# EDI Metadata Template (2020)[[1]](#footnote-1)

Data should be in csv text file. If starting with an Excel spreadsheet, please make sure it does not contain any formulas and comments on cells. If you need comments put them in their own column. If data were used in a database and major table linking is necessary to analyze, please de-normalize into a flat file, not just database table exports.

## Dataset Title

(be descriptive, more than 5 words):

Co-located population, sewage indicator, periphyton, and benthic macroinvertebrate data from Lake Baikal’s (Siberia) southwestern shoreline

## Abstract

(include what, why, where, when, and how)

## Investigators

(list in order as for a paper with e-mail addresses, organization and preferably ORCID ID, if you don’t have one, get it, it’s easy and free: <http://orcid.org/>) add table rows as needed

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| First Name | Middle Initial | Last Name | Organization | e-mail address | ORCID ID (optional) |
| Michael | F | Meyer | School of the Environment, Washington State University | michael.f.meyer@wsu.edu |  |
| Stephanie | E | Hampton | Center for Environmental Research Education & Outreach, Washington State University | s.hampton@wsu.edu |  |
| Ted |  | Ozersky | Large Lakes Observatory, University of Minnesota – Duluth | tozersky@d.umn.edu |  |
| Kara | H | Woo | Center for Environmental Research Education & Outreach, Washington State University | kara.woo@gmail.com |  |

## Other personnel names and roles

(dataset creators & contact, field crew, data entry etc. with e-mail addresses, organization and ORCID ID)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| First Name | Middle Initial | Last Name | Organization | e-mail address | ORCID ID (optional) | Role in project |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

## License

(Select a license for release of your data. We have 2 recommendations: [CCO – most accommodating of data reuse](https://creativecommons.org/publicdomain/zero/1.0/), & [CCBY – requires attribution](https://creativecommons.org/licenses/by/4.0/))

CCO

## Keywords

(List keywords and separate with commas. Using keywords from a controlled vocabulary (CV) will improve the future discovery and reuse of your data. The LTER CV is effective at describing ecological and environmental data. [Access the LTER CV here](http://vocab.lternet.edu/vocab/vocab/index.php). [Try this text mining service to extract LTER CV keywords from your abstract or methods](http://vocab.lternet.edu/keywordDistiller/). Additionally, please determine one or two keywords that best describe your lab, station, and/or project (e.g., Trout Lake Station, NTL LTER). This will help others discover your data by site/project).

Pharmaceuticals and Personal Care Products, Amphipoda, diatom, spirogyra, ulothrix, fatty acids, foodwebs

## Funding of this work:

Add rows to table if several grants were involved, list only the main PI, start with main grant first:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| PI First Name | PI Middle Initial | PI Last Name | PI ORCID ID (optional) | Title of Grant | Funding Agency | Funding Identification Number |
| Stephanie | E | Hampton |  | Dimensions of Biodiversity | NSF |  |
| Michael | F | Meyer |  | Graduate Research Fellowship | NSF |  |
| Michael | F | Meyer |  | Fulbright Fellowship | US State Department |  |

## Timeframe

* Begin date: August XX 2015
* End date August XX 2015
* Data collection ongoing/completed: completed

## Geographic location

* Verbal description: Southwestern shore of Lake Baikal between the town of Listvyanka and the village of Bolshoe Goloustnoe
* North bounding coordinate (decimal degree): 52.02693
* South bounding coordinate (decimal degree): 51.85530
* East bounding coordinate (decimal degree): 105.4724
* West bounding coordinate (decimal degree): 104.8148

## Taxonomic species or groups

## Methods

(please be specific, include instrument descriptions, or point to a protocol online, if this is a data compilation please specify datasets used, preferably their DOI or URL plus general citation information)

## Data Table

* Column name: exactly as it appears in the dataset. Please avoid special characters, dashes and spaces.
* Description: please be specific, it can be lengthy
* Unit: please avoid special characters and describe units in this pattern: e.g. microSiemenPerCentimeter, microgramsPerLiter, absoptionPerMolePerCentimeter
* Code explanation: if you use codes in your column, please explain in this way: e.g. LR=Little Rock Lake, A=Sample suspect, J=Nonstandard routine followed
* Date format: please tell us exactly how the date and time is formatted: e.g. mm/dd/yyyy hh:mm:ss plus the time zone and whether or not daylight savings was observed.
* Missing value code: If a code for ‘no data’ is used, please specify: e.g. -99999

Please add rows as needed

**Table name:** Chlorophyll data (chlorophylla.csv)

**Table description:** This .csv contains chlorophyll a data for each littoral and pelagic sampling location along Lake Baikal’s southwestern shoreline.

