Thompson IMET Data Quality Control Report

Jennifer Lovell and Shawn R. Smith

Research Vessel Surface Meteorological Data Center

Center for Ocean Atmospheric Prediction Studies

Florida State University

October 24, 2000

Report 00-07

Version 1.0

*1.0 **Introduction***

This report summarizes the quality of surface meteorological data collected by the research vessel *Thompson* (identifier: WSRY) IMET system during twenty-one cruises which occurred over a four year span, beginning 18 September 1994 and ending 3 March 1997. The data sets were provided to the Florida State University Research Vessel Surface Meteorological Data Center (RVSMDC) in electronic format by Bill Martin and were converted to standard RVSMDC netCDF format. The data were then processed using an automated screening program, which added quality control flags to the data, highlighting potential problems. Finally, the Data Quality Evaluator (DQE) reviewed the data and current flags, whereby flags were added, removed, or modified according to the judgment of the DQE and other RVSMDC personnel. Details of the RVSMDC quality control procedures can be found in Smith et al. (1996). The data quality control report summarizes the flags for the *Thompson* IMET surface meteorological data, including those added by both the preprocessor and the DQE.

*2.0 **Data Variables***

The *Thompson* IMET data are expected to include observations taken every minute on these cruises. Values for the following variables were collected:

Time	(TIME)
Latitude	(LAT)
Longitude	(LON)
Platform Heading (Gyrocompass)	(PL_HD)
Platform Course	(PL_CRS)
Platform Speed Over Ground	(PL_SPD)
Platform Speed Over Water	(PL_SPD2)
Platform Relative Wind Direction (IMET)	(PL_WDIR)
Platform Relative Wind Speed (IMET)	(PL_WSPD)
Earth Relative Wind Direction (IMET)	(DIR)
Earth Relative Wind Speed (IMET)	(SPD)
Atmospheric Pressure	(P)
Air Temperature	(T)
Sea Temperature	(TS)
Relative Humidity	(RH)
Atmospheric Radiation	(RAD)
Rain Rate (mm/min)	(RRATE)
Rain Rate 2 (mm/hr)	(RRATE2)
Precipitation	(PRECIP)

*3.0 **Cruise Identifiers and Dates***

Note: The Cruise Identifiers and Cruise Dates were assigned to the Thompson cruises by the DQE for cruise identification in the quality control report. The beginning and ending dates of each cruise were determined by the Thompson's departure and return dates to port successively.

Cruise Identifiers	Cruise Dates
1994	
94-A	09/18/94 – 10/07/94
94-B	10/11/94 – 10/25/94
94-C	10/28/94 – 11/21/94
94-D	11/28/94 – 12/19/94
1995	
95-A	02/08/95 – 02/28/95
95-B	04/13/95 – 04/29/95
95-C	06/21/95 – 07/13/95
95-D	09/19/95 – 10/11/95
95-E	10/14/95 – 10/25/95
1996	
96-A	01/18/96 – 01/21/96
96-B	01/22/96 – 01/26/96
96-C	01/30/96 – 03/10/96
96-D	03/13/96 – 04/02/96
96-E	04/03/96 – 04/11/96
96-F	04/15/96 – 05/14/96
96-G	05/18/96 – 06/27/96
96-H	07/08/96 – 07/22/96
96-I	08/12/96 – 08/27/96
96-J	10/12/96 – 10/20/96
1997	
97-A	01/07/97 – 01/25/97
97-B	01/31/97 – 03/03/97

4.0 Overall Quality

The overall quality of the entire data set was good. A total of 10,871,990 values were evaluated with 390,421 flags added by both the preprocessor and the DQE resulting in a total of 3.59% of the values being flagged. Specific details of each cruise are covered in the following sections.

5.0 Major Problems

5.1 H-flags:

Discontinuous data values were assessed the H-flag. Variables such as, latitude (LAT), longitude (LON), and sea temperature (TS) received H-flags during periods of discontinuous data. The erroneous values that occurred occasionally between discontinuous data values were given K-flags to show caution to the user.

<u>5.2 K-flags</u>:

The K-flag represents suspect data and should be used with caution. Throughout each of the four years of data, numerous data were assessed the K-flag. The most significant use of the K-flag was to reveal signatures of ship motion in the variables. Variables such as earth relative wind direction (DIR), earth relative wind speed (SPD), atmospheric pressure (P), temperature (T), and atmospheric radiation (RAD) showed stair steps in the data. These stair steps are related to a change in platform course (PL_CRS), heading (PL_HD), and/or platform speed (PL_SPD) and should not exist in earth relative data. Subsequently, the data was flagged as suspect.

