†]	PRECIP	Precipitation amount	xxx.xx	5	*	
RR[†]	RRATE	Rain rate	xxx.xx	5	*	
SW[†]	RAD (sw)	Shortwave radiation	xxxx.xx	6	*	
LW[†]	RAD (lw)	Longwave radiation	xxx.xx	5	*	

[†]not included in current version of IMMA records

The format for a single data group would be: *IDpdatasdevnncchhhNG*. When multiple sensors exist, the *IDp* will not be duplicated, just the *datasdevnncchhhNG* part of the group

For example, a ship providing a single value of atmospheric pressure (1024.4, *sdev* = 25.5) at a 14 m sensor height (*nn* = 11, *NG* = 2, *cc* = 2) and two true winds (240° at 13.4 m/s and 255.1° at 15.5 m/s; *sdev* = 3.4 m/s and 1.2 m/s) from sensors at 14 and 25 m (*nn* = 10, *NG* = 0 (speed only), *nn* = 6, *NG* = 2 (speed only), *cc* = 6 (dir) and 2 (spd) for both values) would result in the following data groups in the C6:

PA1102440255011 2140 2WD224000   10 614025510   6 6250WS21340 34010 2140 01550 120 6 2250 2

Other rules that apply when writing the C6 include:
1. Each 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 1).
2. Fill zeros will be used to the right of the decimal for each data value (this simplifies the write and read code). Unused spaces in all fields to the left of a decimal will be filled with blanks, unless the integer value for the field is exactly zero (0).
3. In the FSU C6, longitudes are always reported as positive east values (0. to 359.9999°E)
4. A C6 record will only exist if a valid time, latitude, and longitude exist for the record.

At present, the IMMA record length is limited to 1024 characters. With the current data group format, we are limited to including ~35 parameters for each vessel. With 35 parameters, the record length of the C6 would be between 700 and 800 characters. When added to the core (108) + C1 (65) + C4 (57) = 230 characters, the records should stay under the limit. RV data sets at FSU typically have 30 or fewer parameters.

As is mentioned above, FSU will extract and store the convers_units (*cc*) metadata for each parameter from the FSU netCDF files and store this value in the C6. The convers_units documents the original units and sometimes the precision of the data value originally provided to FSU. All data were then converted to standard SI units. The convers_units are very similar to several of the indicator values in the IMMA core, and in fact were developed based upon earlier versions of the ICOADS documentation (see Appendix 3). The one exception is the *cc* value for sea temperature (TS) which is extracted from the "type" attribute in the original netCDF files. The "type" attribute defines the general category of sensor (e.g., intake, thermosalinograph) use to measure the sea temperature. This netCDF attribute equates more directly to the IMMA ICOADS sea temperature indicator (SI; see Appendix 3).

5. Filling the IMMA core record
The values in the IMMA core are selected from the C6 record using an automated and objective process. The IMMA fields typically filled in the core are highlighted in Table 2 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 C6.

5.1 Selection of "best" values for IMMA core
The selection process will follow a set of rules that will apply to each variable to be assigned to the core (see Table 2). Primarily this process will apply to fill the core variables LAT, LON, D, W, SLP, AT, WBT, DPT, and SST, since cloud and weather information are rarely recorded by automated research vessel data systems. The rules include:
1. A superob to be included in the IMMA core must be derived from a minimum of 5 one-minute observations (only1 observation needed for latitude and longitude). A superob created from less than 5 values may not be representative of the physical atmospheric conditions over the 11 minute averaging period.
2. When multiple superobs are available that have passed rule 1 for one IMMA core field, the superob with the minimum sample standard deviation will be stored in the core.
3. If none of the superobs available for a core element pass rules 1 and 2, the core value will be omitted (set to blanks).

5.2 Setting indicator values for the core
Indicators routinely filled by FSU in the IMMA include: TI, LI, II, DI, WI, IT, WBTI, DPTI, and SI. For all files, the TI = 2 since all times provided are precise to the hour and minute, and the II=1 since all identifications are vessel call signs. For the remaining indicators, a value is chosen based on the convers_units (*cc*) stored in the C6. A one-to-one match does not exist for every convers_units for each indicator value (see Appendix 3) and multiple values and convers_units can exist for each parameter (e.g., 2 air temperatures); therefore, the following rules apply to assigning an indicator value based on the convers_units.
1. Convers_units are only assigned to an indicator when the coded values are an exact match (see Appendix 3).
2. The convers_units assigned to the indicator is the convers_units associated with the value selected as "best" for the core.

