Malcom Baldrige AWS Data Quality Control Report

Jesse Enloe and Shawn R. Smith

World Ocean Circulation Experiment (WOCE)

Surface Meteorological Data Assembly Center

Center for Ocean-Atmospheric Prediction Studies

Florida State University

June 16, 1999

Report WOCEMET 99-08

Version 1.0

Introduction:

This report summarizes the quality of surface meteorological data collected by the research vessel *Malcom Baldrige* (identifier: WTER) automated weather system (AWS) during three WOCE cruises beginning 21 March 1995 and ending 25 October 1995. The data were provided to the Florida State University Data Assembly Center (DAC) in electronic format by D. Shields at NOAA Corps. Conversion to the FSU internal format required data averaging and time matching detailed in the Appendix. The data were then processed using an automated screening program, which adds 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 judgement of the DQE and other DAC personnel. Details of the WOCE quality control procedures can be found in Smith et al. (1996). The data quality control report summarizes the flags for the *Malcom* AWS surface meteorological data, including those added by both the preprocessor and the DQE.

Statistical Information:

The *Malcom Baldrige* AWS data are expected to include observations taken every minute on all three WOCE cruises. Values for the following variables were collected:

Time	TIME
Latitude (GPS)	LAT
Longitude (GPS)	LON
Latitude (Trim)	LAT2
Longitude (Trim)	LON2
Platform Heading (ashtk)	PL_HD
Platform Heading (gyrocompass)	PL_HD2
Platform Course (GPS)	PL_CRS
Platform Speed (GPS)	PL_SPD
Platform Speed (EDO Log)	PL_SPD2
Platform Relative Wind Direction (R.M. Young)	PL_WDIR
Platform Relative Wind Speed (R.M. Young)	PL_WSPD
Platform Relative Wind Direction (Bendix)	PL_WDIR2
Platform Relative Wind Speed (Bendix)	PL_WSPD2
Earth Relative Wind Direction (R.M. Young)	DIR
Earth Relative Wind Speed (R.M. Young)	SPD
Earth Relative Wind Direction (Bendix)	DIR2
Earth Relative Wind Speed (Bendix)	SPD2
Sea Temperature (thermosalinograph)	TS
Sea Temperature (seachest)	TS2
Sea Temperature (WAHL)	TS3
Atmospheric Pressure	P
Air Temperature	T
Relative Humidity	RH

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 2,285,300 values were evaluated with 108,823 flags added by the preprocessor and the DQE for a total of 4.76% of the values being flagged. Note that the four earth relative wind parameters were found to be of poor quality and are not included in the statistical results. These values will not be released with the data for this cruise (discussion below).

Table 1: Statistical Cruise Information

CTC	Dates	Number of Records	Number of Values	Number of Flags	Percentage Flagged
IR_03_/01; ISS02_/05;ISS01_/09	03/21/95 – 04/22/95	26,166	523,320	8,591	1.64
IR_01W/03; IR_04_/04	05/31/95 – 06/30/95	40,977	819,540	43,164	5.27
IR_04_/06	09/22/95 – 10/25/95	47,122	942,440	57,068	6.06

Summary:
Though only 4.76% of the values were flagged, the overall quality of the meteorological data collected from the *Malcom Baldrige* proved to be poor. After reevaluation of the four calculated true wind variables it was determined that they should not be included in the public release. Table 2 details the distribution of flags among the remaining variables.

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

Variable	B	F	G	I	J	K	L	S	Total Number of Flags	Percentage of Variable Flagged
TIME									0	0.00
LAT		21					69		90	0.08
LON		21					69	1	91	0.08
LAT2									0	0.00
LON2									0	0.00
PL_HD						17		325	342	0.30
PL_HD2									0	0.00
PL_CRS								3	3	0.00*
PL_SPD								12	12	0.01
PL_SPD2								3	3	0.00*
PL_WDIR						17,489		16	17,505	15.32
PL_WSPD						17,489		1	17,490	15.31
PL_WDIR2					13,581	7			13,588	11.89
PL_WSPD2					13,581	7		2	13,590	11.89
TS						92		14	106	0.09
TS2			1416			113		8	1,537	1.35
TS3			1781			98			1,879	1.64
P				2				13	15	0.01
T				2		25,514		22	25,538	22.35
RH	2				42	16,978		12	17,034	14.91
Total Number of Flags	2	42	3197	4	27,204	77,804	138	481	108,823	
Percentage of All Values Flagged	0.00*	0.00*	0.14	0.00*	1.19	3.40	0.01	0.02	4.76	

*Percentage < 0.01

Deleted Data:

True Winds:
The *Malcom Baldrige* provided two sets of calculated true wind values, one set calculated using wind data from the R.M. Young wind vane and the other from the Bendix wind vane. The quality of these winds proved to be less than mediocre, demonstrating such problems as acceleration spikes and echoes of ship movement. Correct true winds show no signal of ship motion. The true winds were recalculated by DAC personnel using the data provided by the *Malcom Baldrige*, employing a tested code. A detailed description of the true wind calculation procedures can be found in Smith et al. (1999). The DAC found the same problems in the recalculated winds as in the provided wind data. This would suggest a problem with the input data, possibly a units conversion problem. At present, the true winds data will not be released, as the problem is still being investigated, but there is potential for a future release date upon resolution of the problem.