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or  code explanation or date format | Missing value code |
| site | Locational identifier for each sampling site |  |  |
| replicate | Replicate number for a given sampling site |  |  |
| filtered\_volume\_ml | Volume filtered for a given chlorophyll replicate | milliliters |  |
| sample\_volume\_ml | Volume of sample analyte during the chlorophyll extraction process | milliliters |  |
| raw\_fluo | Raw, uncorrected fluorometric reading for chlorophyll analysis |  |  |
| adjusted\_raw | Corrected fluorometric reading for chlorophyll analysis |  |  |
| chl\_conc | Chlorophyll a concentration | milligramsPerLiter |  |

**Table name:** Inverse-distance-weighted population (distance\_weighted\_population\_metrics.csv)

**Table description:** This .csv contains populaton data for each of the sampled locations. Although the majority of sites do not contain adjacent developments, we calculated inverse-distance-weighted population for each location based on neighboring settlements. A full description of the methods can be found in the associated manuscript (Meyer et al., XXXX).

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or  code explanation or date format | Missing value code |
| site | Locational identifier for each sampling site |  |  |
| distance\_weighted\_population | Inverse distance weighted population | NumberOfPeople |  |

**Table name:** Inverse-distance-weighted population (distance\_weighted\_population\_metrics.csv)

**Table description:** This .csv contains populaton data for each of the sampled locations. Although the majority of sites do not contain adjacent developments, we calculated inverse-distance-weighted population for each location based on neighboring settlements. A full description of the methods can be found in the associated manuscript (Meyer et al., XXXX).

