**MMI – 2024 DOCUMENTATION**

**SLH = lead**

[**SITE LEVEL INFO:**](file:///C:\Users\shubler\Desktop\BioMonORDEQ\bugs%20analyses\MMI\Bio_MlocIDs_AWQMS_Most%20disturbed.SYB.xlsx)

* AWQMS interface was down, so as in O/E models, we based the site pulls off of work done outside of the Stations table
* Ref/Disturbed Designations – we used what was in this file, originally pulled from Stations/One.Table\_Rule.All
  + But we recognized the Most Disturbed population never had any sort of BPJ oversight.
  + Sabine visually scanned all sites originally labeled as Most Disturbed AND bugs were available (n = 286) and brought to the Reference Council for review any site she was uncertain of.
  + 23 sites were changed by Ref Council BPJ to moderately disturbed. Final Most Disturbed population n = 263.
  + SLH manually edited the Bio\_MlocIDs\_AWQMS\_Most disturbed.SYB.xlsx file to update the 23 sites to Moderately Disturbed.
    - These sites need to be updated in the Stations Table. (see file in email from SYB)
* COMID: use same “COMID” and “Nearby\_COMID” fields for associated WS /CAT metrics, as was done in O/E model.
* Reference Sites: use the same dataset as used for O/E
* Most disturbed:
  + NEED to winnow down the sites to one per stream and eliminate spatial overlap

**BUG DATA**

* AWQMS queries broken at this time, so using the data download that Lesley complied from AWQMS through (?) SQL queries (?). (‘raw\_bugs.Rdata’)
* NEED METRICS, not raw data
  + Use BCG attribute table?
  + BCG table is vetted and maintained. Better to use that one.
    - Requires adapting current METRICS code
    - DECISION: use BioMonTools from Erik Leppo?
* Tetra Tech PNW BCG attributes/taxonomy
* USU verified that bugs needs to be rarified. At least for richness metrics.
  + Begin, start with rarified bugs for all metrics. If poor model performance, we can explore running metrics with and without rarified bugs.
  + Travis fix for rarify: Need to change object to data.frame to work with rarify, otherwise this error comes up
* Error in Ops.data.frame(inbug[, sample.ID], isamp) :
* ‘==’ only defined for equally-sized data frames
* Drop all samples with < 200 total count (after rarifying).

**PREDICTORS**

FINAL DATA SET

Limited ref pop to same as used for O/E