Table 2: ICOADS values that FSU will provide in the IMMA core (shaded in the first column). The FSU variable and attribute names identify the location in the FSU RV netCDF files from which the values were extracted. Static values are assigned to every current FSU IMMA record.

ICOADS Abbr.	Element description	FSU Variable	FSU Attribute (Variable or Global)	Static value
YR	year UTC	time (using invtime)		
MO	month UTC	time (using invtime)		
DY	day UTC	time (using invtime)		
HR	hour UTC	time (using invtime)		
LAT	latitude	lat or latitude (depending on netCDF version)		
LON	longitude	lon or longitude (depending on netCDF version)		
IM	IMMA version			0
ATTC	attm count			3
TI	time indicator			2
LI	latitude/long. indic.		convers_units (V)	
DS	ship course†			
VS	ship speed†			
NID	national source indic.			
II	*ID* indicator			1
ID	identification/call sign		ID (G)	
C1	country code			
DI	wind direction indic.		convers_units (V)	
D	wind direction (true)	DIR		
WI	wind speed indicator		convers_units (V)	
W	wind speed	SPD		
VI	*VV* indic.			
VV	visibility			
WW	present weather	WX		
W1	past weather			
SLP	sea level pressure	P		
A	characteristic of *PPP*			
PPP	amt. pressure tend.			
IT	indic. for temperatures		convers_units (V)	
AT	air temperature	T		
WBTI	indic. for WBT		convers_units (V)	
WBT	wet-bulb temperature	TW		
DPTI	*DPT* indic.		convers_units (V)	
DPT	dew-point temp.	TD		
SI	*SST* meas. method		convers_units (V)	
SST	sea surface temp.	TS		
N	total cloud amount	TCA		
NH	lower cloud amount	LMCA		
CL	low cloud type	LCT		
HI	*H* indic.			
H	cloud height	ZCL		
CM	middle cloud type	MCT		
CH	high cloud type	HCT		
WD	wave direction			
WP	wave period			
WH	wave height			
SD	swell direction			
SP	swell period			
SH	swell height			

†Note that values for vessel heading, course, and speed can be extracted from the C6.

6. Filling the C1 attachment

FSU will only fill two fields, the deck number (DCK) and the source ID (SID), in the ICOADS attachment (C1). The remainder of this attachment will be filled by the ICOADS program when these RV records are integrated into the ICOADS.

The DCK will be assigned:

740   Research Vessel (R/V) Data Quality-Evaluated by FSU/COAPS

And the SID will be:

130   Research Vessel (R/V) Data Quality-Evaluated by FSU/COAPS: WOCE ver.3.0
131   Research Vessel (R/V) Data Quality-Evaluated by FSU/COAPS: SAMOS
132   Research Vessel (R/V) Data Quality-Evaluated by FSU/COAPS: Other

The DCK and SID were assigned by the ICOADS group.

7. Filling the C4 attachment

Filling fields in the metadata attachment (C4) resulted in a long discussion with the ICOADS group. The C4 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. Many of the research vessels for which data are available at FSU are not well represented in the VOS program. In addition, some critical metadata, primarily the instrument heights, for the research vessel data from FSU are not part of the IMMA core. If these elements were not placed into the C4, access to this information could only be obtained by reading the complicated C6 format. As a compromise, FSU and the ICOADS group decided to fill a select set of the fields in the C4.

First off, we set the KOV (kind of vessel) field to 'RV', since all of the data provided by FSU were collected by oceanographic research vessels. The KOV is a useful parameter to have filled since it can be used to sort data by vessel type.

FSU also fills the instrument height or depth fields for air temperature (HOT), atmospheric pressure (HOB), winds (HOA), and sea temperature (DOS). The heights are associated with the AT, SLP, D or W, and SST values from the core, respectively. When multiple values exist in the C6 for a parameter (e.g., two air temperatures), then the height assigned (HOT) is for the values selected as "best" by the procedure outlined in section 5.1. Note that heights are stored to the nearest 0.1 m in the C6, but get rounded to the nearest whole meter before being written to the C4.

