**Geneva QAQC Notes:**

**Most up-to-date files on GEISHA FTP site (as of 13 August 2018)**

SpeciesList\_Geneva\_QAQC\_07Jul2018.xlsx

Lake Geneva QAQC Notes\_01Jun2020.docx

RawPhytoData\_Geneva\_QAQC\_17May2020.csv

RawSondeData\_Geneva\_QAQC\_27May2020.csv

DailyWeatherData\_Geneva\_INRA\_QAQC\_17May2020.csv

RawNutrientData\_Geneva\_QAQC\_31May2020.csv

LakeMetaData\_Geneva\_GEISHA.xlsx

Hypsometry\_Geneva\_QAQC\_18Apr2018.xlsx

PhytoMetadata\_Geneva\_QAQC\_04Jun2018.docx

In all files, special characters were removed and replaced i.e. Võrtsjärv to Vortsjarv. See the “Log” of changes made to the files since first QAQC’d. These are located at bottom of this document.

**RawWeatherData\_Geneva\_INRA\_QAQC\_06Sep2017.csv**

* Inserted “lake” and “stationid” columns. Modified data and time to get correct “date\_time” format using R.
* Deleted last 7 rows of data as only had date and time.
* 682 observations had rel\_hum\_perc > 100 so these values were changed to 100.
* Radiation units were given in joules/cm2. This appears to be the total for the hour. As we want joules/sec/m2, we multiplied the given value by 10,000 (to convert to m2) and then divided by 3,600 (to convert to per second).
* Anomalous max\_wind\_m\_s on 21-06-95 01:00 (82 m/s) and 19-10-06 01:00 (86 m/s), when no other wind data were recorded or very low avg\_wind\_m\_s (< 5 m/s). These two observations were changed to NaN.
* On 23-06-15, there were 38 observations where air\_temp\_c ranged from -19.1 to -35.3. These observations were changed to NaN.
* There were 27,833 observations where avg\_wind\_m\_s was <= max\_wind\_m\_s as expected, but 33,215 observations where avg\_wind\_m\_s was > max\_wind\_m\_s, which is not expected. Deleted all max\_wind\_m\_s values. Will not use.
* Daily average wind speed (avg\_daily\_wind\_m\_s) and daily maximum wind speed (max\_daily\_wind\_m\_s) and the total wind (not sure what this means, Orlane is checking) for N, NE, E, SE, S, SW, W, and NW were available from a meteorological station. Because we do not have any wind direction data for the hourly observations we grabbed these data. The daily observations were put on their respective day at 12:00.

**RawPhytoData\_Geneva\_QAQC\_06Sep2017.csv**

* Removed all cyst data.
* Went through all phytoplankton names with data providers.
* Some very high values (39 mg/L) for Mougeotia.

**RawSondeData\_Geneva\_QAQC\_18Aug2017.csv**

* Changed 369 instances of “leman” to “Geneva” for “lake” column
* On 10-07-95, 33 observations had depths >= 310 m. Another 115 observations also had depths >= 310 m. These were all deleted as the maximum depth of Lake Geneva is 309 m.
* 16 observations had abnormally low water temperatures at depths greater than 0 meters (values of -3.03).
	+ These points were changed to NaN
* On 19-08-96 there is a depth of -23 meters, other data points did not make sense.
	+ This entire data entry (row) was removed
	+ Data on 21-04-09 had missing data from 0 to 8 m, but these were found at deeper depths (almost 300 m) and were replaced.
	+ Only 2 observations for 22-04-14 so these were deleted.
	+ Temperature on 11-12-96 at 166 m was 17.82, which was far greater than rest of water column (~5 C) so this value was changed to 5.94, which is same as values adjacent to 166 m.
* Many days where oxygen percentage is greater than 120%. Multiple oxygen probes were used and not always calibrated, etc. Data person recommended not using oxygen percent data at all. So these are all NaN.
	+ However, the oxygen data (mg/L) from probes were calibrated against chemically-derived oxygen values (Winkler method) and these data are considered good.
* Conductivity was provided in mS/cm instead of uS/cm, so the values were multiplied by 1,000
* On 06-04-09, the pH values were close to 6 in upper waters, so making all pH values for this day as NaN.
* pH on 06-07-09 at depths 298, 299, 300 were very different from hypolimnetic values. These 3 depths were changed to NaN for pH.
* Conductivity profiles were wonky on 23-08-20, 12-02-01, 26-03-01, 24-04-01, 12-11-13, 05-05-14, 02-06-14. Conductivity values on these dates were changed to NaN.