Each year temperature received K-flags for problems other than stair stepping. The first of which was due to radiational heating of the ship. When the platform relative wind speed was low, ~3 ms^{-1} or less, significant increases in temperature were occurring during daylight hours. The second problem was ventilation, which occurred when the platform wind direction was from around 180 degrees. This likely affected the flow of the air prior to reaching the bow-mounted thermometer. In these instances, significant increases in temperature were flagged as cautionary.

Note: Other K-flag occurrences will be addressed in the yearly summaries below.

<u>5.3 J-flags</u>:

Data of poor quality by visual inspection were given the J-flag and should NOT be used. Each year had some data of poor quality, but 1995 was extreme with 17,421 J-flags. J-flags were assigned to many different variables of the 1995 data that had smooth, yet rigid signatures, which lacked expected variability. Another common use of the J-flag was to show areas of flat-lined data. For example, during the 95-D cruise, sea temperature (TS) had a problem recording data and was "stuck" on one particular value (35.0 degrees Celsius) for a few hours. These occurrences were J-flagged because the DQE believed the data were physically unrealistic. The J-flag was also used to show erroneous data. During the 97-B cruise, the platform heading (PL_HD) data flat-lined on the value 9999. The data received 164 J-flags and should NOT be used.

<u>5.4 Spikes</u>:

Isolated spikes occurred in most of the variables throughout the data. Spikes are a relatively common occurrence with automated data, caused by various factors (e.g. electrical interference, ship movement, etc.). These individual points were assigned the S-flag.

6.0 <u>1994 FLAG SUMMARY</u>

Statistical Information:

Details of each 1994 cruise are listed in Table 1 and include cruise dates, number of records, number of values, number of flags, and total percentage of data flagged. A total of 2,103,568 values were evaluated with 71,471 flags added by both the preprocessor and the DQE resulting in a total of 3.40% of the values being flagged.

Table 1: Statistical Cruise Information

Cruise Identifier	Cruise Dates	Number of Records	Number of Values	Number of Flags	Percent Flagged
94-A	09/18/94 – 10/07/94	27,510	495,180	2,022	0.41
94-B	10/11/94 – 10/25/94	19,968	379,392	28,416	7.49
94-C	10/28/94 – 11/21/94	34,470	654,930	27,086	4.14
94-D	11/28/94 – 12/19/94	30,214	574,066	13,947	2.43

Summary:

The 1994 IMET data from the *Thompson* proves to be of good quality with 3.40% of the reported values being flagged for potential problems. The distribution of flags for the remaining variables is detailed in Table 2.

Table 2: Number of Flags and Percentage Flagged for Each Variable

Variable	G	H	J	K	S	Total Number of Flags	Percentage of Variable Flagged
LAT						0	0.00
LON						0	0.00
PL_HD				356	6	362	0.32
PL_CRS					4	4	0.00*
PL_SPD					524	524	0.47
PL_SPD2					8	8	0.01
PL_WDIR						0	0.00
PL_WSPD						0	0.00
DIR			37	25,907	47	25,991	23.17
SPD			39	27,265	2	27,306	24.35
P				8,003	28	8,031	7.16
T	4			3,257	112	3,373	3.01
TS		4		5,298	80	5,382	4.80
RAD				487	3	490	0.44
RRATE						0	0.00
RRATE2						0	0.00
PRECIP						0	0.00
Total Number Of Flags	4	4	76	70,573	814	71,471	
Percent Of All Values Flagged	0.00*	0.00*	0.00*	3.31	0.04	3.40	

*Percentages<0.01

G-flags:

There were four consecutive data values that were assessed the G-flag on temperature (T) during the 94-C cruise. The DQE felt these values were realistic, as they were approximately 4 degrees

Celsius lower than the given data trend. The G-flags were left in place to highlight values that are greater than four standard deviations from the climatological mean (da Silva et al. 1994).

