Quality Control of Automatic Weather Station Data

Mark W. Seefeldt University of Colorado - Boulder

Linda M. Keller University of Wisconsin - Madison

Outline

- Current AWS data processing
- Weaknesses in the current data processing
- New automated quality control process
- New manual quality control process
- New AWS data features
- New AWS data products

Current AWS Data Processing

089191205.r

1	0891	91205	5.r - Note	pad					
	File Edit	Forma	at View H	Help					
こう ういい はいない かいろいい	Dec Lat 335 335 335 335 335 335 335 335 335 33	05 78. 1 2 3 4 5 6 7 8 9 101 12 13 14 15 167 18	Statio 505 Lo -11.8 -11.2 -11.4 -10.8 -10.5 -10.4 -10.2 -10.2 -10.0 -9.6 -9.6 -9.6 -9.6 -9.6 -9.6 -9.5	n : 8919 982.3 982.5 982.5 982.5 982.7 982.7 982.7 982.7 982.7 982.7 982.7 982.8 982.9 982.9 982.9 982.9 982.9 982.9 982.9 982.9	<pre>emil 12E 10.4 6.9 8.2 8.4 8.2 8.4 8.2 8.4 8.4 8.4 8.4 8.4 8.4 8.4 8.4 8.9 6.2 5.7 7.4 8.2 6.7 9.2</pre>	ia Elev: 212.0 210.0 208.0 215.0 207.0 205.0 208.0 204.0 204.0 204.0 205.0 201.0 204.0 205.0 201.0 205.0	52 M 74.2 73.8 78.2 73.4 73.4 73.4 71.6 70.3 71.2 72.5 68.1 69.8 67.6 71.6 72.9 74.2 74.2	-1.9 -1.8 -2.0 -1.5 -1.4 -1.6 -1.9 -1.9 -1.9 -1.9 -1.9 -1.9 -1.9 -1.9	

- The 10-minute observations from ARGOS are processed to eliminate extreme outliers, and to format the data into .r 10-minute files
- This is done typically within a few weeks after the end of a month
- There are no quality control procedures performed on the observations, other than the extreme outliers

Current AWS Data Processing

9191205h.dat

11	B 91912	05h.dat -	Notepad						×
Ĩ.	File Edit	Format Vie	w Help						
「「「「「「「「「「」」」、「「」」、「「」」、「」、「」、「」、「」、「」、「	File Edit Dec C Lat : -11.8 -9.5 -9.1 -10.6 -12.6 -11.8 -10.2 -9.5 -8.2 -7.6 -7.6 -7.6 -7.6 -8.9 -10.2 -12.8 -12.8 -12.5	Format Vie 5 Sta 78.505 982.3 982.8 983.2 983.7 983.5 983.9 984.2 985.2 995	W Help Long : 10.4 9.9 5.2 6.2 6.2 4.7 8.4 8.4 9.4 8.4 9.4 8.7 7.4 7.2 7.7 4.2 7.7	8919 Em 173.12E 212.0 205.0 195.0 205.0 195.0 195.0 195.0 177.0 186.0 184.0 170.0 181.0 179.0 177.0 188.0 179.0	ilia Elev 74.2 74.7 77.3 85.2 87.9 87.9 87.9 87.9 87.0 91.8 90.5 83.5 69.4	: 52 M -1.9 -1.6 -1.0 -1.4 -2.1 -0.5 -1.1 -1.6 -1.7 -1.6 -1.7 -1.5 -0.4 -1.1 -1.2 -2.2 -1.9		i B	
	-9.8 -6.6	989.8 989.8	3.5 4.0	173.0 150.0	51.8 60.2	-1.2 -0.2			*
1	<							>	

- 3-hourly files are created and then scanned manually by Linda Keller
- If an observation looks suspect, in relation to the surrounding observations, the 10-minute observations are checked
- A bad observation is replaced with a surrounding 10-minute observation, if one exists within 40 minutes of the hour
- The process is extremely time consuming

AWS Data Processing - Dilemma

- The current method is very time consuming
- The expanding AWS network is requiring more AWS data to be processed every year
- Due to the subjective nature, it is best for one person to perform the quality control, if possible
- The quality controlled 3-hourly data is currently five years behind
- In order to include the AWS data in climatology archives a more expedient system is needed