**RawNutrientData\_Geneva\_QAQC\_15Aug2017.csv**

* Chlorophyll was provided in mg/L instead of ug/L, so values were multiplied by 1,000.
* Several TP values were very high.



* Orlane checked these values and has sent corrections which are the following and have been made:



* Many days and depth when the sum of nitrogen species are greater than total nitrogen
	+ There were 45 observations where TN = 0. Orlane indicated these are likely missing values and not 0, so the 45 observations were changed from 0 to NaN. The corresponding NO3, NO2, and NH4 values were also changed to NaN. The dates were 19-08-02, 14-04-04, 15-10-07, 03-03-08, 31-03-08, and 29-03-10.
* Many days and depth when phosphate is greater than total phosphorus
	+ TP and PO4 are measured separately so there is likely several layers of variability that contribute to TP being lower than PO4. However, large differences were very rare. The following table shows all cases when difference > 4. We chose only to modify the values on 29-03-04 (PO4 changed from 200 to 20), 17-09-75 (TP changed from 0 to 100 based on adjacent values), and 06-08-08 (changed PO4 from 67 to 54 based on adjacent values) when differences ranged from 171 to 11. All other differences were <= 10 so we let these be.



* On 13-11-78 at 309 m, nh4\_ug\_l was 267 and was changed to NaN.
* On 06-09-04 at 10 m, no2\_ug\_l was 100 and was changed to NaN.
* On 27-07-11 at 309, sio2\_ug\_l wsa 8960 and was changed to NaN.

**RawLakeHFData**

* No HF data

**Hypsometry\_Geneva\_QAQC\_18Apr2018.xlsx**

* Added this file for Schmidt stability and lake number estimates.

**Log of Changes**

12 November 2017

**LakeMetaData\_Geneva\_GEISHA.xlsx was added to FTP site.**

* Forgot to upload this file originally.

16 November 2017

**RawWeatherData\_Geneva\_INRA\_QAQC\_06Sep2017.csv was updated to RawWeatherData\_Geneva\_INRA\_QAQC\_16Nov2017.csv.**

* Found a number of dates duplicated but with no data (all NaN). These were deleted.

12 March 2018

**RawSondeData\_Geneva\_QAQC\_18Aug2017.csv was updated to RawSondeData\_Geneva\_QAQC\_12Mar2018.csv.**

* On 17-03-97, there were two observations at each of 55m, 93m, 100m, 102m, 123 m, 107m, 147m, and 158m. Most of these observations had a missing value at the shallower depth adjacent, so the first of the two depths was changed (e.g., the first 55 m observation was changed to 54 m). For the 107 m double observation, the 105 m was missing, so everything was adjusted accordingly the 106 m was changed to 105 m, and the following 107 m to the 106 m. The same was done for the double 123 m (there was no 121 m observation).
* On 21-04-09, double observations for 121m, 129m, 135m, 141m (138m missing), 148m (145m missing), 152m (150m missing), 158m (155m missing), 167m (161m missing). These were all handled as in previous bullet.

18 April 2018

**Hypsometry\_Geneva\_QAQC\_18Apr2018.xlsx was added to the FTP site to be used to estimate Schmidt stability and lake number.**

11 May 2018

**RawWeatherData\_Geneva\_QAQC\_16Nov2017.csv was updated to WeatherData\_Geneva\_QAQC\_Daily\_11May2018.csv.**

* Wind direction data was finally converted to a daily wind direction using total km per day by eight wind directions (N, NE, E, SE, S, SW, W, NS). Because this was daily, converted all of the weather values from hourly to daily values. These were averages for all variables, except for rain which was summed for each day. The date\_time was changed to 12:00 for time on each day to be consistent with format of all weather data from other lakes.

