What the RMN QC package does (& doesn't do!)

The automated QC scripts generate QC reports for temperature, pressure and sensor depth which we will describe in detail in ensuing slides.

They do **NOT** -

- Automatically 'fix' the data. You still have to go through the process of interpreting the QC reports and -
 - Checking flagged data
 - Doing visual checks on the time series plots
 - Deciding if and how to make corrections
- Catch everything. There may be instances where the tests miss data points that should be flagged (we would love to get your feedback when this happens! we are still learning ourselves & want to keep improving the scripts).

Tips

Start by skimming through the QC Word document.

- Are there problems? If so, what kind, and what parameters do they affect?
- Then open the .csv file.
- Reformat the file so that it is easier to navigate (unfortunately we can't get .csv files to retain this formatting each time you open them)
 - Format Autofit Column Width
 - Alignment center the column entries
 - Freeze top row
 - Filter
- Filter for the flagged values. Evaluate whether changes should be made.
- Document changes (we are currently considering a couple different options for this; one is a field/data processing form – see handout – that gets submitted to the continuous data coordinator after each download).
- Records of any changes will also be visible in the QC .csv file, in the Comments field(s)

QC tests

- Unrealistic values ('Gross range')
 - Entries are flagged if values are above or below upper and lower limits
- Spikes
 - Entries are flagged if adjacent points change by more than 'x' amount
- Rate of change (RoC)
 - Entries are flagged if the RoC exceeds a given threshold (e.g., ≥ 3 st dev within 25 hrs)
- Flat line
 - Entries are flagged if a certain # of consecutive measurements are within a certain amount of each other (e.g., >10 consecutive temperature measurements are within 0.01 degrees C of one another)

Flags

Flags are assigned to each data point (a single measured parameter at a unique point in time)

- **P** = Pass,
- **S** = Suspect,
- **F** = Fail,
- X = No Data or Not Applicable (NA).

Fail - extreme, unrealistic

• Example - water temp values flagged 'F' if ≥ 30 or ≤ -2

Suspect - highly unlikely, but not as extreme as above

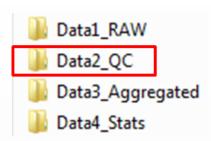
• Example - water temp values flagged 'S' if ≥ 25 or ≤ -1

Thresholds

Table 1. Updated thresholds (7/31/2017) based on evaluation of continuous air temperature, water temperature, air pressure, water pressure and sensor depth data that have been collected at freshwater, high quality wadeable RMN stream sites in the eastern US.

Most of the data are being recorded at 30-minute intervals by Onset HOBO U20 Water Level Data Loggers.

Test	Water Temp		Air Temp		Sensor depth		Water Pressure		Air (Barometric) Pressure	
	Fail	Suspect	Fail	Suspect	Fail	Suspect	Fail	Suspect	Fail	Suspect
Gross	>30°C or <- 2°C	>25°C or <- 0.1°C	>38°C or <- 25°C	>35°C or <- 23°C	>6 ft or <-1 ft	>6 ft or <-1 ft	>17 psi or <13 psi	>16.8 psi or <13.5 psi	>15 psi or <13 psi	>14.8 psi or <13.5 psi
Spike	≥1.5°C (±)	≥1.0°C (±)	≥10°C (±)	≥8°C (±)	≥5 <u>ft</u> (±)	≥3 <u>ft</u> (±)	≥0.7 psi (±)	≥0.5 psi (±)	≥0.25 psi (±)	≥0.15 psi (±)
Rate of Change	NA	≥3 <u>stdev</u> within 25 <u>hrs</u>	NA	≥3 <u>stdev</u> within 25 <u>hrs</u>	NA	≥3 <u>stdev</u> within 25 <u>hrs</u>	NA	≥3 <u>stdev</u> within 25 <u>hrs</u>	NA	≥3 <u>stdev</u> within 25 <u>hrs</u>
Flat Line	>30 consecutive measureme nts within 0.01 units of one another	>20 consecutive measureme nts within 0.01 units of one another	>15 consecutive measureme nts within 0.01 units of one another	>10 consecutive measureme nts within 0.01 units of one another	>60 consecutive measureme nts within 0 units of one another	>20 consecutive measureme nts within 0 units of one another	>15 consecutive measureme nts within 0.001 units of one another	>10 consecutive measureme nts within 0.001 units of one another	>15 consecutive measureme nts within 0.001 units of one another	>10 consecutive measureme nts within 0.001 units of one another



