Polarstern Data Quality Control Report

Melissa Griffin and Shawn R. Smith

Research Vessel Surface Meteorological Data Center

Center for Ocean Atmospheric Prediction Studies

Florida State University

January 10, 2001

Report 01-01

Version 1.0

Introduction:

This report summarizes the quality of surface meteorological data collected by the research vessel *Polarstern* (identifier: DBLK) during thirteen cruises beginning 18 May 1993 and ending 8 November 1996. The data were provided to the Florida State University Research Vessel Surface Meteorological Data Center (RVSMDC) in electronic format by G.Koenig-Longlo 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 *Polarstern* meteorological data, including those added by both the preprocessor and the DQE.

Data Variables:

The *Polarstern* data are expected to include observations averaged over ten minute periods on these cruises. Values for the following variables were collected:

Time	(TIME)
Latitude	(LAT)
Longitude	(LON)
Platform Heading	(PL_HD)
Platform Course	(PL_CRS)
Platform Speed	(PL_SPD)
Platform Relative Wind Direction	(PL_WDIR)
Platform Relative Wind Speed	(PL_WSPD)
Earth Relative Wind Direction	(DIR)
Earth Relative Wind Speed	(SPD)
Ocean Relative Wind Direction	(DIR2)
Ocean Relative Wind Speed	(SPD2)
Sea Temperature	(TS)
Atmospheric Pressure	(P)
Air Temperature	(T)
Dewpoint Temperature	(TD)
Relative Humidity	(RH)
Specific Humidity	(Q)
Rain Rate	(RRATE)

Cruise Identifiers and Dates:

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

1993

Cruise Identifiers	Cruise Dates
93-A	05/18/93 - 10/05/93
93-B	10/18/93 - 11/27/93
93-C	12/10/93 - 01/29/94

1994

94-A	02/11/94 - 02/28/94
94-B	04/01/94 - 05/19/94
94-C	05/31/94 - 06/17/94
94-D	07/05/94 - 10/06/94
94-E	11/23/94 - 06/10/95

1995

95-A	07/07/95 - 10/29/95
95-B	11/09/95 - 03/08/96

1996

96-A	05/21/96 - 06/20/96
96-B	07/12/96 - 09/22/96
96-C	10/05/96 - 11/08/96

Overall Quality:

The overall quality of the entire data set from the *Polarstern* proves to be excellent. A total of 2,686,163 values were evaluated with 15,582 flags added by both the preprocessor and the DQE resulting in 0.58 % of the values being flagged. Specific details of each cruise are covered in the following sections.

Major Problems:

Missing Data:

On the first day of each cruise, many of the variables had missing data. Earth Relative Wind Direction (DIR), Earth Relative Wind Speed (SPD), Platform Heading (PL_HD), and Platform Course (PL_CRS) were variables that reported some missing data throughout the data set. Due to the missing data, the DQE was unable to accurately identify occurrences of stair stepping in the data (frequently associated with flow distortion). Without radiation data, the DQE was unable to correctly identify occurrences of ship heating.

Spikes:

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.

Relative Humidity:

Our initial analysis of the data showed that throughout the four years, when Temperature (T) and Dew Point Temperature (TD) were missing from the data the relative humidity would spike to 100%. The relative humidity values would flat-line at 100% until the next time values for temperature and dewpoint temperature would appear in the data set. We became curious about this situation and consulted the data provider on this issue. The provider informed us that the relative humidity, during this type of situation, should have missing data as well. A program was created to change the erroneous data to the correct missing value of –9999, whenever T and TD were missing.

Precipitation:

Precipitation measurements on board the *Polarstern* started on 18 October 1994. Only rain events are normally quantified. Due to several problems with the sensor, the data should be used with caution. Data, which was obviously wrong, was excluded from the database. Nevertheless, it is still possible that some rain events are missing. Extreme sea spray events could affect the measurements and accumulated snow would melt within the sensor. The melting snow would be measured as liquid precipitation hours or days after the snowfall.

1993 FLAG SUMMARY

Statistical Information:

Details of the cruises 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 629,850 values were evaluated with 2,892 flags added by both the preprocessor and the DQE resulting in 0.46% 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
93-A	05/18/93 - 10/05/93	20,011	380,209	1,676	0.44
93-B	10/18/93 - 11/27/93	5,796	110,124	154	0.14
93-C	12/10/93 - 01/29/94	7,343	139,517	1,062	0.76

Summary:

