**Acton Notes for QAQC Files**

**Most up-to-date files on GEISHA FTP site (as of 9 June 2018)**

**Updated list on FEMC (as of 1 June 2020)**

Lake\_Acton\_QAQC\_Notes\_1Jun2020.docx

LakeMetadata\_Acton\_QAQC\_10Jan2018.xlsx

RawSondeData\_Acton\_QAQC\_27May2020.csv

RawWeatherData\_Acton\_OXF122\_QAQC\_09Jun2018.csv

RawNutrientData\_Acton\_QAQC\_31May2020.csv

RawPhytoData\_Acton\_QAQC\_8Apr2020.csv

RawLakeHFData\_Acton\_QAQC\_22May2020.csv

SpeciesList\_Acton\_QAQC\_14Jan2018.xlsx

Hypsometry\_Acton\_QAQC\_16Apr2018.xlsx

PhytoMetadata\_Acton\_GEISHA.docx (still waiting on this)

DailyWeatherData\_Acton\_OXF122\_QAQC\_09Jun2018.csv

See the “Log” of changes made to the files since first QAQC’d. These (if any) are located at bottom of this document.

**LakeMetadata\_Acton\_QAQC\_14Jan2018.xlsx**

* Under worksheet “ManualSonde”
* Changed “water\_temperature” and “dissolved\_oxygen” from “n” to “y” for both rows.
* Under worksheet “BuoyHighFrequency”
* Added “y” to “water\_temperature”, “dissolved\_oxygen”, “chla\_fluorescence”, and “ph”, and “no” to “phyco\_fluorescence”.
* Still need sensor models for “water\_temperature”, “dissolved\_oxygen”, “chla\_fluorescence”, and “ph”
* Under worksheet “MethodForPhytoplankton”
* Still need “count\_method”
* Added “n” to “biomass\_y\_n” and “NaN” to “biomass\_method”
* Still need “methods\_change\_y\_n”, “taxonomic\_expertise”, “data\_checked\_y\_n”, and “comments”
* Under worksheet “MeteorologicalMetadata”
* Added “OXF122” as another met\_stationid but need the latitude and longitude and location.
* Under worksheet “LakeType”
* Still need “colour\_type”

**RawWeatherData\_Acton\_Multi\_QAQC\_5Mar2018.csv**

* Data are collected at 1-hour time scale for the met\_stationd OXF122 and at 15-minute time scale for met\_stationid Outflow and Inflow. There was some issue with the data at 15-min scale (precipitation) so data provider re-compiled data at hourly scale for all weather stations.
* 150 wind\_dir\_deg were 360 so these were changed to 0 (because 360 and 0 are same values)
* 430 radiation\_j\_sec\_m2 observations were “#VALUE!”. These were changed to NaN.
* Air pressure data was in inches Hg so converted to mb by multiplying by 33.8639.
* Changed datetime format to be yyyy-mm-dd HH:MM:SS

**DailyWeatherData\_Acton\_OXF122\_QAQC\_09Jun2018.csv**

* Using this single station for our analyses.
* Used the R package “openair”, function “timeAverage” to generate daily averages, including the vector average for wind direction. Rain data were summed for each day. See Technical Note by Stuart K. Grange “Average wind speeds and directions”.
* Some observations from raw data set had fewer than the maximum possible per day so that daily average values could be skewed if too many were missing. For any daily average values based on less than 83% of available observations in a given day, these values were switched to NaN.
* On 05-06-96, the radiation daily average was a high outlier. It had 20 observations, but the 4 missing were all at night (00:00-03:00) which skewed the daily average high. Changed the radiation value on this day to NaN.
* Changed date format to YYYY