QC reports

2 documents

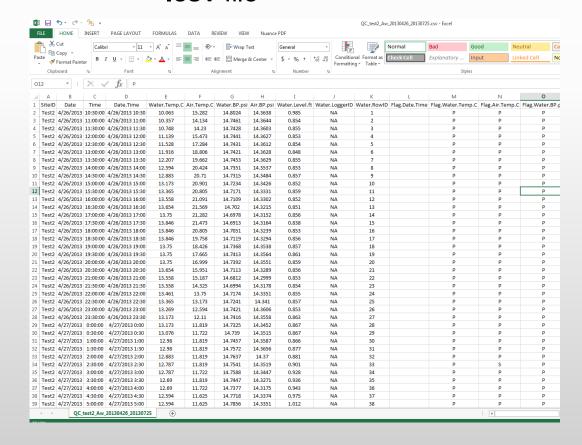
QC report (Word document)

Data Quality Control Report Report Date: 2016-02-05 DATA FILE INFORMATION Filename: QC_test1_Aw_20130725_20131015.csv SiteID: test1 Period of Record, Requested: 2013-07-25 to 2013-10-15 Period of Record, Actual: 2013-07-25 to 2013-10-15 Recording Interval: 30 minutes Data Type: Aw Parameters Included: Water.Temp.C, Air.Temp.C, Water.BP.psi, Air.BP.psi, Water.Level.ft, DATA SUMMARY, OVERALL Month and Day versus number of records. 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 10 48 48 48 48 48 48 48 48 48 48 48 48 49 48 21

16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

0 0 0 0 0 0 0 0 0 27 48 48 48 48 48 48

.csv file



QC Word file

QC report (Word)

4 main components

- 1. Data file information (siteID, date range, parameters, recording interval)
- 'Count' tables (# measurements/day) so that you can find inconsistencies (too few measurements (=missing data) or too many measurements)
- 3. Results from QC tests # of entries marked as pass (P), suspect (S), fail (F), missing data (X) or not available (NA)
- **4. Time series plots** for each individual parameter & several combined parameters (water & air temperature; water temperature and water level)

QC report (Word)

Organizational scheme

1. Data file information (siteID, date range, parameters, recording interval)

2. Overall summary

- Count table
- Overall flag (# entries marked as pass, suspect, fail or missing data, based on the worst flag across all of the QC tests)

3. Individual parameters

- Count table
- Flags by QC test (gross, spike, rate of change and flat line)
- Time series plot
- 4. Multi-parameter time series plots (as available)

QC report – Section 1

Data Quality Control Report

Report Date: 2017-06-17

DATA FILE INFORMATION

Filename: QC_ECO6702_AirWater_20130910_20140115.csv

SiteID: ECO6702

Period of Record, Requested: 2013-09-10 to 2014-01-15

Period of Record, Actual: 2013-09-10 to 2014-01-15

Recording Interval: 30 minutes

Data Type: AirWater

Parameters Included: Water.P.psi, Water.Temp.C, Air.BP.psi, Sensor.Depth.ft, Air.Temp.C,

DATA SUMMARY, OVERALL

Estimated number of records per day is 48.

Number of records by year and month (with totals).

	1	9	10	11	12	Sum
2013	0	984	1488	1440	1488	5400
2014	696	0	0	0	0	696
Sum	696	984	1488	1440	1488	6096

Number of records by day and month (with totals).

						-
	1	9	10	11	12	Sum
1	48	0	48	48	48	192
2	48	0	48	48	48	192
3	48	0	48	48	48	192
4	48	0	48	48	48	192
5	48	0	48	48	48	192
6	48	0	48	48	48	192
7	48	0	48	48	48	192

QC report – Section 2

'Count' tables (# measurements/day)

Month

Sensor was set to record at 30-minute intervals. This equates to 48 measurements per day.