AWS Data Processing - Solution

- Attempts to create an automated quality control method have proven to be difficult and not refined enough (good points are excluded, and bad points are included)
- Through the use of IDL software a semi-automated quality control process has been developed
 - An automated procedure removes data points based on a statistical evaluation of the observations
 - A manual procedure is performed to review and correct the automated quality control procedure
- This new semi-automated quality control procedure results in a faster quality control process, and additional products

AWS Data Quality Control - Automated

selected observation

•a window is defined around both sides of the observation (ie. +/- 24)

•the mean of the observations within the window is calculated

•the standard deviation of the observations within the window is calculated

•threshold values are established based on multipliers of the standard deviation (ie. +/- 3 SD)

•if the observation is outside of the threshold values it is removed

•the entire process is repeated for every observation





- Up to three different and progressive filters can be used
- The size of the window and the threshold multiplier can be custom specified for each sensor
- This is not able to applied to wind direction due to the more random variability and the discontinuity between 360° and 0°

AWS Data Quality Control – Automated Features



- The automated filter can be configured to remove all of the observations of a specific sensor (ie. bad wind direction at an AWS)
- A consecutive observation filter can be used to remove observations from a specific sensor which remain constant over a set value

AWS Data Quality Control – Manual

clicking on any observation point will
'flip' it from good to bad,or bad to good

•clicking and dragging a box will 'flip' all observations in the box

 controls are provided to adjust the day and data ranges

•all sensors, and nearby observations can be viewed in data table



table displays observation information from all the sensors for the point at the cursor

AWS Data Quality Control – Manual



- The manual quality control process is repeated for each sensor
- At the completion of all the sensors the end products are created

AWS Data – New Features

- New data files and formats have been introduced
- AWS sites are identified by a six-letter ID in the filenames: ex. Emilia: emilia Elizabeth: elzbth Gill: gill_____ Note: The official six-letter IDs have not been determined yet.
- Three-letter IDs are included in the data files to establish a common set of IDs for use in station plots, or other data displays.

ex. Emilia: EML Gill: GIL Carolyn: CRL

- The day is indicated as the day of month, and the observation as the time of the observation
- All data files will be given the .txt extension

New AWS Data Formats – 10-minute files (q10)

emilia200512q10.txt

2	🗾 en	nl20051	2q10.txt	- Notepad	Ĵ.					
Ċ.	File	Edit Forn	nat View	Help						
ころう ビー 液合 たいかく	γear: 111111111111111111111111111111111111	: 2005 78.50 0000 0010 0020 0030 0040 0050 0100 0110 0120 0140 0150 0200 0210 0220 0240	Month S Lon:: -11.8 -11.2 -11.4 -10.8 -10.4 -10.4 -10.2 -10.0 -10.0 -9.6 -9.6 -9.6 -9.6 -9.6	12 IC 173.12E 982.3 982.5 982.5 982.7 982.7 982.7 982.7 982.7 982.7 982.7 982.7 982.7 982.8 982.9 982.9 982.9 982.9 982.9	2: EML 10.4 6.9 8.2 8.4 8.2 8.4 8.2 8.4 8.2 8.4 8.4 8.4 8.9 7.9 6.2 5.7 7.4 8.7	ARGOS: : 52m 212.0 210.0 208.0 210.0 215.0 205.0 205.0 205.0 204.0 211.0 204.0 205.0 201.0 201.0 201.0 204.0 200.0 2	8919 74.2 73.8 78.2 73.4 73.4 73.4 73.4 71.6 70.3 71.2 72.5 68.1 69.8 67.6 71.9 74.2	Name: -1.9 -1.8 -2.0 -1.5 -1.4 -1.7 -1.9 -1.9 -1.8 -1.9 -1.8 -1.9 -1.9 -1.9 -1.9 -1.9 -1.9 -1.9 -1.9	Emilia	
	1	0250	-9.5	982.9	9.2	205.0	74.2	-1.7		~