The 1993 data from the *Polarstern* proves to be of excellent quality with 0.46 % 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	F	G	K	S	Total Number of Flags	Percentage of Variable Flagged
LAT	4				4	0.01
LON	4				4	0.01
PL_HD					0	0.00
PL_CRS					0	0.00
PL_SPD					0	0.00
PL_WDIR					0	0.00
PL_WSPD					0	0.00
DIR					0	0.00
SPD					0	0.00
DIR2					0	0.00
SPD2		122			122	0.37
TS		827	20		847	2.56
P			182		182	0.55
T		673	181		854	2.58
TD			176	5	181	0.55
RH		1	489	23	513	1.55
Q			180	5	185	0.56
RRATE					0	0.00
Total Number Of Flags	8	1,623	1,228	33	2,892	
Percent Of All Values Flagged	0.00*	0.26	0.19	0.01	0.46	

*Percentages < 0.01

F-Flag:

Latitude and Longitude were both assessed four F-flags by the preprocessor during the 93-A and 93-B cruises. These flags show that the platform speed computed by the preprocessor exceeds a realistic speed for a research vessel (15 meters per second). This may have been caused by uncertainties or truncation error in the navigation data.

G-Flag:

Of the 1,623 G-flags assessed during the three cruises in 1993, 1,336 of them occurred during the 93-A cruise. The *Polarstern* spent a lot of the time exploring the Arctic Circle, reaching as far East as the Kara Sea, during the 93-A cruise. The climatology in this area is somewhat questionable due to the lack of data at high latitudes.

Ocean Relative Wind Speed (SPD2) was assessed 122 G-flags by the preprocessor. The DQE felt these values were realistic, even though they were up to 10 meters per second greater than climatology.

Sea Temperature (TS) was assessed 827 G-flags by the preprocessor. The DQE felt these values were realistic, although they were between 2 to 5 degrees Celsius lower than the given climate data.

Temperature (T) was assessed 673 G-flags by the preprocessor. The DQE felt these values were realistic, despite the fact the values were up to 10 degrees Celsius lower than the available climatology.

Relative Humidity (RH) was assessed one G-flag by the preprocessor. The DQE felt this value was realistic, although it was approximately 35% lower than climatology.

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).

K-flag:

During the 93-C cruise, all of the meteorological variables were assessed K-flags from 10:55 on 11 December 1993 to 14:30 on 12 December 1993 due to a possible instrument malfunction.

1994 FLAG SUMMARY

Statistical Information:

Details of the cruises 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 1,017,640 values were evaluated with 10,122 flags added by both the preprocessor and the DQE resulting in 0.99% of the values being flagged.

Table 3: Statistical Cruise Information

Cruise Identifier	Cruise Dates	Number of Records	Number of Values	Number of Flags	Percent Flagged
94-A	02/11/94 - 02/28/94	2,567	48,773	1	0.00*
94-B	04/01/94 - 05/19/94	6,576	124,944	1,636	1.31
94-C	05/31/94 - 06/17/94	2,461	46,759	248	0.53
94-D	07/05/94 - 10/06/94	13,373	254,087	674	0.27
94-E	11/23/94 – 06/10/95	28,583	543,077	7,563	1.39

Summary:

The 1994 data from the *Polarstern* proves to be of excellent quality with 0.99 % of the reported values being flagged for potential problems. The distribution of flags for the remaining variables is detailed in Table 4.

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

Variable	B	F	G	I	K	S	Total Number of Flags	Percentage of Variable Flagged
LAT		4					4	0.01
LON		4					4	0.01
PL_HD							0	0.00
PL_CRS							0	0.00
PL_SPD						6	6	0.01
PL_WDIR							0	0.00
PL_WSPD						1	1	0.00*
DIR						1	1	0.00*
SPD					28	1	29	0.05
DIR2						1	1	0.00*
SPD2			1,659		28		1,687	3.15
TS			3,320				3,320	6.20
P	282		2,177	9			2,468	4.61
T			2,114				2,114	3.95
TD					66	1	67	0.13
RH			243		109	1	353	0.66
Q					66	1	67	0.13
RRATE							0	0.00
Total Number Of Flags	282	8	9,513	9	297	13	10,122	
Percent Of All Values Flagged	0.03	0.00*	0.93	0.00*	0.03	0.00*	0.99	

*Percentages < 0.01

B-Flag:

Pressure (P) received 282 B-flags from the preprocessor during the 94-E cruise. During three separate occurrences in the middle of the cruise, the pressure dropped below 950mb. **See I-flag.*

F-Flag:

Latitude and Longitude were both assessed four F-flags by the preprocessor during the 94-D and 94-E cruises. These flags show that the platform speed computed by the preprocessor exceeds a realistic speed for a research vessel (15 meters per second). This may have been caused by uncertainties or truncation error in the navigation data.

G-Flag:

In 1994, during the five cruises, five different variables received G-flags. Of the 9,513 values that were assessed G-flags, 7,206 of them occurred during the 94-E cruise, 1,635 occurred on the 94-B cruise and on the 94-D cruise, 623 values were assigned G-flags. The 94-B and 94-E cruises traveled through the South Atlantic, mainly collecting data around Antarctica. Cruise 94-D traveled back and forth through the Greenland and Norwegian Seas before returning to Germany. Climatological data is very limited in these high latitude regions, resulting in more uncertainty in the climatology.