Day

Overall flags by parameter

Water.P.psi

Water.Temp.C

Air.BP.psi

P	Sum
6096	6096

Sensor.Depth.ft

F	P	S	Sum
1	6068	27	6096

Air.Temp.C

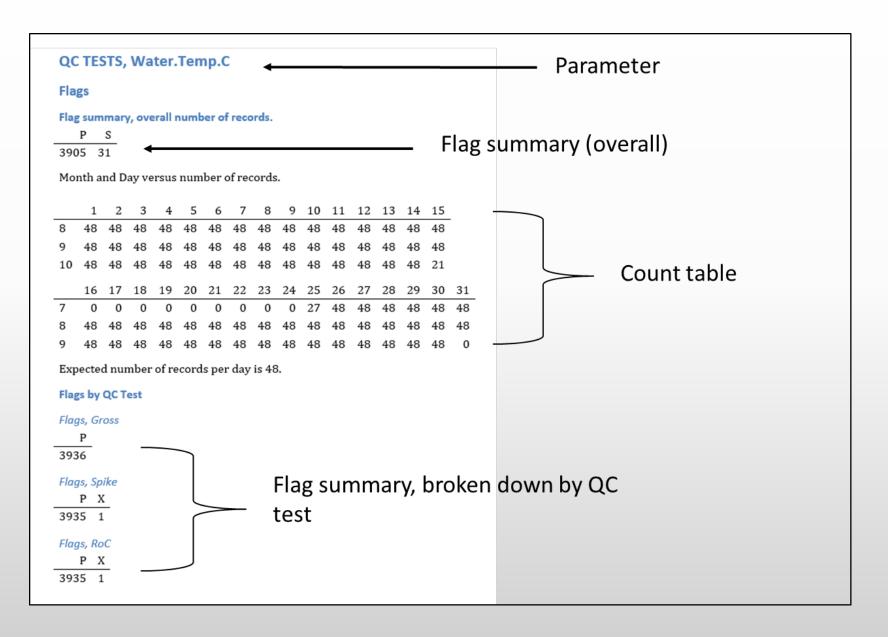
QC report – Section 2

Overall summary -

For each parameter, # of entries marked as pass (P), suspect (S), fail (F) or missing data (X), for any of the QC tests.

In this example, 6047 of the water temperature entries were marked as pass, and 49 were marked as suspect.

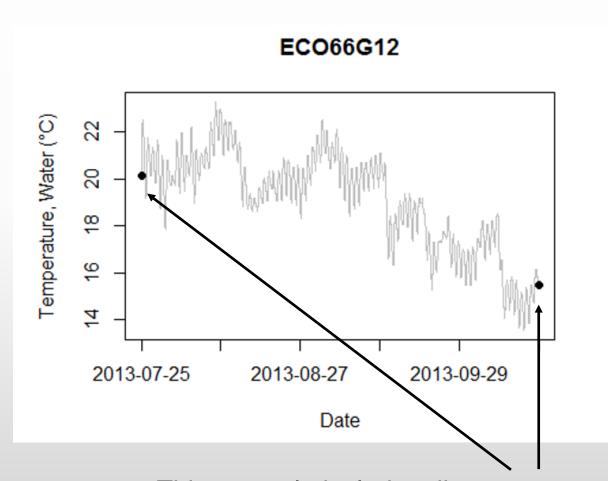
QC report – Section 3



Individual parameters – e.g., water temperature

QC report – Section 3

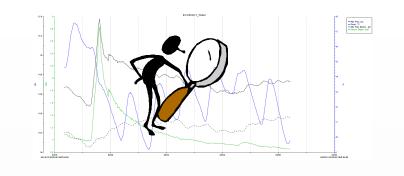
Individual parameters continued...



Time series plot (all values plotted vs. date/time) for water temperature

This example includes discrete measurements (this feature is not currently available in TNCON)