H-flags:

Sea temperature (TS) received four H-flags during the 1994 cruises. Two of the H-flags occurred on the first day of the 94-A cruise when the ship was leaving port. The sea temperature showed a discontinuity in the data when it dropped approximately three degrees Celsius in five minutes. An H-flag was placed at the end of the 94-A cruise where the sea temperature increased approximately four degrees Celsius and then remained higher than normal for the rest of the cruise and into the beginning of the 94-B cruise. Sea temperature returned to normal values about four days into the 94-B cruise where an H-flag was placed to indicate the end of discontinuity. The data values between the two H-flags were assessed K-flags and should be used with caution.

J-flags:

Earth relative wind direction (DIR) and earth relative wind speed (SPD) were assessed a total of 76 J-flags during the 94-D cruise. These J-flags were used to highlight unrealistic steps that resembled a block-like pattern in the data and were not associated with ship movement.

Deleted data:

It was determined by the DQE that the relative humidity (RH) data in the 94-A cruise not be reported in the public release of the 1994 data. During the 94-A cruise, the relative humidity data recorded values between 100% and 180% for the majority of the cruise. This was most likely caused by improper calibration or failure of the instrument.

Missing data:

During the entire 94-A cruise and first five days of the 94-B cruise, platform relative wind direction (PL_WDIR), platform relative wind speed (PL_WSPD), earth relative wind direction (DIR), and earth relative wind speed (SPD) were missing. Due to the missing winds, deciphering the quality of other meteorology variables was difficult. The DQE cautions the use of the meteorological data, as much of the data left unflagged may still be questionable. *Note: The meteorological data was NOT K-flagged because of the missing winds.*

Spikes: See *5.4 Spikes.*

7.0 **1995 FLAG SUMMARY**

Statistical Information:

Details of each 1995 cruise are listed in Table 3 and include cruise dates, number of records, number of values, number of flags, and total percentage of data flagged. A total of 2,457,023 values were evaluated with 125,872 flags added by the preprocessor and the DQE resulting in 5.12% of the values being flagged.

Table 3: Statistical Cruise Information

CTC	Cruise Dates	Number of Records	Number of Values	Number of Flags	Percent Flagged
95-A	02/08/95 - 02/28/95	28,394	539,486	10,725	1.99
95-B	04/13/95 - 04/29/95	21,635	411,065	4,750	1.16
95-C	06/21/95 - 07/13/95	31,688	602,072	76,665	12.73
95-D	19/19/95 - 10/11/95	31,676	601,844	27,813	4.62
95-E	10/14/95 - 10/25/95	15,924	302,556	5,919	1.96

Summary:

The 1995 IMET data from the *Thompson* proves to be of good quality with 5.12% of the reported values being flagged for potential problems. Table 4 details the distribution of flags among the variables.

Table 4: Number of Flags and Percentage Flagged for Each Variable

Variable	G	H	J	K	S	Total Number of Flags	Percentage of Variables Flagged
LAT		24		29	6	59	0.05
LON		27		29	5	61	0.05
PL_HD					14	14	0.01
PL_CRS					17	17	0.01
PL_SPD					1,058	1,058	0.82
PL_SPD2					89	89	0.07
PL_WDIR			872	4,930	28	5,830	4.51
PL_WSPD			1,541	5,067	1	6,609	5.11
DIR			287	24,253	125	24,665	19.07
SPD			314	31,969	13	32,296	24.97
P	2,547		3,546	11,260	4	17,357	13.42
T			2,728	10,735		13,463	10.41
TS	408		1,956	1,079	304	3,747	2.90
RH			3,426	8,016	6	11,448	8.85
RAD			2,751	5,030	1	7,782	6.02
RRATE				439		439	0.34
RRATE2				499		499	0.39
PRECIP				439		439	0.34
Total Number of Flags	2,955	51	17,421	103,774	1,671	125,872	
Percentage of All Values Flagged	0.12	0.00*	0.71	4.22	0.07	5.12	

*Percentages < 0.01

G-flags:

The G-flag represents a value that is greater than four standard deviations from the climatological mean (da Silva et al. 1994). Pressure (P) was assessed 2,547 G-flags during the 95-D cruise. These flags were left in place since the values were only 1 to 3 millibars higher than the given data trend.

Sea temperature (TS) received 408 G-flags during the 95-D cruise, too. The DQE felt these flags were realistic, as they were approximately 1 degree Celsius higher than the given data trend.

H-flags:

Latitude (LAT) and longitude (LON) received a total of 51 H-flags throughout the 95-A, 95-B, and 95-C cruises. These H-flags illustrate a discontinuity in the ship's position. The data values between H-flags were given K-flags and should be used with caution.