Ocean Relative Wind Speed (SPD2) was assessed 1,659 G-flags by the preprocessor. The DQE felt these values were realistic, even though they were up to 15 meters per second greater than climatology.

Sea Temperature (TS) was assessed 3,320 G-flags by the preprocessor. The DQE felt these values were realistic, despite the fact they were approximately five degrees Celsius lower and/or greater than the given climate data.

Pressure (P) was assessed 2,177 G-flags by the preprocessor. The DQE felt these values were realistic measurements within areas of deep polar lows. At times pressure values were upwards of 50 mb less than the available climate data.

Temperature (T) was assessed 2,114 G-flags by the preprocessor. The DQE felt these values were realistic, although they were approximately 10 degrees Celsius lower and/or greater than climatology.

Relative Humidity (RH) was assessed 243 G-flags by the preprocessor. The DQE felt some of these values probably not realistic, as they were up to 50% lower than the given climatological data. Relative Humidity sensors are not ideal in high latitude locations. When temperatures are below freezing, sublimation takes place. Days when changes occur in the pressure and wind data at the same time of relative humidity drops could show realistic values.

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).

I-flag:

Pressure (P) was assessed nine I-flags by the DQE during the 94-E cruise. At the time of all I-flags, the *Polarstern* was traveling through the Bellingshausen Sea, off the Western Antarctic coast. The I-flags were put in place to highlight instances of when the Polarstern encountered deep polar low-pressure centers. When the Polarstern was off the coast of Antarctica, a pressure of 949.1 was recorded on 11 April 1995 at 02:05. In addition, pressures of 945.3 and 939.5 were record on 16 April 1995 at 12:45 and on 19 April 1995 at 12:45, respectively. These values can be found during the same occurrences of B-flags with the pressures below 950mb.

K-flag:

During the 94-C cruise, Relative Humidity, Dewpoint Temperature (TD) and Specific Humidity (Q) were K-flagged when winds were high enough to possibly cause a sea-spray effect, interfering with the sensors.

Relative Humidity received K-flags during the 94-D cruise, when the relative humidity increased at the same time that the temperature increased.

K-flags were used to reveal signatures of ship motion in certain variables. Variables such as Earth Relative Wind Speed (SPD) and Ocean Relative Wind Speed (SPD2) 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 or ocean relative data. Subsequently, the data was flagged as suspect.

1995 FLAG SUMMARY

Statistical Information:

Details of the cruises 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 662,530 values were evaluated with 2,060 flags added by both the preprocessor and the DQE resulting in 0.31% of the values being flagged.

Table 5: Statistical Cruise Information

Cruise Identifier	Cruise Dates	Number of Records	Number of Values	Number of Flags	Percent Flagged
95-A	07/07/95 - 10/29/95	16,432	312,208	1,313	0.42
95-B	11/09/95 - 03/08/96	18,438	350,322	747	0.21

Summary:

The 1995 data from the *Polarstern* proves to be of excellent quality with 0.31 % of the reported values being flagged for potential problems. The distribution of flags for the remaining variables is detailed in Table 6.

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

Variable	B	G	I	K	S	Total Number of Flags	Percentage of Variable Flagged
LAT						0	0.00
LON						0	0.00
PL_HD					1	1	0.00*
PL_CRS						0	0.00
PL_SPD	1					1	0.00*
PL_WDIR						0	0.00
PL_WSPD			1			1	0.00*
DIR				174		174	0.50
SPD			1			1	0.00*
DIR2				274		274	0.79
SPD2		97	1			98	0.28
TS		1,089		19		1,108	3.18
P		31	1		2	34	0.10
T		218				218	0.63
TD					2	2	0.01
RH		144			2	146	0.42
Q					2	2	0.01
RRATE						0	0.00
Total Number Of Flags	1	1,579	4	467	9	2,060	
Percent Of All Values Flagged	0.00*	0.24	0.00*	0.07	0.00*	0.31	

*Percentages < 0.01

<u>*B-flag:*</u>

Platform Speed (PL_SPD) was assessed one B-flag by the preprocessor during the 95-A cruise. The Platform Speed jumped from 4.5 m/s to 15.2 m/s within 10 minutes.

<u>*G-Flag:*</u>

The two cruises in 1995, 95-A and 95-B, covered both the Arctic and the Antarctic Circles. Due to both of the cruises taking place in such high latitudes, climatological data maybe questionable due to the lack of available data for both regions.

Ocean Relative Wind Speed (SPD2) was assessed 97 G-flags by the preprocessor. The DQE felt these values were realistic, even though they were approximately 10 meters per second greater than the climate data.