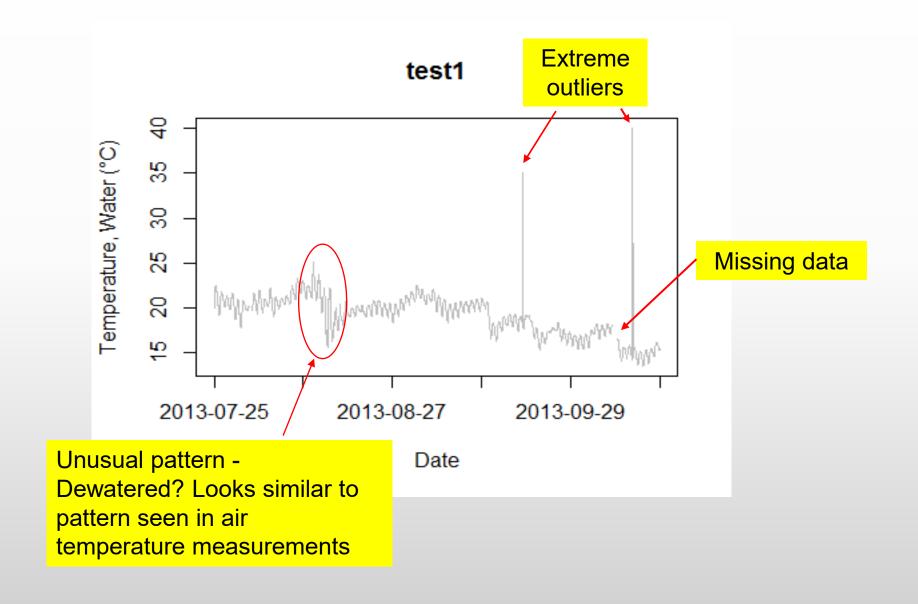
Visual checks on time series plots



- Missing data
- Closely matching water and air temperature values (this indicates that the stream sensor may have been out of the water)
- Diel fluxes with flat tops, or "chatter" in the gage height record (this indicates that the sensor may have been buried in sediment)
- Water level values of 0 (this could mean that the pressure transducer was dewatered. With vented transducers, another possibility is that moisture got into the cable and caused readings of zero water depth)
- Water level values that are negative
- Outliers or rapidly fluctuating values (it is possible that the sensor moved, e.g., due to a high flow event or vandalism)
- Values recorded before the sensor was correctly positioned & stabilized, or after the sensor was removed from its location ("trimming")

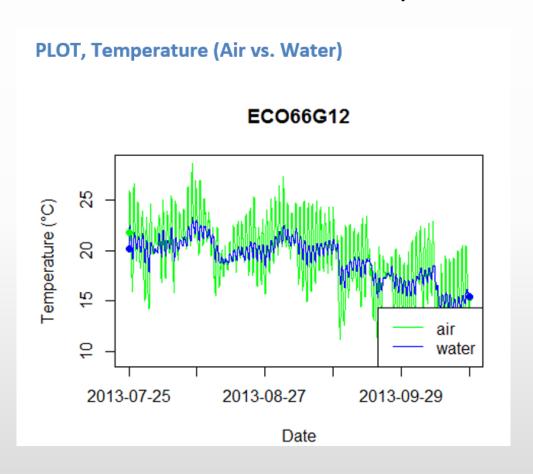
Note: you can also do these visual checks on plots in HOBOware or with other software

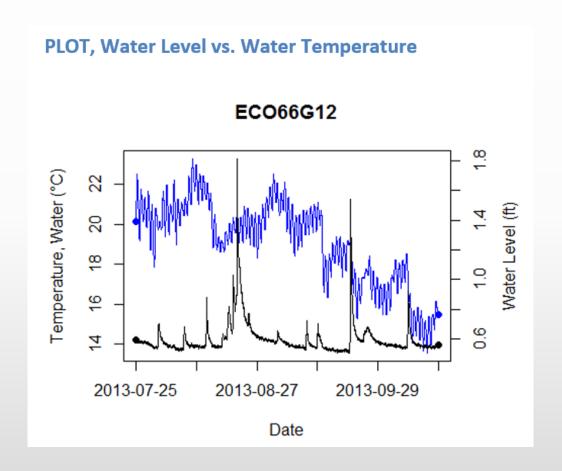
Examples of visual checks on time series plots



QC report – Section 4

Multi-parameter time series plots





Do the patterns make sense? (e.g., water temperature should fluctuate less than air temperature; increases in water level may correspond with decreases in water temperature)

QC .csv file

Lots of columns!

The R scripts add in many additional fields – flags (overall & each parameter/QC test), plus RAW and Comment fields

The following tips will make them easier to work with...

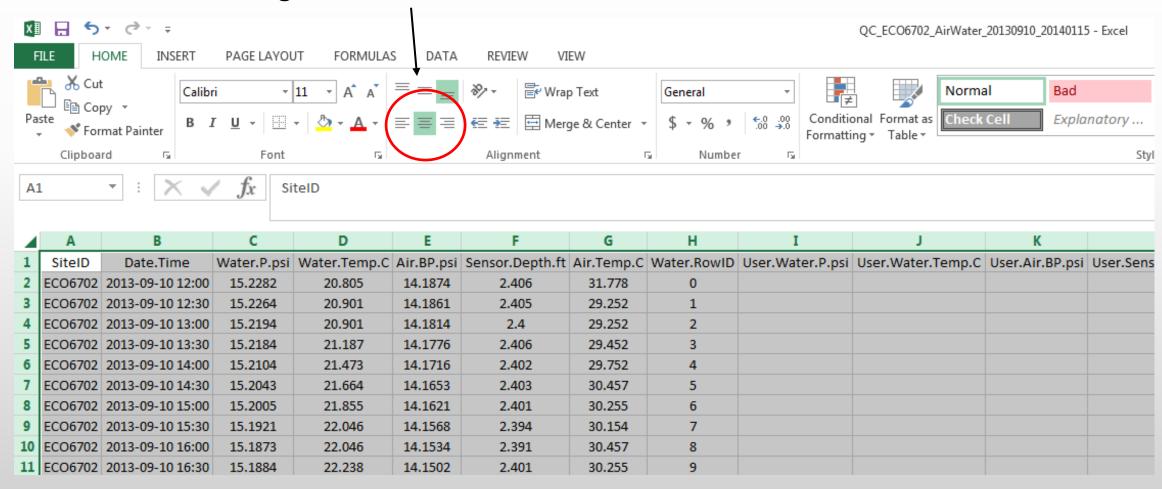
Tip for easier navigation

Highlight all the fields, then click **Format – Autofit Column Width**. Excel then automatically adjusts column widths so that you can see the full field names.