Appendix A

Details of the programs used to convert the WOCE data from FSU one-minute netCDF files to hourly superobs in IMMA format. The programs are driven by a wrapper called mkcoapsimma and all codes are written in either FORTRAN 77 or 90.

Complete list of programs and supplementary codes for mkcoapsimma process:

  BarretProgs.f
  GetsFromRead_netDoesAvm12.f90
  Makefile
  mkcoapsimma.f
  Module11.f90
  netcdf_inc.f
  NETCDF_INC.mod
  purge
  Read_netcdfm12.f90
  TimeProgs.f
  wrtimma_v003

All other programs necessary to this process are created upon running the Makefile with the command at the unix prompt "make".

Disclaimer:
mkcoapsimma.f and all codes that it calls are open source and not operationally supported. Any questions may be forwarded to samos@coaps.fsu.edu.

Executing code:
mkcoapsimma is the executable of mkcoapsimma.f. It has three required input arguments and an optional fourth option:

1. The directory where mkcoapsimma and the rest of the auxiliary programs reside on the local system.
  e.g. ../makeimma/

2. The directory that contains the list files to be run
  e.g. ../datafiles/KCEJ/

3. This input has two options:
  (a) The listfile that contains all of the files of one minute data from which the hourly averages are created for each cruise.
  e.g. KCEJP__19C-00.list

  (b) When executed, the command 'all' will search the directory for all files named '*.list' and proceed to run mkcoapsimma all list files.

4. '-l' if the user would like to retain all intermediate files created through the process (for debugging purposes). As this is an optional input, it does not need to be entered for general use of the program; simply enter the required three arguments.

Examples:
(NOTE: Below directory structure examples are from generic local directories)

Sample command line:
../makeimma/mkcoapsimma ../makeimma/ ../datafiles/KCEJ/ KCEJ__19C-00.list
Runs all data files in KCEJ__19C-00.list.

Sample command line:
../makeimma/mkcoapsimma ../makeimma/ ../datafiles/KCEJ/ all
Runs all data files in all list files within the /KCEJ/ directory.

Sample command line:
../makeimma/mkcoapsimma ../makeimma/ ../datafiles/KCEJ/ KCEJ__19C-00.list -l
Runs all data files in KCEJ__19C-00.list, retaining all intermediate files.

> NOTE: The contents of the *.list files provided to ICOADS for archival contain paths to local COAPS directories and files. These must be modified in order to run mkcoapsimma in a non-COAPS environment.

Compiling code:
To create the mkcoapsimma executable and all the other auxiliary executables, run the command "make" in the directory in which the FORTRAN programs reside. Running "make" once will create all the necessary programs.

After execution of mkcoapsimma, run "purge" to clean the working directory of all executable files, if desired. The executables are easily re-created by running "make".

Code steps:
mkcoapsimma calls 3 successive programs; Reads12.exec, GetsFrom12.exec, and wrtimma_v003. The first two programs were originally written by James Lamm at COAPS. Barrett Olafson at COAPS modified the first two programs and was the sole author of wrtimma_v003 and the mkcoapsimma wrapper. A schematic of the process is shown in Figure A1.

Reads12.exec is compiled from Read_netcdfm12.f90, TimeProgs.f, and Module11.f90. The purpose of this program is to extract the data values that comprise the super-obs for each 11 minute period at the top of each hour during a cruise. The superobs are created starting at 50 minutes past the hour and continuing to the top of the hour (the 00 minute). The data extracted include the values themselves (and their quality flags), the heights of the sensors, and the conversion units. This program searches each file looking for any parameter on a hard-coded list of variables and attributes that is specific to the original WOCE data files. If one of these hard-coded variables or parameters is not found, a netCDF error message prints to standard out.

These error messages are benign and will not affect the resulting output of the mkcoapsimma procedure. These are output into three separate files, (w/example names): KCEJ__19C-00.list.DataFor.int, KCEJ__19C-00.list.HeightsFor.int, and KCEJ__19C-00.list.ConvUsFor.int, respectively; these are all used by GetsFrom12.exec. The *.int suffix signifies that the file will be scrubbed upon completion of the processing (if the –l argument has not been passed to mkcoapsimma).