- Quality controlled 10-minute data
- Includes all of the processed AWS data

New AWS Data Formats – 1-hour files (q1h)

emilia200512q1h.txt

2	🗾 en	nl20051	2q1 h. txt	- Notepad	j						×
G.	File	Edit Forn	nat View	Help							
2	∣γear	: 2005	Month	: 12 IC	: EML	ARGOS:	8919	Name:	Emilia		^
「「「「「「		0000 0100 0200 0300 0400 0500 0600 0700	-11.8 -10.4 -9.6 -9.5 -8.8 -8.9 -9.1 -9.5	982.3 982.7 982.9 982.8 982.9 983.0 983.2 983.3	10.4 8.4 6.2 9.9 6.9 7.7 5.2 6.9	212.0 207.0 204.0 205.0 205.0 191.0 195.0 197.0	74.2 73.4 69.8 74.7 71.6 72.9 77.3 83.0	-1.9 -1.7 -1.9 -1.6 -1.1 -1.2 -1.0 -1.0			
	1 1 1 1 1 1 1 1	0800 0900 1000 1200 1300 1400 1500 1600 1700	-10.0 -10.6 -10.8 -11.0 -12.6 -14.1 -12.5 -11.8 -11.2 -10.5	983.6 983.7 983.6 983.5 983.5 983.5 983.7 983.9 983.9 983.9 983.9	5.5 6.2 6.9 6.2 5.5 6.4 4.7 6.2 7.4	197.0 201.0 207.0 205.0 200.0 200.0 195.0 200.0 195.0 200.0	85.7 85.2 88.3 93.6 87.9 87.4 90.5 87.9 89.2 89.2 84.8	-1.1 -1.4 -0.9 -0.7 -2.1 -2.0 -0.3 -0.5 -0.9 -0.7			×
5	< .									>	11

- Quality controlled 1-hour data
- The observation, within 10 minutes of the hour (0, +10, -10), with the most reporting sensors (T,p,WS,WD) is selected for the hourly observation
- Any missing sensors are filled in with valid observations within 10 minutes of the hour

New AWS Data Formats – 3-hour files (q3h)

emilia200512q3h.txt

2	🗖 en	nl20051	2q3h.txt	- Notepad	j						×
2	File	Edit Forn	nat View	Help							
こうちょう 「「「「「「「「「」」」」」」」」	File Year 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Edit Forn : 2005 78.50 0300 0300 0900 1200 1200 1800 2100 0300 0000 0300 0400 0300 0400 0300 04000 0400 04000 0400 0400 0400 0400 0400 00	nat View Month S Lon: -11.8 -9.5 -9.1 -10.6 -12.6 -11.8 -10.2 -9.5 -8.2 -7.6 -7.6 -7.6 -12.8 -12.8 -12.5	Help : 12 IC 982.3 982.8 983.2 983.7 983.5 983.9 984.2 985.2 985.2 985.2 986.0 987.2 988.7 990.2 991.0 991.0 991.0	2: EML 10.4 9.9 6.2 6.2 4.7 8.4 9.4 8.4 9.4 8.7 7.4 7.7 4.2	ARGOS: 52m 212.0 205.0 195.0 201.0 205.0 197.0 197.0 197.0 197.0 197.0 197.0 197.0 197.0 198.0 197.0	8919 74.2 74.7 77.3 85.2 87.9 87.9 87.9 87.9 73.8 71.6 77.3 87.0 91.8 87.0 90.5 83.5 69.4	Name: -1.9 -1.6 -1.0 -1.4 -2.1 -0.5 -1.16 -1.7 -1.5 -0.4 -1.12 -2.2 -1.9	Emilia		
1	2 3 3	2100 0000 0300	-11.9 -9.8 -6.6	990.0 989.8 989.8	5.7 3.5 4.0	173.0 173.0 150.0	61.5 51.8 60.2	-1.3 -1.2 -0.2			~
	<									>	

- Quality controlled 3-hour data
- The observation, within 40 minutes of the hour (0,+10,-10,+20, -20,+30,-30,+40,-40), with the most reporting sensors (T,p,WS,WD) is selected for the hourly observation
- Any missing sensors are filled in with valid observations within 40 minutes of the hour

AWS Data Formats – 3-hour files (dat)