J-flags:

Most of the J-flags and K-flags assessed to the 1995 cruises were assigned to values of data severely lacking expected variability and unrealistically out of the given data trend. The main difference between the use of the J-flag and K-flags was that data which was J-flagged did not resemble true data, while data that was K-flagged looked liked true data but lacked expected variability (e.g. Fig 7.1).

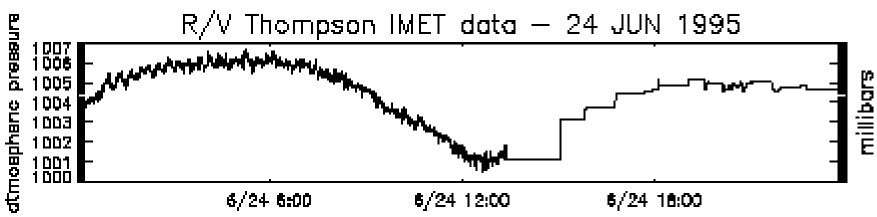

Figure 7.1 Example of K-flagged Thompson atmospheric pressure (P) on June 24, 1995. Data look good, until approximately 13:00 UTC. After 13:00 UTC, the data contains the expected diurnal pattern, but lacks variability (K-flags applied from13:00 to 23:59 UTC).

During two consecutive days in the 95-C cruise, platform relative wind direction (PL_WDIR) and platform relative wind speed (PL_WSPD) received numerous J-flags for repeated data. These variables flat-lined at 50 degrees Celsius for PL_WDIR and 11 ms^{-1} for PL_WSPD.

K-flags:

Nearly all of the 103,774 suspect data flags (K) assigned to atmospheric pressure (P), earth relative wind direction (DIR), earth relative wind speed (SPD), temperature (T), and atmospheric radiation (RAD) were due to signatures of ship motion in the variables. The discontinuous stair steps in the data that were related to a change in platform course, heading, and/or speed should not exist in earth relative data and were subsequently flagged as suspect.

Stair stepping occurred with pressure (P) throughout the data sets. There were some stair steps in the pressure data that were a result of a change in either forward speed or direction. These stair steps were associated with approximately a ½ millibar (mb) decrease in pressure relative to both the forward speed and direction change of the ship. However, there were some stair steps in the pressure data that were not a result of the ship motions. These stair steps were related to the ship relative winds and increased pressure approximately ½ mb when the platform wind direction was approximately 0 degrees.

The earth relative wind direction (DIR) and earth relative wind speed (SPD) had stair steps occurring throughout the data sets. The cause was likely due to flow distortion. Flow distortion is the disturbance of airflow from other objects or instruments upstream from the anemometer. The significance of the stair stepping varied throughout the data set; therefore, the earth relative winds should be used with caution.

Temperature (T) had problems that were more specific. The first of which was due to radiational heating of the ship. When the platform relative wind speed (PL_WSPD) was low, ~3 ms^{-1} or less, significant increases in temperature were occurring during daylight hours. The second problem was ventilation, which occurred when the platform wind direction (PL_WDIR) was from around 180 degrees. This likely affected the flow of the air prior to reaching the bow-mounted thermometer. In these instances, significant increases in temperature were flagged as cautionary.

Spikes: See *5.4 Spikes.*

8.0 1996 FLAG SUMMARY

Statistical Information:

Details of each 1996 cruise are listed in Table 5 and include cruise dates, number of records, number of values, number of flags, and total percentage of data flagged. A total of 4,935,269 values were evaluated with 154,411 flags added by both the preprocessor and the DQE resulting in a total of 3.13% of the values being flagged.

Table 5: Statistical Cruise Information

CTC	Dates	Number of Records	Number of Values	Number of Flags	Percent Flagged
96-A	01/18/96 – 01/21/96	4,473	84,987	69	0.08
96-B	01/22/96 – 01/26/96	6,066	115,254	239	0.21
96-C	01/30/96 – 03/10/96	56,818	1,079,542	14,873	1.38
96-D	03/13/96 – 04/02/96	28,246	536,674	942	0.18
96-E	04/03/96 – 04/11/96	11,466	217,854	2,958	1.36
96-F	04/15/96 – 05/14/96	42,122	800,318	44,003	5.50
96-G	05/18/96 – 06/27/96	58,101	1,103,919	73,436	6.65
96-H	07/08/96 – 07/22/96	19,239	365,541	10,833	2.96
96-I	08/12/96 – 08/27/96	21,511	408,709	6,587	1.61
96-J	10/12/96 – 10/20/96	11,709	222,471	471	0.21

Summary:

The 1996 IMET data from the *Thompson* proves to be of good quality with 3.13% of the reported values being flagged for potential problems. The distribution of flags for the variables is detailed in Table 6.