Major Problems:

Temperature and Relative Humidity:
The quality of the air temperature (T) and the relative humidity (RH) data are questionable at best. Viewing the data as a time series using the Visual Data Assessment Tool (VIDAT) shows evidence of ship motion induced changes in both of the variables. The obvious anomalies caused by ship motion received the cautionary flag (K), the spike flag (S), and, in extreme cases where the data should be completely disregarded, the J flag. Some of the other anomalies were left unflagged because of insufficient meteorological backing due to the lack of trust in the true winds data. In these cases, there was not enough evidence to say whether the data should be flagged. The DQE attempted to relate the problem to platform relative wind, ship speed, and ventilation problems but a specific cause was not determined. It is possible that data left unflagged are also questionable and should be used with caution.

Platform Relative Winds:
On the last cruise, beginning 25 September the first set of platform relative winds (PL_WDIR and PL_WSPD) became highly questionable. Both PL_WDIR and PL_WSPD were stuck at zero, sporadically jumping to various values and returning immediately to zero. These data values were highly suspect and appropriately flagged with the K flag. The problem, likely an instrumental error, was fixed on 8 October.

The Bendix platform relative winds (PL_WDIR2 and PL_WSPD2) from the IR_01W/03 cruise were missing from 10 June through 17 June. When the winds started reporting again, on the eighteenth, the data were extremely noisy and very unrealistic. The data remained as such for the rest of the cruise and were all flagged unusable (J) by the DQE. These data were missing all together in the last cruise.

Other Problems:

GPS Latitude and Longitude:
The GPS Latitude and Longitude (LAT and LON) proved not to be as accurate as the Trimble Latitude and Longitude (LAT2 and LON2), as they demonstrated a higher amount of variability. While still close to port at the beginnings of the second and fourth cruises, the ship was too close to land for its position to be resolved in the land mask used by the preprocessor; thus, LAT and LON were assessed L flags. The L flag is to bring attention to a position value over land. In other areas, the preprocessor applied F flags to LAT and LON, indicating unrealistic platform velocity as determined by position data. Typically, the DAC only releases one set of navigation data. In this case we chose the Trimble Latitude and Longitude position data for release to the public, because of its slightly higher data quality. If for a particular minute, the Trimble position data was missing and the GPS existed, the GPS value stood proxy for the position data for that minute (25 March 1995 at 04:52 and on 06 October 1995 from 17:29-23:58). The GPS data is available upon request. For public release, the Trimble Latitude and Longitude were renamed LAT and LON respectively.

Platform Heading:
The ashtk platform heading (PL_HD) was considerably noisy as compared to the gyrocompass platform heading (PL_HD2). There were a total of 362 spike flags (S) applied to PL_HD by the DQE for values radically different from the general trend. Seventeen K flags were also assessed to the data to indicate suspect data. If the user chooses to use PL_HD, a smoother is recommended.

Sea Temperature:
At the beginning of the second cruise, the vessel began recording data before its departure from port. The measurements of sea temperature were that of the water in the harbor and therefore drastically different from that of the temperatures at sea, once the ship departed. The sea temperature values taken at port were all flagged with the K flag.

On the vessel's fourth cruise, it traveled south of the 40-degree south latitude line. While travelling through this region of the globe the seachest and WAHL sea temperature recorded values greater than four standard deviations from the climatological mean (see da Silva et al. 1994). These values received the G flag from the preprocessor. This region, plagued with harsh conditions, is rarely traveled, and therefore the climatology is not well known. The data values flagged G indicate extreme temperatures, but are believed to be realistic and of good quality.

Atmospheric Pressure:
Atmospheric Pressure received 645 K flags on 15 July for values that echoed ship motion. Reasons for this occurring on only one day could not be determined.

Worthy of Mention:
A possible brown out took place on 14 July. All instruments except the sea temperature instruments experienced a radical change in the data for a period of about fifteen minutes.

All meteorological variables affected by this event where flagged as suspect (K). Due to limited power resources on a vessel, over taxation of power resources can cause errors in electronic measurements of meteorological variables.

A particular event took place in the morning hours of 1 June. A series of oscillations occurred in atmospheric pressure and air temperature. It is believed that the vessel traveled just behind, and parallel to, a passing front where it passed through a series of convective bands separated by areas of down drafts. The event is bounded on the temperature and pressure data at the beginning and end with an interesting feature flag (I).

Final Comments:

There are suspicious temperature and relative humidity data that were not flagged by the DQE. However, in many cases, there was little information to help decipher if these events were legitimate occurrences. The DQE strongly cautions use of these variables, as much of the data left unflagged may still be questionable.

If the user should choose to employ use of the PL_HD data, a smoother is suggested, due to its excessive noisiness.

Isolated spikes occurred in most of the variables throughout the data. Spikes are a relatively common occurrence in automated data, caused by such factors as electrical interference and ship accelerations. These individual points were assigned the S flag.

Though only 4.76% of the total values were flagged, the over all quality of the meteorological data is average.

Appendix:

The different variables' raw ship data were recorded and stored in separate files. There was a wide spectrum of time intervals amongst the data, ranging from 15 seconds to one minute. All ship data with time intervals less than one minute were converted to one minute averages in their respected files – scalar values (e.g., air temperature, sea temperature, and atmospheric pressure) linearly averaged and vector values (e.g., platform course, earth and ship relative winds) vector averaged.

The next step was time matching. The position data (i.e., the Trimble latitude and longitude) values were used as the time template to which the other data values were matched. Meaning, for every minute average a value existed for the position data, the corresponding variables (i.e., temperature, pressure, etc.) were matched to that minute. If both position data were missing for a particular minute, that minute was not included in the data set for any variable. In the case of a duplicate time stamp, the first of the two values was used and the second dismissed

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

Smith, S.R., M.A. Bourassa, and R.J. Sharp, 1999: Establishing More Truth in True Winds. *J. Atmos. Oceanic Technol*., in press.

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.