9191205h.dat

205h.dat - I	Notepad						
Format View	v Help						
05 Stat 78.505 982.3 982.8 983.2 983.7 983.5 983.5 984.2 985.2 990.2 990.2 990.2 990.0 990.2 990.0 991.0 990.0	ion : Long : 10.4 9.9 5.2 6.2 4.7 8.2 8.4 9.4 8.4 9.4 8.4 9.4 8.7 7.7 4.2 7.7 4.2 5.7 4.0	8919 Em 173.12E 212.0 205.0 195.0 205.0 195.0 195.0 177.0 186.0 177.0 181.0 179.0 177.0 188.0 179.0 177.0 188.0 179.0 173.0 173.0 150.0	ilia Elev 74.2 74.7 77.3 85.2 87.9 87.9 87.9 87.9 87.0 91.6 77.3 87.0 91.8 90.5 83.5 69.4 61.5 51.8 60.2	: 52 M -1.9 -1.6 -1.0 -1.4 -2.1 -0.5 -1.1 -1.6 -1.7 -1.5 -0.4 -1.7 -1.2 -2.2 -1.9 -1.3 -1.2 -0.2			
	205h.dat - 1 Format View 05 Stat 78.505 982.3 982.8 983.2 983.7 983.5 983.9 984.2 985.2 985.2 985.2 987.2 988.7 987.2 988.7 988.7 988.7 988.7 988.7 988.7 988.7 988.7 988.7 988.7 988.7 988.7 988.7 988.7 988.7 988.7 988.7 988.7 990.2 991.0 991.0 999.0 989.8 989.8	205h.dat Notepad Format View Help 05 Station: 78.505 78.505 Long: 78.505 982.3 10.4 982.8 9.9 983.2 5.2 983.7 6.2 983.9 4.7 984.2 8.2 985.2 8.4 986.0 8.4 986.0 8.4 986.7 8.7 987.2 9.4 988.7 8.7 990.2 7.4 991.0 7.2 991.0 4.2 991.0 5.7 989.8 3.5 989.8 3.5 989.8 4.0	205h.dat - Notepad Format View Help 05 Station: 8919 Em 78.505 Long: 173.12E 982.3 10.4 982.3 10.4 982.3 10.4 982.3 10.4 982.3 10.4 982.3 10.4 982.3 10.4 982.3 10.4 982.3 10.4 982.3 10.4 982.3 10.4 982.3 10.4 982.3 10.4 982.3 10.4 982.5 195.0 983.5 6.2 983.9 4.7 984.2 8.2 898.7 8.4 184.0 986.0 8.4 181.0 990.2 7.4 991.0 7.2 991.0 7.2 991.0 7.7 989.8 3.5 989.8 3.5 989.8 4.0 <	205h.dat - Notepad Format View Help 05 Station: 8919 Emilia 78.505 Long: 173.12E Elev 8982.3 10.4 212.0 74.2 982.8 9.9 205.0 74.7 983.2 5.2 195.0 77.3 983.7 6.2 201.0 85.2 983.9 4.7 195.0 87.9 984.2 8.2 177.0 83.0 985.2 8.4 186.0 73.8 986.0 8.4 184.0 71.6 987.2 9.4 170.0 77.3 988.7 8.7 181.0 87.0 990.2 7.4 179.0 91.8 991.0 7.2 177.0 90.5 991.0 7.7 188.0 83.5 991.0 7.7 173.0 61.5 989.8 3.5 173.0 51.8 989.8 3.5 1	205h. dat - Notepad Format View Help 05 Station: 8919 Emilia 78.505 Long: 173.12E Elev: 52 982.3 10.4 212.0 74.2 -1.9 982.8 9.9 205.0 74.7 -1.6 983.2 5.2 195.0 77.3 -1.0 983.5 6.2 201.0 85.2 -1.4 983.5 6.2 205.0 87.9 -2.1 983.9 4.7 195.0 87.9 -0.5 984.2 8.2 177.0 83.0 -1.1 985.2 8.4 186.0 73.8 -1.6 986.0 8.4 184.0 71.6 -1.7 987.2 9.4 170.0 77.3 -1.5 988.7 8.7 181.0 87.0 -0.4 990.2 7.4 179.0 91.8 -1.1 991.0 7.2 177.0 80.5	205h.dat - Notepad Format View Help 05 Station: 8919 Emilia 78.505 Long: 173.12E Elev: 52 M 8 982.3 10.4 212.0 74.2 -1.9 982.8 9.9 205.0 74.7 -1.6 983.2 5.2 195.0 77.3 -1.0 983.7 6.2 201.0 85.2 -1.4 983.5 6.2 205.0 87.9 -2.1 8 983.9 4.7 195.0 87.9 -0.5 984.2 8.2 177.0 83.0 -1.1 985.2 8.4 186.0 73.8 -1.6 986.0 8.4 184.0 71.6 -1.7 987.2 9.4 170.0 77.3 -1.5 988.7 8.7 181.0 87.0 -0.4 990.2 7.4 179.0 91.8 -1.1 991.0 7.2 1	205h.dat - Notepad Format View Help 05 Station : 8919 Emilia 78.505 Long : 173.12E Elev : 52 M 8 982.3 10.4 212.0 74.2 -1.9 982.8 9.9 205.0 74.7 -1.6 983.2 5.2 195.0 77.3 -1.0 983.5 6.2 201.0 85.2 -1.4 983.5 6.2 205.0 87.9 -2.1 8 983.9 4.7 195.0 87.9 -0.5 984.2 8.2 177.0 83.0 -1.1 985.2 8.4 186.0 73.8 -1.6 986.0 8.4 184.0 71.6 -1.7 987.2 9.4 170.0 77.3 -1.5 988.7 8.7 181.0 87.0 -0.4 990.2 7.4 179.0 91.8 -1.1 991.0 7.2 177.0 80.5 -2.2 991.0 7.7 188.0 83.5 -2.2