Table 6: Number of Flags and Percentage Flagged for Each Variable

Variable	B	G	H	J	K	S	Total Number of Flags	Percentage of Variable Flagged
LAT			8				8	0.00*
LON			6				6	0.00*
PL_HD						7	7	0.00*
PL_CRS						3	3	0.00*
PL_SPD						2,177	2,177	0.84
PL_SPD2						2	2	0.00*
PL_WDIR							0	0.00
PL_WSPD						2	2	0.00*
DIR					15,178	109	15,287	5.89
SPD					19,014	6	19,020	7.32
P		96,836			6,336	9	103,181	39.72
T		4			5,745	387	6,136	2.36
TS		3,916		584		905	5,405	2.08
RH	1,035						1,035	0.40
RAD	2			161	1,977	2	2,142	0.82
RRATE							0	0.00
RRATE2							0	0.00
PRECIP							0	0.00
Total Number Of Flags	1,037	100,756	14	745	48,250	3,609	154,411	
Percent Of All Values Flagged	0.02	2.04	0.00*	0.02	0.98	0.07	3.13	

*Percentages<0.01

Bounds Flags:

Relative humidity (RH) received 260 B-flags during the 96-C cruise and 775 B-flags during the 96-H cruise. In both instances, the flagged values were above 100%, but below 101%.

There were two B-flags assessed to atmospheric radiation (RAD). One flag occurred on the 96-C cruise and the other on the 96-F cruise, both with a value of $-0.1 Wm^{-2}$. This physically unrealistic negative radiation value is likely the result of the instrument not being tuned to low radiation values.

G-Flags:

Temperature (T) was assessed four G-flags during the 96-D cruise. The data was approximately ½ degree Celsius lower than the given data trend. The DQE left the flag in place to highlight a value that is greater than four standard deviations from the climatological mean (da Silva et al. 1994).

Pressure (P) was extremely high (i.e. 1035 millibars) and was given 96,836 G-flags over several cruises. These flags do not signify bad data, but rather draw attention to data that is greater than four standard deviations from the climatological mean (da Silva et al. 1994). During these cruises, the ship traversed the South Atlantic seas, south of 40 degrees South Latitude. In this region of the globe, little is known of climatology, as the data is sparse. Consequently, the G-flagged data values may be realistic, though extreme observations.

Sea temperature (TS) was given 3,916 G-flags by the preprocessor during two different cruises. Similar to pressure (P), these data are flagged because the climatology is poorly known in the Thompson's region of operation.

H-flags:

Latitude (LAT) and longitude (LON) received a total of 14 H-flags throughout the 96-E, 96-G, and 96-H cruises. These H-flags illustrate a discontinuity in the ship's position.

J-flags:

Atmospheric radiation (RAD) received 161 J-flags during the 96-C cruise. These J-flags occurred because data was flat-lined at zero Wm^{-2} on the last day of the cruise.

Sea temperature (TS) received 584 J-flags the first day of the 96-D cruise. The data was extremely noisy, varying from zero to 35 degrees Celsius within a few hours.

K-flags:

K-flags were used to reveal signatures of ship motion in certain meteorological variables. Variables such as earth relative wind direction (DIR), earth relative wind speed (SPD), atmospheric pressure (P), temperature (T), and atmospheric radiation (RAD) showed stair steps in the data. These stair steps were related to a change in platform course (PL_CRS), heading (PL_HD), and/or platform speed (PL_SPD) and should not exist in earth relative data. Subsequently, the data was flagged as suspect.

Temperature (T) was assessed several K-flags due to radiational heating of the ship. When the platform relative wind speed was low, ~3 ms^{-1} or less, significant increases in temperature were occurring during daylight hours. The second problem in the temperature (T) data was a ventilation problem, which occurred when the platform wind direction was from around 180 degrees. This likely affected the flow of the air prior to reaching the bow-mounted thermometer. In these instances, significant increases in temperature were flagged as cautionary.