x∄ FIL	E HOME	T FAGE LA	YOUT	FORMULAS DATA	REVIEW VIEW Nu	ince PDF	QC_EC	O66G12_Aw_20130426_a	20130725 - Excel		ηη		0	
Past	Clipboard	Calibri B I U +	- 11 H - 2	A^ A = = = = = = = = = = = = = = = = = =	 Wrap Text Merge & Center Alignment	General \$ - % 9 6.0	Conditional Form	at as Check Cell	Bad Good Explanatory Inpu		ition Inse	t Delete Format	AutoSum Z Z Fill Y Sort & Filter Y	Find &
A1	▼	\times \checkmark f_x	SiteID									‡□ Row I	Height Fit Row Height	
	Α	В	С	D	E	F	G	Н	I	J	K		mn Width	
1	SiteID 🔻	Date 🔻 T	ime 💌	Date.Time	Water.Temp.C 🔻	Air.Temp.C	Water.BP.psi 🔻	Air.BP.psi 🔻	Water.Level.ft 🔻	Discrete.Water.Temp.C	Discrete.Air.1			⁄el.ft 🔻
2	ECO66G12	2013-04-26 10:	30:00	2013-04-26 10:30	10.063	15.282	14.8024	14.3638	0.985	10.055	15.24		ult Width	
3	ECO66G12	2013-04-26 11:	00:00	2013-04-26 11:00	10.357	14.134	14.7461	14.3644	0.854	NA	NA	Visibility		
4	ECO66G12	2013-04-26 11:	30:00	2013-04-26 11:30	10.748	14.23	14.7428	14.3603	0.855	NA	NA	Hide	& <u>U</u> nhide	
5	ECO66G12	2013-04-26 12:	00:00	2013-04-26 12:00	11.139	15.473	14.7441	14.3627	0.853	NA	NA	Organize	Sheets	
6	ECO66G12	2013-04-26 12:	30:00	2013-04-26 12:30	11.528	17.284	14.7431	14.3612	0.854	NA	NA	<u>R</u> enai	me Sheet	
7	ECO66G12	2013-04-26 13:	00:00	2013-04-26 13:00	11.916	18.806	14.7421	14.3628	0.848	NA	NA	<u>M</u> ove	or Copy Sheet	
8	ECO66G12	2013-04-26 13:	30:00	2013-04-26 13:30	12.207	19.662	14.7453	14.3629	0.855	NA	NA	<u>T</u> ab (Color	
9	ECO66G12	2013-04-26 14:	00:00	2013-04-26 14:00	12.594	20.424	14.7351	14.3537	0.853	NA	NA	Protectio	n	
10	ECO66G12	2013-04-26 14:	30:00	2013-04-26 14:30	12.883	20.71	14.7315	14.3484	0.857	NA	NA	Prote	ct Sheet	
11	ECO66G12	2013-04-26 15:	00:00	2013-04-26 15:00	13.173	20.901	14.7234	14.3426	0.852	NA	NA	<u> </u>	Cell	
12	ECO66G12	2013-04-26 15:	30:00	2013-04-26 15:30	13.365	20.805	14.7171	14.3331	0.859	NA	NA	Form	at Cells	
13	ECO66G12	2013-04-26 16:	00:00	2013-04-26 16:00	13.558	21.091	14.7109	14.3302	0.852	NA	NA		- NA	
14	ECO66G12	2013-04-26 16:	30:00	2013-04-26 16:30	13.654	21.569	14.702	14.3215	0.851	NA	NA		NA	

Tip for easier navigation

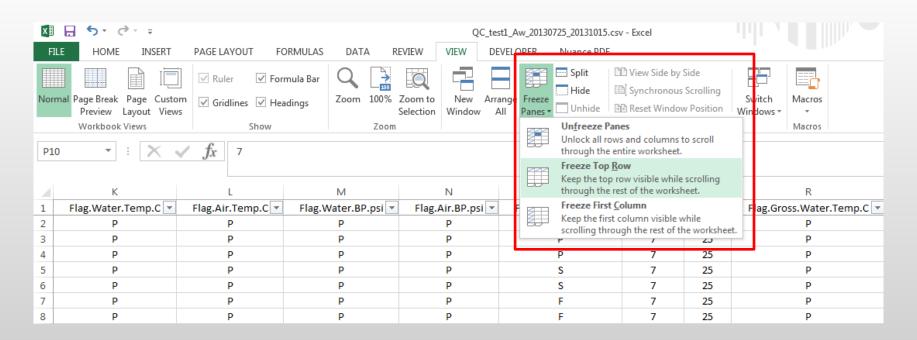
Alignment - center the entries in the columns