- Quality controlled 3-hour data same data as the q3h files
- The file naming and data format is the exact same as the current 3-hourly files
- This product will be discontinued in the future and should no longer be used

New AWS Data Product – Monthly Meteograms



• Plot of temperature, pressure, wind speed, and wind direction, from the quality controlled 10-minute observations, for a month

Future AWS Data Product – Detailed 1h/3h data

🖪 en	nl200510	.q1h - Notepa	d							-	
File Year Lat: 10 10 10 10 10 10 10 10 10 10 10 10 10	Edit Forma : 2005 78.505 1 1 1 1 1 1 1 1 1 1 1 1 1	t View Help Month: 10 Lon: 172. 0000 0100 0200 0300 0400 0500 0600 0700 0800 0900 1000 1200 1300	ID: EML 12E Elev -23.2 -25.0 -26.1 -27.6 -29.5 -30.2 -31.5 -31.1 -32.8 -33.6 -34.5 -35.4 -34.0 -34.6	ARGOS: 52m 0000 0100 0200 0400 0500 0600 0700 0800 0900 1000 1100 1200	8919 N 974.2 974.7 975.2 975.3 975.2 975.2 975.2 975.2 975.2 975.2 974.4 974.5 974.3 974.3 974.3 974.3 974.3	Jame: 6 0000 0200 0300 0400 0500 0700 0700 0700 0700 0700 07	Emilia 9.4 7.7 4.7 1.2 2.7 2.2 3.2 444.0 444.0 444.0 444.0 444.0 1.7 2.0 2.0	0000 0100 0200 0400 0500 0600 4444 4444 4444 4444 44	240.0 226.0 231.0 266.0 276.0 276.0 233.0 444.0 444.0 444.0 444.0 241.0 285.0 271.0 269.0	0000 0100 0200 0300 0400 0500 0600 4444 4444 4444 4444 44	
10 10 10 10	1 1 1	1400 1500 1600 1700	-34.2 -32.4 -28.6 -31.4	1400 1500 1600 1650	972.5 971.4 971.3 970.3	1400 1500 1600 1650	2.0 4.0 3.7 1.7	1400 1500 1600 1650	288.0 298.0 298.0 330.0	1400 1500 1600 1650	
<											2

- In the future, a FORTRAN program will be provided which creates an output file which includes the time of the observation from each sensor
- The input files will be the q10 file and either the q1h or q3h files

Summary

- A new semi-automated data quality control system has been establish to process AWS data
- The new system provides easy creation of new data products (10-minute, 1-hour, 3-hourly, meteograms)
- Additional changes will be implemented with the new data products (use of six-letter IDs, new filenameing, etc.)
- The new system is not quite operational, an announcement will be made when the new system and products are implemented

Questions

- Mark Seefeldt
- Linda Keller

mark.seefeldt@colorado.edu Imkeller@wisc.edu