5 June 2018

**Uploaded PhytoMetadata\_Geneva\_QAQC\_04Jun2018.docx**

7 July 2018

**RawPhytoData\_Geneva\_QAQC\_06Sep2017.csv updated to RawPhytoData\_Geneva\_QAQC\_07Jul2018.csv**

* Original phyto\_name Synechocystis sp 4 µm and Synechocystis sp 2 µm had been removed from data and named XXXX in Algebase name. Data were placed back in data set and Algaebase names of Synechocystis (4 um) and Synechocystis (2 um) were used.
* The following taxa had density observation but no biovolume or biomass estimates for some observations. For these specific observations, the observations were removed per data provider: Achnanthidium exile (1 of 1 observation); Amphora pediculus (1 of 7 observations); Ankyra inerme (1 of 17 observations); Ankyra judayi (2 of 143 observations); Chlamydomonas globose (1 of 179); Chlorella vulgaris (18 of 523); Chlorophyceae intermediate (small) (1 of 106); Choricystis parasitica (1 of 133); Chrysochromulina parva (7 of 598); Craticula accomoda (1 of 7); Cyclotella (3 of 32); Desmodesmus bicellularis (2 of 6); Drepanochloris nannoselene (6 of 16); Elakatothrix (1 of 2); Elongated flagella (chlorophyceae zoospores) (1 of 3); Gymnodinium lantzschii (7 of 252); Hyaloraphidium contortum (7 of 167); Lagerheimia longiseta (1 of 3); Lanceola spatulifera (4 of 110); Messastrum gracile (1 of 4); Micractinium pusillum (colony 4 cells) (1 of 3); Monoraphidium minutum (3 of 131); Nephrochlamys rostrate (2 of 26); Nitzschia acicularis (2 of 164); Oocystis lacustris (1 of 95); Phacotus lendneri (1 of 149); Pseudanabaena arcuate (5 of 11); Pseudodidymocystis planctonica (2 of 22); Salpingoeca frequentissima (1 of 30); Scenedesmus (2 of 32); ;Stephanodiscus minutulus (5 of 411 observations).

13 August 2018

**WeatherData\_Geneva\_QAQC\_Daily\_11May2018.csv was updated to DailyWeatherData\_Geneva\_QAQC\_05Jun2018.csv**

* To make naming convention the same as with other lakes.
* Also, and important, the wind data was at a daily time step whereas other variables were hourly. To streamline the data for derived physics calculation, all weather data were converted to daily average except rain which is daily sum. We do not present the hourly weather data on the FTP site, but it is available for a subset of the data.

23 April 2020

**RawPhytoData\_Geneva\_QAQC\_23Apr2020.csv**

* Fragil***i***aria acus(extra i) was changed to Fragilaria acus
* Chlorophyceae intermediate (big) changed to Chlorophyceae indetermined\_big
* Chlorophyceae intermediate (small) changed to Chlorophyceae indetermined\_small

17May2020

**RawPhytoData\_Geneva\_QAQC\_23Apr2020.csv updated to RawPhytoData\_Geneva\_QAQC\_17May2020.csv**

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

**DailyWeatherData\_Geneva\_QAQC\_05Jun2018.csv updated to DailyWeatherData\_Geneva\_QAQC\_ 17May2020.csv**

* Changed date to four numbers for year YYYY (dd-mm-yyyy)
* Round air temp, solar radiation, relative humidity and wind variables to all have just 2 decimal places

27May2020

**RawSondeData\_Geneva\_QAQC\_12Mar2018.csv updated to RawSondeData\_Geneva\_QAQC\_27May2020.csv**

* Changed date to four numbers for year YYYY (dd-mm-yyyy)
* Rounded depth, water temperature, DO, pH, and conductivity to have 2 decimal places

31May2020

**RawNutrientData\_Geneva\_QAQC\_15Aug2017.csv updated to RawNutrientData\_Geneva\_QAQC\_31May2020.csv**

* Changed date to four numbers for year YYYY (dd-mm-yyyy)
* Some chl-a data had many decimal places so rounded to 2 decimal places