Tip for easier navigation

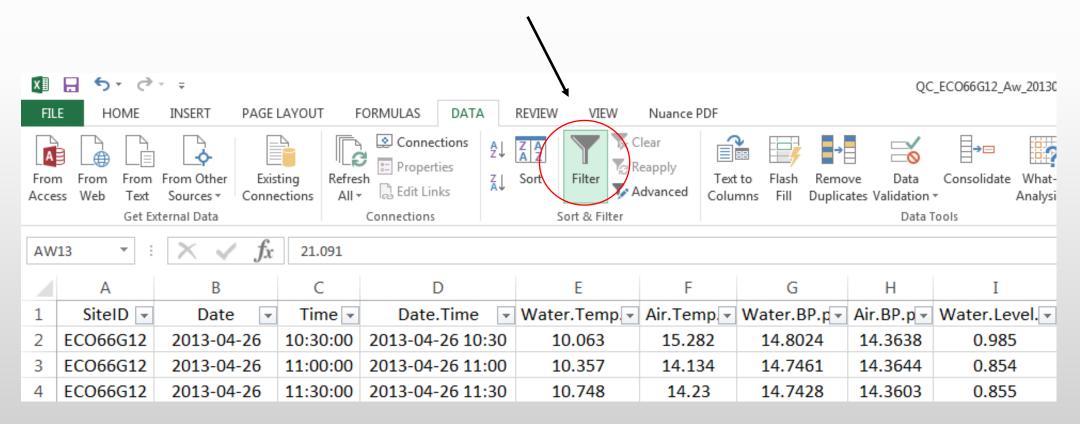
You can also **Freeze Panes – Freeze Top Row** to make viewing easier (this allows you to view the column headings while you scroll down).

You can also use the **Split** function (this allows you to split the screen so that you can see the column headings while scrolling down, and the left-hand columns (e.g., date-time field) while scrolling to the right.



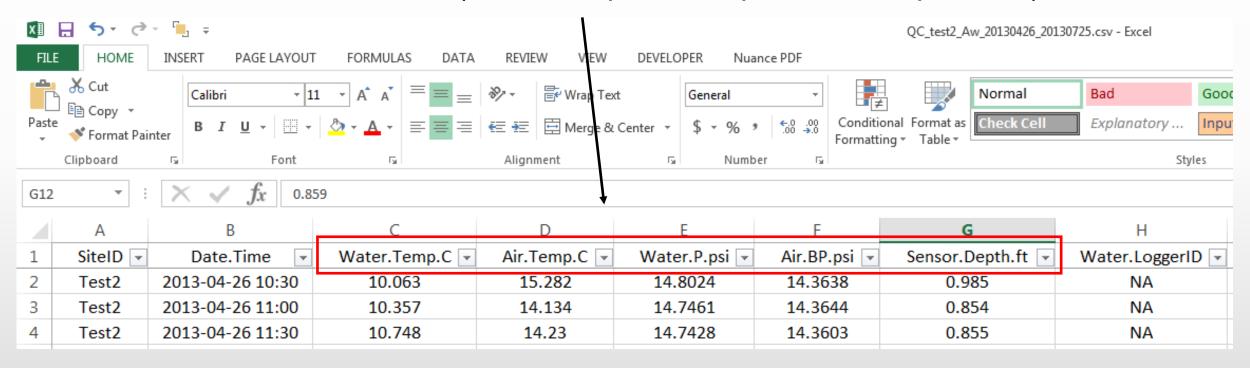
Tip for easier navigation

Using the **Filter function** can help too (e.g., you can filter for entries flagged as 'F' or 'S' and evaluate)



Interpretation

Values (water temp, air temp, sensor depth, etc.)