**RawSondeData\_Acton\_QAQC\_11Jan2018.csv**

* Only two observations (0 and 0.5 m) on 02-06-97. Both were removed.
* On 03-09-03, surface (0 m) was 2 C lower than at 0.5 m (22.5 vs 24.5 C). From data provider Tanner Williamson: “Correct according to the field data book, but I would discard. It's not usual for the 0M to be a few tenths of a degree colder than the 0.5M in the spring and fall. But I have difficulty believing >1C difference could be sustained. Same goes for the rest of the data point where 0.5M temp is > 0M temp by >1C”. Accordingly, the temperature value at 0 m was removed.
* On 04-09-07, surface (0 m) was 1.4 C lower than at 0.5 m (24.2 vs 25.6 C). The temperature value at 0 m was removed.
* Only three observations (0, 0.5, and 1 m) on 06-06-11. The three were removed.
* Only one observation (0 m) on 08-06-15. Removed.
* On 08-08-01, temperature at 2.5 m is 36.31 C, about 10 C higher than temps above and below. Data provider indicated this value should be 26.31. Value was changed accordingly.
* On 10-09-08, temperature at 1.5 m is 36.6 C, about 13 C higher than temps above and below. Data provider indicated this value should be 23.6. Value was changed accordingly.
* On 10-09-92, surface (0 m) was 1.8 C lower than at 0.5 m (20.2 vs 22 C). The temperature value at 0 m was removed.
* On 12-09-07, surface (0 m) was 4.5 C lower than at 0.5 m (18.9 vs 23.4 C).Data provider indicated that the values for this profile got shifted by one value. The profile should be:
* 0M     = 23.4
* 0.5M  = 23.8
* 1M     = 24.1
* 1.5M  = 24.1
* 2M     = 24.2
* 2.5M  = 24.2
* 3M     = 24.2
* 3.5M  = 24.2
* 4M     = 24.2
* 4.5M  = 24.2
* 5M     = 24.2
* 6M     = 24.1

The values were changed accordingly.

* Only one observation (0 m) on 13-07-15. Removed.
* On 17-05-10, surface (0 m) was 3.4 C lower than at 0.5 m (15.2 vs 18.6 C). The temperature value at 0 m was removed.
* On 20-04-06, temperature at 3 m is 7 C, about 5-6 C lower than temps above and below. Correct from the log book, but data provider suggested to remove the temperature value at 3 m, so it was removed.
* Only one observation (0 m) on 20-07-15. Removed.
* On 21-04-00, temperatures at 0 and 0.5 m were 6.5 C and all deeper temperatures were 11 C or 11.2 C. The temperature values at 0 and 0.5 m were removed.
* On 24-06-13, temperature at 1.5 m is 29.8 C, about 4 C higher than temps above and below. Data provider indicated this value should be 25.8 C. Value was changed accordingly.
* On 28-05-14, temperature at 3.5 m is 12.9 C, at least 4 C lower than temps above and below. Data provider indicated this value should be 17.9 C. Value was changed accordingly.
* On 29-06-06, temperature at 3 m is 27.7 C, higher than temps above and below. Correct from the log book, but data provider suggested to remove the temperature value at 3 m, so it was removed.

**RawPhytoData\_Acton\_QAQC\_18Jan2018.csv**

* 32 taxa had “cf.” before the genus or between the genus and species. The “cf.” was removed in each case as best representation of the taxa.
* See “SpeciesList\_Acton\_QAQC\_14Jan2018.xlsx” to see name changes which include updates from Algaebase, name corrections, etc.
* The following dates had NaN for their max\_depth\_m value: 01-09-10, 02-07-14, 08-10-98, 13-06-11, 15-09-97, 17-05-99, 17-10-03, 20-07-13, 21-07-03, 22-10-98, 25-05-99, 26-06-95, 29-10-98, 30-06-03, 30-09-07, 31-08-09. Data provider found values for three of the dates: 01-09-10 = 1.75 m; 13-06-11 = 1.25 m; 21-07-03 = 1.0 m. These were replaced and remainder will remain NaN
* Some names are ambiguous. Tanner is trying to get clarification.
* JB - Email from Mike Vanni on 6 March 2020 – don’t use phyto data after 2012. Change in methods/personnel so removed data after 2012 and made a new file with updated date: RawPhytoData\_Acton\_QAQC\_8Apr2020.csv

**RawNutrientData\_Acton\_QAQC\_5Mar2018.csv**

* All dates were changed to the correct dd\_mm\_yy format
* All nutrient measurements were integrated from the 1% PAR depth to surface, except TP on 24-04-96, 09-07-96, 26-08-96, 04-11-96, 09-06-97, 17-05-99, 06-07-99, 25-10-99, 02-06-00, and 02-08-00 was measured at 0 m.
* Changed all “NA” values to “NaN”
* NH4 (mgN/L) values were multiplied by 1,000 and variable name changed to nh4\_ug\_l.
* N03 (mgN/L) values were multiplied by 1,000 and variable name changed to no3\_ug\_l.
* TN (ugN/L) was changed to variable tn\_ug\_l
* TP (ugP/L) was changed to variable tp\_ug\_l
* SRP (ugP/L) was changed to variable orthophosphate\_ug\_l
* SS (mg/L) was changed to tss\_mg\_l.
* Removed NVSS.
* depth\_m values originally represented the depths of 1% PAR penetration, which was the lower depth from which integrated nutrient samples were taken. For our vocabulary, integrated samples are coded as -999. So all the depth\_m values were changed to this.

