Ongoing list of questions for about the data

2019-06-27

Have data for stations

HA02

HA07 (water and met data)

HA08

HA10

HA12

HA15

CASC (why does this station have a different name?), stands for Cascadia, in the Middle Fork on Flathead River

Data from Jeremy (and hence in the database) extends back to 2013-04-01. The data from Phil for HA07 extends back to 2012-05-15. Is there earlier data from these wells?

Is there data from other wells? Maybe ones that have been discontinued?

The variables are numbered at the end for the column headers. For example, for HA10, each column header’s name has the number “4” at the end. Why is that? Is that supposed to identify the specific sensor?

HA 2, 8, and 15 have seemingly a completely duplicated set of column headers with a different number identifying them. For example, HA08 has columns with a “1” at the end, then a duplicate set of columns with a “2” at the end.

HA 08\_01 is in the well

HA 08\_02 is in Beaver Creek

HA 15\_07 is in the well

HA 15\_08 is in Wally Spring, which is connected to Beaver Creek

Each sensor type has a temperature sensor coupled to it

HA07 may have a longer term record (sensor sweet may have been located at HA06 for a year prior to current location).

Need to talk to Jeremy or Tom about where the longer term data lives (prior to winter 2011, start of 2012).

Ask Jeremy for manually downloaded datasets

Jeremy is sending me all the old data. Note that CR1000 files tend to include Table1 and Table2 spreadsheets. The Table1 data includes the sample data. The Table2 data is usually sampled at a different frequency and includes things like the battery temperature data. For my purposes, I only need the Table 1 data.

Jeremy said to also be weary that some of the data files may be duplicated or testing data to see if the CR1000 was working correctly. Be especially weary of short data files for this reason.

Looking at the metadata files

Data from HA07 started on 2012-05-15 for water and met

Data from HA08 in the well started on 2011-12-14

Data from HA08 in Beaver Creek started on 2012-5-2

Data from HA10 in well was collecting by 2012-02-03, first installed 2011-12-16

Data from HA12 in well was collecting at 1 hour intervals from 2012-04-26, from 2012-04-24-2012-04-26 was recording in 5 minute intervals

Data from CASC started at 1 hour intervals on 2012-09-06, was doing testing prior to that at 15 minute time intervals

Data from HA15 started at 1 hour time intervals on 2012-04-26

Data from HA02 for both well and creek started recording at 1 hour intervals on 2012-06-08, 2012-05-08 is when it actually started, but the DO concentration looked wonky

Is there any reason to include the record number in the final files? Is that useful to anyone?

All the data, across all sites and sensors, seem to have one missing timestamp each year that occurs around March 10 and always at 02:00:00. Could this be some kind of system maintenance thing? All the other data is present and seems consistent with the observations before and after this, only the timestamp is missing. **Answer:** This is when the clocks change for daylight savings. Not sure why the system would still collect an observation between 1 am and 3 am when 2 am technically doesn’t exist.

What’s going on with sensor 10 at HA02? It looks like it collected data for part of 2013, but was discontinued after the fact, but there are still CR1000 rows for the data that persist to today. Looks like these sensor were originally deployed in a parafluvial area, but were pulled (2012-09-21) when this area went dry.