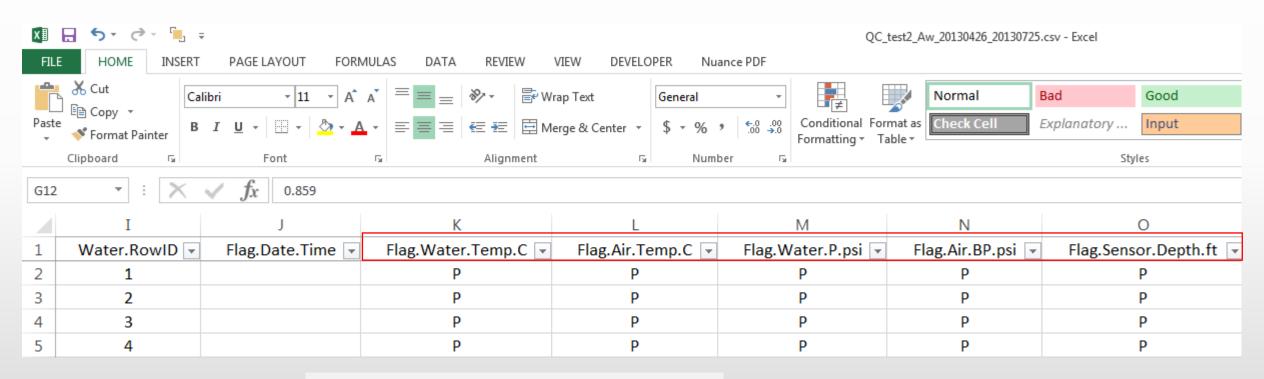


Scrolling to the right...



.csv file continued...

Overall flags (entries marked as pass, suspect, fail or missing data, based on the *worst* flag across all of the QC tests).



F = fail

P = pass

S = suspect

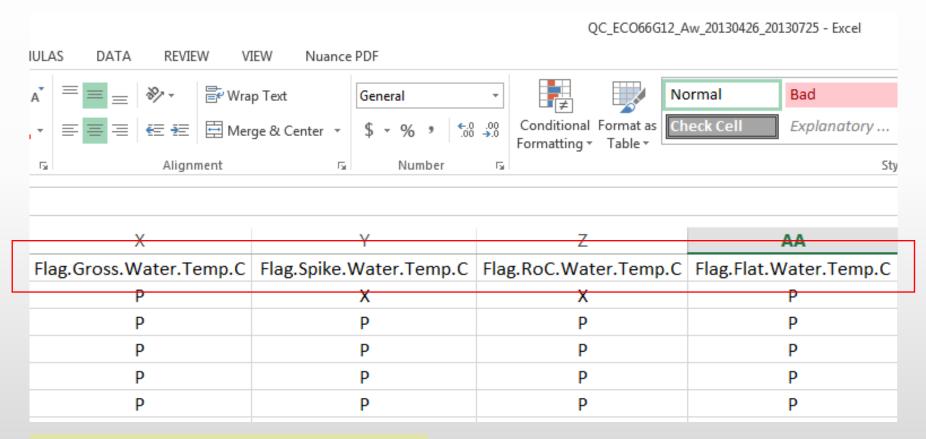
X = Not Data or Not Applicable (NA)

Continue scrolling to the right



.csv file continued...

Flags for each parameter and each QC test (gross, spike, RoC, flat)



Example - water temperature



.csv file continued...

RAW = original values; if you end up making corrections in the primary value fields, the RAW fields allow you to retain the original records and track any changes you make.

COMMENT.MOD = allows you to comment on any changes you make.

ΧI	QC_ECO66G12_/-											
FIL	FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW Nuance PDF											
*	Calibri 11 A A A S S S Wrap Text Paste Format Painter Clipboard Font Alignment Alignment Alignment General Conditional Format as Formatting Table Alignment Alignment											
	AS	AT	AU	AV	AW							
1	RAW.Date.Time	Comment.MOD.Water.Temp.C	RAW.Water.Temp.C	Comment.MOD.Air.Temp.C	RAW.Air.Temp.C							
2	2013-04-26 10:30		10.063		15.282							
3	2013-04-26 11:00		10.357		14.134							
4	2013-04-26 11:30		10.748		14.23							
5	2013-04-26 12:00		11.139		15.473							
6	2013-04-26 12:30		11.528		17.284							

Corrections

In general, 3 possible actions can be taken:

- Leave data as is
- Apply correction factor
- Remove data

Corrections should not be made unless the cause(s) of error(s) can be validated or explained in the field notes or by comparison with information from nearby stations.

Accurate field notes and accuracy check logs are essential in the data correction process.

Any **discrepancies should be documented** in your data file and any actions you take should also be carefully documented.

Corrections

Right now we don't have any automated corrections programed into the QC package (this is a future 'wish list' item).