GetsFrom12.exec is compiled from GetsFromRead_netDoesAvm12.f90, BarretProgs.f, and Module11.f90. This program will take the 11 minute observations and create one superob for each hour. The program also makes some logic decisions based on the "goodness" of the flags attributed to each value (see section 5.1). If the value has a bad flag, it is excluded from the averaging calculation. All periods must have at least 5 good values to create a super ob, except for latitude and longitude, which need only one observation. The standard deviation of each period is also calculated. A time stamp of the hourly super-ob period is output first. These values are all output in the following order:

CC   - Two character variable for type of value (e.g. LA = latitude).
- Number of records for the variable. This option allows for up to 3 values for a single variable type (e.g. LA, LA2, LA3).
####### - 5 to 7 digit value of the 11-minute averaged data value.
- 4 digit value of the standard deviation for the 11 minute period.
- Two digit value of the number of values from the 11 minute period used in the average.
- Two digit value of the conversion unit from COAPS netCDF files. Only values with a one-to-one map to an ICOADS indicator value are inserted into the core IMMA record.
- Three digit value of the height/depth for the variable's sensor.
- Two digit value of the number of G-flagged values used in the average.

Once each record is written, the program places a sentinel value into the file, to signify what wrtimma_v003 should do as it reads in the values. sentinel values:

-1   - A new variable for the same time period.
-9898 - A new entry (i.e. new time w/new LA, LO, etc.).
-9777 - End of cruise (i.e. end of file).

The output file, e.g. KCEJ__19C-00.IMMA.int, also gets scrubbed (if the -l argument has not been passed to mkcoapsimma).

If any errors occur in the first two steps of the program, the errors are written out to a *.err file; e.g. KCEJ__19C-00.IMMA.err. If there are no errors in there two steps, the error file is deleted.

wrtimma_v003 is compiled from wrtimma_v003.f. This program is a modified version of rdimma0:01G originally written by the ICOADS group and altered by Shawn Smith. wrtimma_v003.f is thoroughly commented, as such this mention will be brief. wrtimma_v003 takes the file, e.g. KCEJ__19C-00.IMMA.int, and actually writes the data into the ICOADS

IMMA format, filing the portions of the core and supplementary blocks, including the C6. If there are multiple values for a single variable, the program decides which is the best value via the decision making process outlined in section 5 above. The final outputs of the entire process are created at this step; the *.sum and *.out files.

The two output files; *.sum and *.out will be placed in the directory from which the program is run (this should be the directory that contains the list files, e.g., ../datafiles/KCEJ/) The *.sum file is a statistical summary file detailing the creation of COAPS IMMA records. The *.sum files are created by code in the original rdimma0:01G code. The *.out file contains the actual IMMA records created from COAPS data. Both of these files will contain the name of the original file w/o *.list as an ending. The naming structure is: (filename).IMMA.(sum or out).

> e.g. KCEJ__19C-00.IMMA.sum and
>   KCEJ__19C-00.IMMA.out

After the entire process has completed, either for a single list file or for all files in a directory, mkcoapsimma scrubs all *.int files (if the -l argument has not been passed to mkcoapsimma). The only files that will remain are the original list files, *.sum and *.out files, as well as any *.err files that may have been created during the process. If the -l option has been entered all *.int files created throughout the running of mkcoapsimma will remain for debugging purposes.

Procedure for placing hourly super-obs from FSU DAC's Research Vessel data files into ICOADS

CruiseFile.list

mkcoapsimma program

Reads12.exec program

CruiseFile.IMMA.err

Error file created if any errors or warnings occur during this step

CruiseFile.list.HeightsFor.int

Height (or depth), if they exist, for all instruments

CruiseFile.list.DataFor.int

All values for all valid instruments within 11 minute super-ob period

CruiseFile.list.ConvUsFor.int

Conversion units (or type), if they exist, for all instruments

GetsFrom12.exec program

CruiseFile.IMMA.err

Error file created if any errors or warnings occur during this step

CruiseFile.IMMA.int

ASCII formatted intermediate file with computed 11-minute super-ob values, standard deviation of the period, number of values in the period, conver sion units, heights of instruments, and the number of g-flagged values used in calculation of the super-ob

