A. Von Humboldt Automated Weather System Data Quality Control Report

Cruises: AR_06_/01
AR_16_/16
AR_06_/02
AR_16_/04
AR_06_/03
AR_16_/05
AR_06_/04
AR_16_/12
AR_16_/07
AR_16_/11

Daniel M. Gilmore and Shawn Smith

World Ocean Circulation Experiment(WOCE)

Surface Meteorological Data Assembly Center

Center for Ocean Atmospheric Prediction Studies

The Florida State University

August 28, 1996

Report WOCEMET 97-2

Version 1.0

Introduction:

The data referenced in this report were collected from the research vessel A. von Humboldt (call sign: Y3CW; data provider: Institute for Baltic Research/P.I.: E. Hagen) Automated Weather System from 10 different cruises. The data were recieved in electronic format and converted to a standard FSU format. Then they were preprocessed using an automated data checking program. Next a visual inspection was completed by a Data Quality Evaluator(DQE) who reviewed, modified and added appropriate quality control (QC) flags to the data. Details of the WOCE QC can be found in Smith et al. (1996). The data quality control report summarizes the flags for the von Humboldt data, including those added by both the preprocessor and the analyst.

Table 1: List of dates and number of records for each cruise

CTC	Dates	Number of Records	Number of Values	Number of Flags	Percentage Flagged
AR_06_/01 AR_16_/16	08/26/91 - 08/29/91	4678	60814	3122	5.13
AR_06_/02 AR_16_/04	09/03/91 - 09/12/91	11064	143806	9838	6.84
AR_06_/03 AR_16_/05	09/17/91 - 10/15/91	27626	359138	23516	6.54
AR_16_/07	10/18/91 - 10/19/91	1402	18226	1211	6.14
AR_16_/11	09/06/92 - 09/09/92	4327	56806	2260	3.98
AR_06_/04 AR_16_/12	09/10/92 - 09/25/92	18807	244491	10322	4.22

Statistical Information:

The data from the von Humboldt were expected to include observations taken every minute from 10 cruises. The start and end dates, the number of observations, and the number and percentage of non-Z flags for each cruise is given in table 2. Time (TIME), latitude (LAT), longitude (LON), platform course (PL_CRS), platform speed (PL_SPD), earth relative wind direction (DIR), earth relative wind speed (SPD), sea temperature (TS), atmospheric pressure (P), air temperature (T), wet-bulb temperature (TW), relative humidity (RH), and atmospheric radiation (RAD) were analyzed. A total of 883,281 values were checked with 50,269 flags added resulting in 5.69 of the data being flagged. The distribution of flags for each variable sorted by flag type is detailed in table 2.

Definition of Flags

B: Data out of bounds
F: Unreal ship movement in lat and lon
G: Data 4 standard deviations from climatological mean
H: Discontinuity in data
I: Interesting data point
J: Bad data point
K: Caution/Suspect data
L: Land error
S: Spike in data
T: Time error or duplicate
Z: Data passing all evaluation

Table 2: Percentage of Flags Assigned by Flag Type and Variable

Variable	B	F	G	H	I	J	K	L	S	T	Total Number of Flags	Percentage of Data Flagged
TIME										2	2	0.00
LAT		796		2			92	100	63		1053	1.44
LON		778						100	1597		2475	3.39
PL_CRS						108			1		109	0.15
PL_SPD						207					207	0.28
DIR							953		77		1030	1.41
SPD			2		1	49	935				987	1.35
TS			16	2		287	544		1		850	1.17
P			18				140				158	0.22
T			577		2	101					680	0.93
TW					3		108				111	0.15
RH			1			99					100	0.14
RAD	42444										42444	58.27
Total Number of Flags	42444	1574	614	4	6	851	2772	200	1739	2	50206	5.30
Percentage of Flags Used	4.48	0.17	0.06	0.00	0.00	0.09	0.30	0.02	0.18	0.00	5.30	

Summary:

A: Significant Problems:

The most obvious problem with this data is that 58.27 percent of the RAD data was flagged by the prescreener as “B”, out of bounds. If the sensor is measuring only incoming radiation, the sensor may be

off calibration; however, if the sensor is measuring net radiation, the values may be accurate. The RAD observations were sometimes down to $-100 W/m^2$. This problem may have carried over into the radiation observations taken during the day. At the time of QC, the RAD instrument type was unknown and no confirmation was available at present. Thus, all the RAD data flagged as being out of bounds should be treated with caution.

Another problem with this data set is the nearly 800 pairs of "F", unreal movement" flags applied by the prescreener to the LAT and LON data. No reason for these observations is known to the DQE, however, slight jumps in measured ship position can cause this problem in minute resolution data. The flags were left as a cautionary note.

In addition to 776 "F" flags, 1597 "S", spike in data, flags were added to LON. These spikes were a result of some points in the original data having degrees longitude, but not minutes or seconds longitude. Thus the data would suddenly jump to the next whole longitude for 1 observation and then return to the pattern of correct

longitude observations. Any jump in the data such as the one described would result in "F" flags being applied to both the latitude and the longitude for the points next to where the jump occurred.

The DQE removed these "F" flags when the "S" flags were added.

The last major problem with this data is that DIR and SPD occasionally mirrored the movement of the ship. That is, the DIR sometimes reflected the changes in PL_CRS, and SPD the changes in PL_SPD. Any time DIR or SPD showed patterns similar to PL_CRS or PL_SPD respectively, DIR and SPD were flagged with "K", caution/suspect data, flags. The result was that 953 values of the DIR data and 935 values of the SPD data were flagged with "K" flags.

B: Other cautionary flags:

- 2 "H" flags added to each LAT and T. These each have one discontinuity, and these flags reflect that. The data immediately following the discontinuity are very questionable, so these values were flagged with "K" flags.

- 577 "G" flags were added to T for various days and in varying amounts over the 2 week period from 09/06/91 - 09/18/91.

These were all temperatures that were far above the climatological mean. A nominal number of "G" flags were also added to SPD, TS, P, and RH.

- Varying numbers of "J" flags were added to PL_CRS, PL_SPD, SPD, TS, T, and RH. These were added for the most part because the data seemed to flat-line (i.e., many observations in a row that were the exact same value.

  Identical values are rare in minute resolution data; thus the data were flagged as erroneous. Another situation that occurred was that PL_CRS was erroneous, but DIR seemed to be correct, and may have only been off slightly. Any DIR observation in the same record as an erroneous PL_CRS observation was flagged with "J" as well.

- 100 "L" flags were left by the DQE. These were all applied by the prescreener during the day of 10/15/91. The longitude data jump several degrees, and by doing so, the positions are shifted over land. These values may be considered erroneous.

- The data also had a small number of “S” flags applied when necessary. An explanation for the 1577 “S” flags added to LON is in part A of this summary. The rest of the “S” flags are a result of typical problems one has with a large or high temporal resolution data set.

- TIME had 2 “T” flags added by the prescreener because of time duplicates.

Final Note:

These data are in good condition. Other than the RAD data, there were no systematic or frequently recurring errors. The DQE cautions the user to regard all the RAD data as suspect. All the other data, with the exception of flagged data, is considered usable.

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 32310.