Spikes: See *5.4 Spikes.*

9.0 1997 FLAG SUMMARY

Statistical Information:

Details of each 1997 cruise are listed in Table 7 and include cruise dates, number of records, number of values, number of flags, and total percentage of data flagged. A total of 1,348,620 values were evaluated with 13,040 flags added by both the preprocessor and the DQE resulting in 0.97% of the values being flagged.

Table 7: Statistical Cruise Information

CTC	Dates	Number of Records	Number of Values	Number of Flags	Percent Flagged
97-A	01/07/97 – 01/25/97	26,258	498,902	2,125	0.43
97-B	01/31/97 – 03/03/97	44,722	849,718	10,915	1.28

Summary:

The 1997 IMET data from the *Thompson* proves to be of excellent quality with 0.97% of the reported values being flagged for potential problems. The distribution of flags for the remaining variables is detailed in Table 8.

Table 8: Number of Flags and Percentage Flagged for Each Variable

Variable	G	H	J	K	S	Total Number of Flags	Percentage of Variable Flagged
LAT		4		2		6	0.01
LON		42		8	1	51	0.07
PL_HD			164		9	173	0.24
PL_CRS					4	4	0.01
PL_SPD					696	696	0.98
PL_SPD2					1	1	0.00*
PL_WDIR					2	2	0.00*
PL_WSPD					20	20	0.03
DIR				3,330	39	3,369	4.75
SPD				3,928	13	3,941	5.55
P	600			2,151	111	2,262	4.03
T				1,319	5	1,924	1.87
TS					442	442	0.62
RH					149	149	0.21
RAD						0	0.00
RRATE						0	0.00
RRATE2						0	0.00
PRECIP						0	0.00
Total Number Of Flags	600	48	164	10,738	1,492	13,040	
Percent Of All Values Flagged	0.04	0.00*	0.01	0.80	0.11	0.97	

*Percentages<0.01

<u>G-flags:</u>

Pressure (P) was given 600 G-flags during the 97-A cruise. The DQE felt these values were realistic as they were only 1 to 5 millibars lower than the given data trend. The G-flags were left in place to highlight values that were greater than four standard deviations from the climatological mean (da Silva et al. 1994).

H-flags:

Latitude (LAT) and longitude (LON) received a total of 46 H-flags throughout the1997 cruises. These H-flags illustrate a discontinuity in the ship's position. The data values between H-flags were given K-flags and should be used with caution.

J-flags:

During the 97-B cruise, platform heading (PL_HD) flat-lined on the value 9999. The data received 164 J-flags and should NOT be used.

K-flags:

K-flags were used to reveal signatures of ship motion in certain meteorological variables. Variables such as earth relative wind direction (DIR), earth relative wind speed (SPD), atmospheric pressure (P), and temperature (T) showed stair steps in the data. These stair steps were related to a change in platform course (PL_CRS), heading (PL_HD), and/or platform speed (PL_SPD) and should not exist in earth relative data. Subsequently, the data was flagged as suspect.

Temperature (T) was assessed several K-flags due to radiational heating of the ship. When the platform relative wind speed was low, ~3 ms^{-1}or less, significant increases in temperature were occurring during daylight hours. The second problem in the temperature (T) data was a ventilation problem, which occurred when the platform wind direction was from around 180 degrees. This likely affected the flow of the air prior to reaching the bow-mounted thermometer. In these instances, significant increases in temperature were flagged as cautionary.

Spikes:

See *5.4 Spikes.*

10.0 Final Discussion

Precipitation (PRECIP), rain rate (mm min^{-1}) (RRATE), and rain rate 2 (mm hr^{-1}) (RRATE2) were not thoroughly quality controlled since insufficient metadata were available about these variables. All three variables should be used with caution.

The *Thompson's* platform speed over ground (PL_SPD) data was extremely noisy. The DQE recommends the final user to apply a smoother on the data.

The quality of the data, especially earth relative data, improved over the course of the four-year span. This is due, impart, to improved communication between the *Thompson* operators and the RVSMDC.

11.0 References:

Smith, S.R., C. Harvey, and D.M. Legler, 1996: *Handbook of Quality Control Procedures and Methods for Surface Meteorology Data*. WOCE Report No. 141/96, Report WOCEMET 96-1, Center for Ocean-Atmospheric Prediction Studies Florida State University, Tallahassee FL 32306-2840

da Silva, A.M., C.C. Young and S. Levitus, 1994: *Atlas of Surface Marine Data 1994, Volume 1: Algorithms and Procedures*. NOAA Atlas Series.