wrtimma_v003 program

Note: Upon completion, mkcoapsimma scrubs all *.int files

CruiseFile.IMMA.out

Complete IMMA-formatted record of FSU DAC super-obs for each cruise

CruiseFile.IMMA.sum

ICOADS summary statistics for each variable within each IMMA cruise file
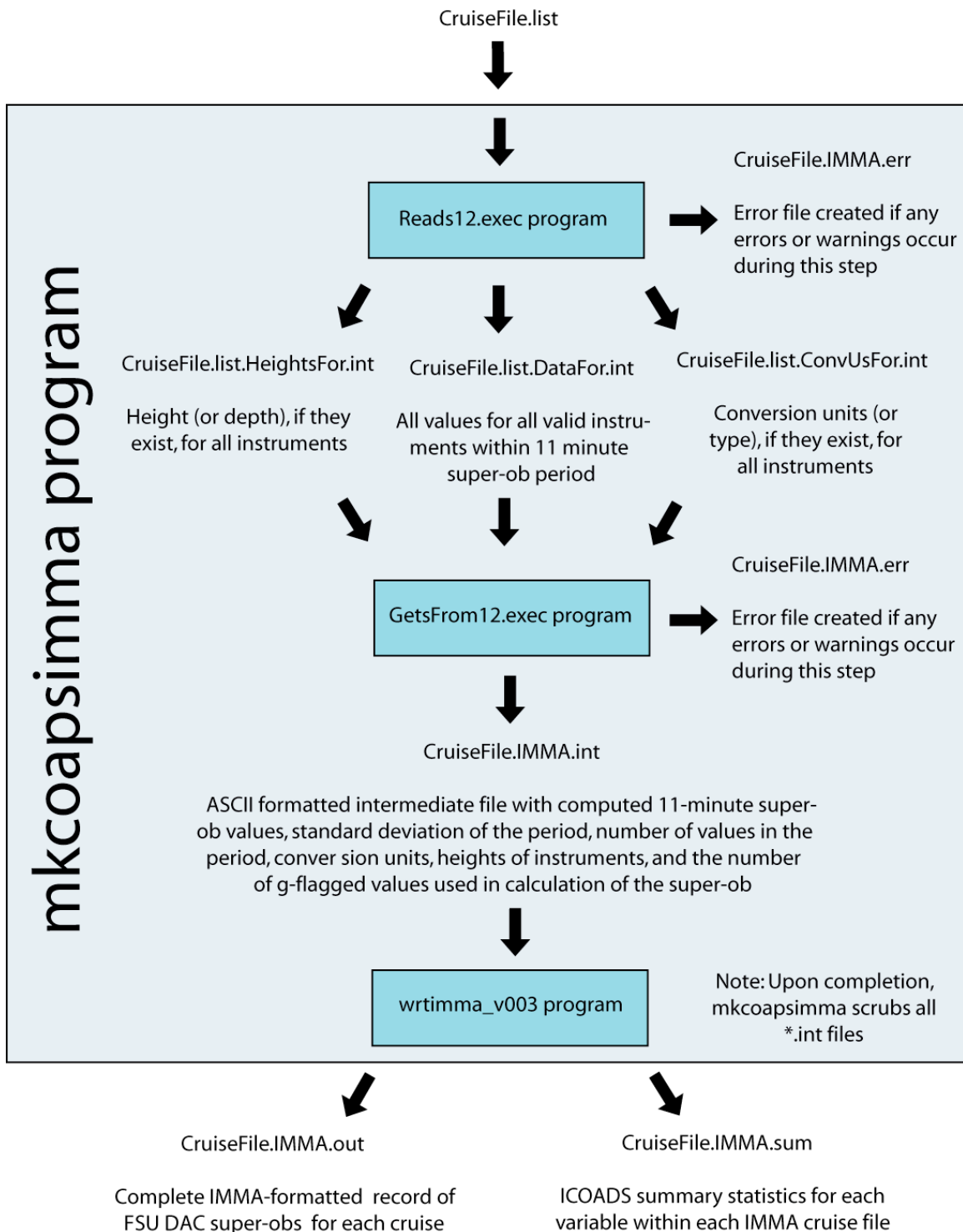
Figure A1: Flow diagram showing steps in the process to convert one-minute sampled WOCE research vessel data from FSU to hourly super-observations for ICOADS. The single input file to mkcoapsimma contains a list of one-minute netCDF files for the cruise to be processed (must be modified to correspond to local file system). The intermediate files (*.int) will be scrubbed unless the -l option is chosen.

Appendix B

A description of FSU DAC quality evaluation flags. A more detailed description of the 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://www.coaps.fsu.edu/RVSMDC/WOCE/qccodes_netcdf.shtml

The flag definitions are:
A: Original data had unknown units. The units shown were determine 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.