**RawLakeHFData\_Acton\_QAQC\_11Jan2018.csv**

* All dates were changed to the correct date\_time format (dd-mm-yy HH:MM)
* Data are collected at 15 minute (from start to 21-04-16 08:00) and at 10 minute (from 28-04-16 13:40) intervals.
* Temperature values from 24-03-15 through 31-03-15 were all over the place (above 40 C). Changed all data during this week to “NaN”.
* Temperature values on 26-12-14 were abnormally high (for those with data) – all in the mid- to high-teens. Chl a values ranged from 5.4 to 497. All 22 observations were changed to NaN.
* On 05-08-10 at 13:45 and 14:00, temperature values were 2 and 4 C less than adjacent temperatures, and other variable values were also strange. Data for these two date\_time observations were changed to NaN.
* On 29-09-10 13:00 temperature was 4C lower than adjacent. Data for this date\_time were changed to NaN.
* The only six observations on 09-04-14 had temperature values > 20 C. Other variables were strange too. All changed to NaN.
* On 29-08-14 at 08:30 and 08:45, temperature values were 6 and 3 C less than adjacent temperatures, and other variable values were also strange. Data for these two date\_time observations were changed to NaN.
* On 11-09-14 08:15 temperature was 6C lower than adjacent. Data for this date\_time were changed to NaN.
* On 20-08-15 10:15 temperature was 3C lower than adjacent. Data for this date\_time were changed to NaN.
* On 28-08-15 14:00 temperature was 2C lower than adjacent. Data for this date\_time were changed to NaN.
* On 28-07-16 09:30 temperature was 3C lower than adjacent. Data for this date\_time were changed to NaN.
* On 26-09-16 10:50 temperature was 5C lower than adjacent. Data for this date\_time were changed to NaN.
* On 02-09-16 08:00 temperature was 8C lower than adjacent. Data for this date\_time were changed to NaN.
* On 10-10-16 08:20 temperature was 4C lower than adjacent. Data for this date\_time were changed to NaN.
* Chl a values hit a constant ceiling for all 2015 observations and most of the 2016 observations. All these Chl a values were changed to NaN.
* From 15-08-14 to 18-08-14, chla a values also maxed out but this was a brief period in 2014 – not sure if this is real or not so I left in.
* Chl a values were < 0 on 27-06-13 11:30 and 11-09-13 12:30. These values were changed to 0.
* Changed datetime format to be yyyy-mm-dd HH:MM:SS and filled NaN dates – updated file name to RawLakeHFData\_Acton\_QAQC\_22May2020.csv

**SpeciesList\_Acton\_QAQC\_14Jan2018.xlsx**

* See this list for original names and updated names used for our QAQC data set to standardize names.

**Hypsometry\_Acton\_QAQC\_16Apr2018.xlsx**

* Added this file for Schmidt stability and lake number estimates.
* Updated column names to remove empty rows at the top

**Log of Changes**

16 April 2018

**Hypsometry\_Acton\_QAQC\_16Apr2018.xlsx was added to the FTP site to be used to estimate Schmidt stability and lake number.**

09 June 2018

**RawWeatherData\_Acton\_Multi\_QAQC\_5Mar2018.csv updated to RawWeatherData\_Acton\_OXF122\_QAQC\_09Jun2018.csv**

* Focusing on single weather station at buoy which has most data.
* Changed datetime format to YYYY-mm-dd HH:MM:SS

**DailyWeatherData\_Acton\_OXF122\_QAQC\_09Jun2018.csv**

* Provide weather data at daily time step for the derived physics data set.
* Changed date to four numbers for year YYYY (dd-mm-yyyy)

8 April 2020

**New file name = RawPhytoData\_Acton\_QAQC\_8April2020.csv**

* Update Raw Phytoplankton dataset to remove data after 2012 due to change in methods/personnel
* Changed date to four numbers for year YYYY (dd-mm-yyyy)

22May2020

**RawLakeHFData\_Acton\_QAQC\_11Jan2018.csv updated to RawLakeHFData\_Acton\_QAQC\_22May2020.csv**

* Changed datetime format to YYYY-mm-dd HH:MM:SS
* Some dates missing so filled in based on surrounding dates

27May2020

**RawSondeData\_Acton\_QAQC\_11Jan2018.csv updated to RawSondeData\_Acton\_QAQC\_27May2020.csv**

* Changed date to four numbers for year YYYY (dd-mm-yyyy)

31May2020

**RawNutrientData\_Acton\_QAQC\_5Mar2018.csv updated to RawNutrientData\_Acton\_QAQC\_31May2020.csv**

Changed date to four numbers for year YYYY (dd-mm-yyyy)