Sea Temperature (TS) was assessed 1,089 G-flags by the preprocessor. The DQE felt these values were realistic, despite the fact they ranged from 5 degrees Celsius greater to 5 degrees Celsius lower than climatology.

Pressure (P) was assessed 31 G-flags by the preprocessor. The DQE felt these values were realistic, although they ranged from 20 mb greater to 12 mb lower than the available climatology.

Temperature (T) was assessed 218 G-flags by the preprocessor. The DQE felt these values were realistic, even though they were approximately 5 degrees Celsius lower and/or greater than the climate data.

Relative Humidity (RH) was assessed 144 G-flag by the preprocessor. The DQE felt some of these values might not be realistic, as they were up to 50% lower than the given climatological data. Relative Humidity sensors are not ideal for explorations into regions of high latitudes.

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).

I-Flag:

Platform Wind Speed (PL_WSPD), Ocean Relative Wind Speed (SPD2), Earth Relative Wind Speed (SPD) and Pressure (P) were each assessed one I-flag by the DQE during the 95-B cruise. On 27 February 1996 at 03:35, platform wind speed was 40.1 m/s, and 38.1 m/s was recorded as the ocean relative wind speed. An extremely high earth relative wind speed of 40.2 m/s was reported at 03:45, with a special value being recorded ten minutes prior. On 27 February 1996 at 03:55, a low pressure of 964.9mb was recorded. The I-flags were put in place to highlight these interesting features.

K-flags:

Sea Temperature (TS) was assessed K-flags during the 95-B cruise for a possible instrument malfunction. From 24-26 January 1996, the sea temperature dropped of rapidly before reporting missing values and then returns to the previous trend once the instrument resumes recording data.

Earth Relative Wind Direction (DIR) and Ocean Relative Wind Direction (DIR2) were both assessed K-flags during the 95-A cruise. From 21:05 on October 19th to 21:05 on October 21st, the wind direction data seems to suspect. After consulting other DQE's on a possible problem, it was not determined what caused this irregularity in the data. This data should be used with caution.

1996 FLAG SUMMARY

Statistical Information:

Details of the cruises 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 376,143 values were evaluated with 508 flags added by both the preprocessor and the DQE resulting in 0.14% of the values being flagged

Table 7: Statistical Cruise Information

Cruise Identifier	Cruise Dates	Number of Records	Number of Values	Number of Flags	Percent Flagged
96-A	05/21/96 - 06/20/96	4,422	84,018	2	0.00*
96-B	07/12/96 - 09/22/96	10,479	199,101	289	0.15
96-C	10/05/96 - 11/08/96	4,896	93,024	217	0.23

Summary:

The 1996 data from the *Polarstern* proves to be of excellent quality with 0.14 % 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	B	F	G	I	S	Total Number of Flags	Percentage of Variable Flagged
LAT		4				4	0.02
LON		4				4	0.02
PL_HD						0	0.00
PL_CRS						0	0.00
PL_SPD						0	0.00
PL_WDIR						0	0.00
PL_WSPD						0	0.00
DIR						0	0.00
SPD						0	0.00
DIR2						0	0.00
SPD2						0	0.00
TS			257			257	1.30
P				2		2	0.01
T			26			26	0.13
TD					3	3	0.02
RH					3	3	0.02
Q					3	3	0.02
RRATE	206					206	1.04
Total Number Of Flags	206	8	283	2	9	508	
Percent Of All Values Flagged	0.05	0.00*	0.08	0.00*	0.00*	0.14	

*Percentages < 0.01

F-Flag:

Latitude and Longitude were both assessed four F-flags by the preprocessor during the 96-A and 96-B cruises. These flags show that the platform speed computed by the preprocessor exceeds a realistic speed for a research vessel (15 meters per second). This may have been caused by uncertainties or truncation error in the navigation data.

G-Flag:

All of the G-flags occurred during the 96-B cruise. The ship was entering the Kara Sea from the Laptev Sea, along the northern part of Russia, when the G-flags appeared.

Sea Temperature (TS) was assessed 257 G-flags by the preprocessor. The DQE felt these values were realistic, as they were only about 2 degrees Celsius lower than the climatological data.

Temperature (T) was assessed 26 G-flags by the preprocessor. The DQE felt these values were realistic, although they were approximately 4 degrees Celsius lower than the trend of the available climate data.

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).

I-flags:

Pressure was assessed two I-flags during the 96-C cruise. On 08 October 1996 at 22:15, a pressure of 970.6mb was recorded. The I-flag were put in place to highlight these interesting low-pressure values.

Final Comments:

The overall quality of the data was excellent with the exception of the missing data in some of the wind variables. This data set proves to be good for scientific studies.

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.