Reference:
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.

Appendix C

This appendix contains tables defining the values for the convers_units (*cc* in the C6, section 4 above) in the FSU netCDF files (leftmost column). For the ICOADS indicators LI, DI, WI, TI, WBTI, DPTI, and SI, the tables also show the indicator values used by ICOADS IMMA (rightmost column). Note that when the FSU convers_units and IMMA indicators match, the FSU values can be assigned to the FSU IMMA core (gray in tables). Some of the convers_units are not likely to appear in the cc field of the C6 simply because they relate more to manual (bridge) observations than to modern automated instrumental data. They are included in these tables since FSU netCDF files do exist for some of these earlier bridge reports and these may be incorporated into the ICOADS IMMA at a future date.

In this IMMA release of FSU research vessel data we chose to only map convers_units to IMMA indicators in the core when a one-to-one mapping exists. This was done as a compromise to keep the conversion codes fairly straightforward but does exclude some rather obvious mappings (e.g., for wind speed measured in m/s). Modifications of this approach may be made in future releases, including the possibility of creating a new IMMA value for each indicator that represents "data averaged from higher-sampling rate observations".

Latitude and longitude (Indicator mapping for LI):

FSU convers_units (*cc*)	FSU Description	IMMA
0	degrees and tenths	0
1	whole degrees	1
2	non-random tenths	2
3	interpolated	3
4	degrees and minutes	4
5	high resolution data (e.g., degrees to seconds, or degrees and hundredths)	5
6	other (refer to supplemental data)	6

Atmospheric pressure (no IMMA mapping):

FSU convers_ units (*cc*)	FSU Description	IMMA
0	millibars	
1	bars	
2	millimeters of mercury	
3	inches of mercury	
4	pascals	
5	high resolution (0.01 mb)	

Specific humidity (no IMMA mapping):

FSU convers_ units (*cc*)	FSU Description	IMMA
0	grams per kilogram	
1	kilograms per kilogram	

Directional parameters (Course, heading, platform-relative and true winds). Also used as indicator mapping for DI:

FSU convers_ units (*cc*)	FSU Description	IMMA
0	36-point compass	0
1	32-point compass	1
2	16 of 36-point compass	2
3	16 of 32-point compass	3
4	8-point compass	4
5	360-point compass	5
6	high resolution data (e.g., tenths of degrees)	6
7	16 point compass*	
8	u,v components	

* a generic 16 point compass with bins centered on 0, 22.5, 45, …315, 335.5 (basically a N, NNE, NE,…NW, NNW compass).

Speed parameters (ship speed, platform-relative and true winds). Also used as indicator mapping for WI:

FSU convers_ units (*cc*)	FSU Description	IMMA
0	meters per second, estimated	0
1	meters per second, unknown	
2	meters per second, measured	1
3	knots, estimated	3
4	knots, unknown	
5	knots, measured	4
6	Beaufort force (based on documentation)	5
7	high resolution measurement (e.g., hundredths of a meter per second)	8
8	u,v components (m/s)	
9	u,v components (knots)	

For both speed and direction, FSU uses codes for when the original data are received as u,v vector components. The direction and speed are calculated from the components.

Air temperatures (air, wet-bulb, and dewpoint temperatures). Also used as indicator mapping for TI, WBTI, DPTI:

FSU convers_ units (*cc*)	FSU Description	IMMA
0	degrees Celsius and tenths	0
1	half degrees Celsius	1
2	whole degrees Celsius	2
3	degrees Fahrenheit and tenths	4
4	half degrees Fahrenheit	5
5	whole degrees Fahrenheit	6
6	high resolution data (e.g., hundredths of a degree Celsius or Fahrenheit)	8
7	other (notes will be found in data quality reports)	9

Sea temperature (Indicator mapping for SI):

FSU convers_ units (*cc*)	FSU Description	IMMA
0	unknown	
1	condenser inlet (intake)	1
2	trailing thermistor	2
3	hull contact sensor	3
4	through hull sensor	4
5	radiation thermometer	5
6	bait tanks thermometer	6
7	others	7
8	bucket	0
9	implied bucket (an HSST SID or any match)	10
10	reversing thermometer	11
11	other electronic sensor	12
12	thermosalinograph	