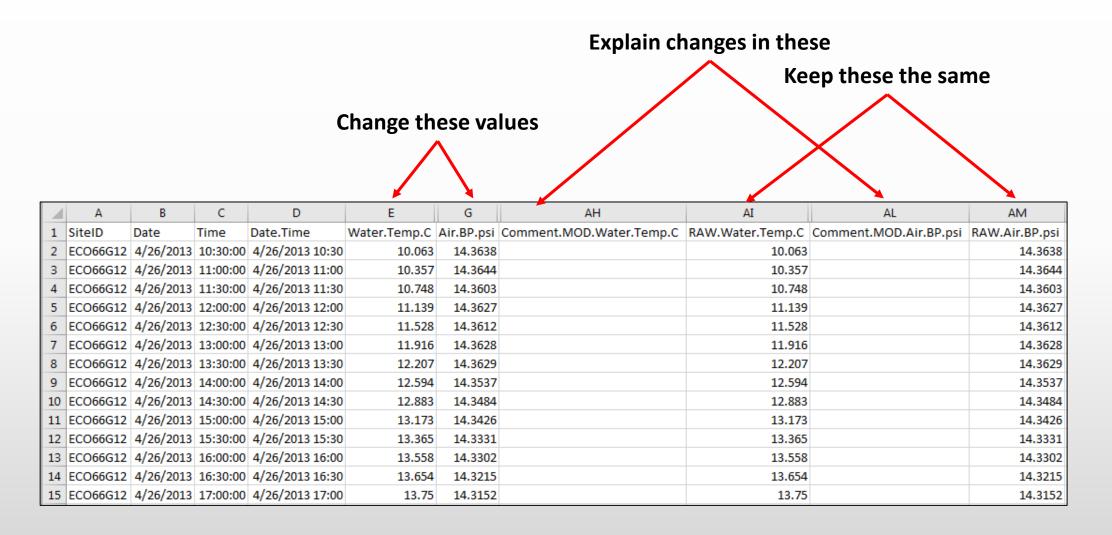
Any corrections you make need to be done manually.

Some types of corrections are fairly straightforward and can be done in Excel (e.g., change values by same amount, or basic drift corrections); others are more complex.

Potential references -

- Wagner et al. 2006 (USGS) (more detailed)
- EPA 'Best Practices' report (more general has a table with a general summary of different types of problems that can occur (e.g., missing data, failed accuracy check) and recommended actions for addressing them)

If you want to make a change in your data...

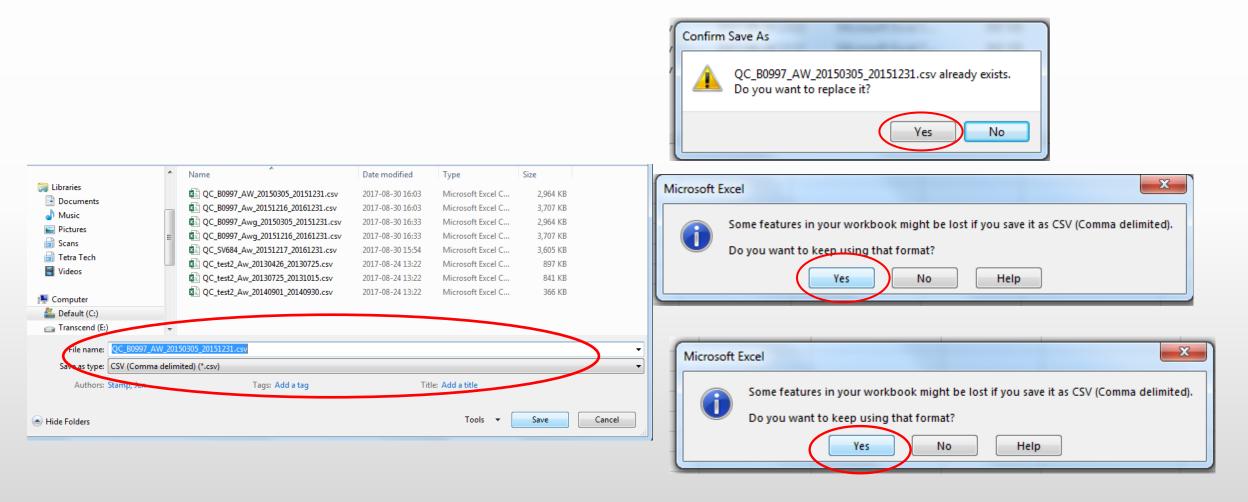


NOTE: Any changes to recorded data (including deleting data) should be done carefully and fully documented.

Resave as .csv file

Make sure you resave the corrected QC .csv file as a CSV (Comma delimited)

Click 'Yes' through the multiple series of prompts.



Seeking feedback!

We are still learning & gaining experience with the QC process ourselves.

Please provide feedback as you go through these files.

- Are there certain types of errors the QC reports are not catching?
- How well are the QC test thresholds working?

There will be some things we can fix and some things we cannot fix (those will go on a 'wish list').