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or  code explanation or date format | Missing value code |
| site | Locational identifier for each sampling site |  |  |
| Genus | Genus of the analyzed organism. Amphipods were identified to species, whereas snails and periphyton were kept as generalized groupings. Drapa spp. is an endemic alga and was likewise not identified to species. |  |  |
| Species | Amphipod species analyzed |  | NA |
| C12.0 | Concentration of C12:0 fatty acid | microgramsPerMilligramOfTissue |  |
| i.14.0 | Concentration of i.14.0 fatty acid | microgramsPerMilligramOfTissue |  |
| C14.0 | Concentration of C14.0 fatty acid | microgramsPerMilligramOfTissue |  |
| C14.4n5 | Concentration of C14.4n5 fatty acid | microgramsPerMilligramOfTissue |  |
| i.15.0 | Concentration of i.15.0 fatty acid | microgramsPerMilligramOfTissue |  |
| a.15.0 | Concentration of a.15.0 fatty acid | microgramsPerMilligramOfTissue |  |
| C15.1w7 | Concentration of C15.1w7 fatty acid | microgramsPerMilligramOfTissue |  |
| i.16.0 | Concentration of i.16.0 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.0 | Concentration of C16.0 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.1w9 | Concentration of C16.1w9 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.1w8 | Concentration of C16.1w8 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.1w7 | Concentration of C16.1w7 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.1w6 | Concentration of C16.1w6 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.1w5 | Concentration of C16.1w5 fatty acid | microgramsPerMilligramOfTissue |  |
| i.17.0 | Concentration of i.17.0 fatty acid | microgramsPerMilligramOfTissue |  |
| a.17.0 | Concentration of a.17.0 fatty acid | microgramsPerMilligramOfTissue |  |
| C17.0 | Concentration of C17.0 fatty acid | microgramsPerMilligramOfTissue |  |
| C17.1n7 | Concentration of C17.1n7 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.2w7 | Concentration of C16.2w7 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.2w6 | Concentration of C16.2w6 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.2w4 | Concentration of C16.2w4 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.3w6 | Concentration of C16.3w6 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.3w4 | Concentration of C16.3w4 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.3w3 | Concentration of C16.3w3 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.4w3 | Concentration of C16.4w3 fatty acid | microgramsPerMilligramOfTissue |  |
| C16.4w1 | Concentration of C16.4w1 fatty acid | microgramsPerMilligramOfTissue |  |
| C18.0 | Concentration of C18:0 fatty acid | microgramsPerMilligramOfTissue |  |
| C18.1w9 | Concentration of C18.1w9 fatty acid | microgramsPerMilligramOfTissue |  |
| C18.1w7 | Concentration of C18.1w7 fatty acid | microgramsPerMilligramOfTissue |  |
| C18.2w6t | Concentration of C18.2w6t fatty acid | microgramsPerMilligramOfTissue |  |
| C18.2w6 | Concentration of C18.2w6 fatty acid | microgramsPerMilligramOfTissue |  |
| C18.3w6 | Concentration of C18.3w6 fatty acid | microgramsPerMilligramOfTissue |  |
| C18.3w3 | Concentration of C18.3w3 fatty acid | microgramsPerMilligramOfTissue |  |
| C18:4w4 | Concentration of C18.4w4 fatty acid | microgramsPerMilligramOfTissue |  |
| C18.4w3 | Concentration of C18.4w3 fatty acid | microgramsPerMilligramOfTissue |  |
| C18:5w3 | Concentration of C18.5w3 fatty acid | microgramsPerMilligramOfTissue |  |
| C20.0 | Concentration of C20.0 fatty acid | microgramsPerMilligramOfTissue |  |
| C20.1w9 | Concentration of C20.1w9 fatty acid | microgramsPerMilligramOfTissue |  |
| C20.1w7 | Concentration of C20.1w7 fatty acid | microgramsPerMilligramOfTissue |  |
| C20.2.5.11 | Concentration of C20.2.5.11 fatty acid | microgramsPerMilligramOfTissue |  |
| C20.2.5.13 | Concentration of C20.2.5.13 fatty acid | microgramsPerMilligramOfTissue |  |
| C20.2w6 | Concentration of C20.2w6 fatty acid | microgramsPerMilligramOfTissue |  |
| C20.3w6 | Concentration of C20.3w6 fatty acid | microgramsPerMilligramOfTissue |  |
| C20.4w6 | Concentration of C20.4w6 fatty acid | microgramsPerMilligramOfTissue |  |
| C20.3w3 | Concentration of C20.3w3 fatty acid | microgramsPerMilligramOfTissue |  |
| C20.4w3 | Concentration of C20.4w3 fatty acid | microgramsPerMilligramOfTissue |  |
| C20.5w3 | Concentration of C20.5w3 fatty acid | microgramsPerMilligramOfTissue |  |
| C22.0 | Concentration of C22.0 fatty acid | microgramsPerMilligramOfTissue |  |
| C22.1w9 | Concentration of C22.1w9 fatty acid | microgramsPerMilligramOfTissue |  |
| C22.1w7 | Concentration of C22.1w7 fatty acid | microgramsPerMilligramOfTissue |  |
| C22.2w6 | Concentration of C22.2w6 fatty acid | microgramsPerMilligramOfTissue |  |
| C22.4w6 | Concentration of C22.4w6 fatty acid | microgramsPerMilligramOfTissue |  |
| C22.5w6 | Concentration of C22.5w6 fatty acid | microgramsPerMilligramOfTissue |  |
| C22.3w3 | Concentration of C22.3w3 fatty acid | microgramsPerMilligramOfTissue |  |
| C22.4w3 | Concentration of C22.4w3 fatty acid | microgramsPerMilligramOfTissue |  |
| C22.5w3 | Concentration of C22.5w3 fatty acid | microgramsPerMilligramOfTissue |  |
| C22.6w3 | Concentration of C22.6w3 fatty acid | microgramsPerMilligramOfTissue |  |
| C24.0 | Concentration of C24.0 fatty acid | microgramsPerMilligramOfTissue |  |

**Table name:** Macroinvertebrate species counts (invertebrates.csv)

**Table description:** This .csv contains benthic macroinvertebrate abundance data. A full description of the methods can be found in the associated manuscript (Meyer et al., XXXX).

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or  code explanation or date format | Missing value code |
| site | Locational identifier for each sampling site |  |  |
| replicate | Invertebrate replicate number for a given sampling site |  |  |
| Acroloxidae | Mollusc genus | NumerOfIndividuals |  |
| Asellidae | Endemic isopod genus | NumerOfIndividuals |  |
| Baicaliidae | Mollusc genus, most of which are endemic | NumerOfIndividuals |  |
| Benedictidate | Mollusc genus, most of which are endemic | NumerOfIndividuals |  |
| Brandtia\_latissima | Endemic amphipod species | NumerOfIndividuals |  |
| Brandtia\_parasitica\_parasitica | Endemic amphipod species | NumerOfIndividuals |  |
| Caddisflies | General grouping; were not identified to species. | NumerOfIndividuals |  |
| Cryptoropus\_inflatus | Endemic amphipod species | NumerOfIndividuals |  |
| Cryptoropus\_pachytus | Endemic amphipod species | NumerOfIndividuals |  |
| Cryptoropus\_rugosus | Endemic amphipod species | NumerOfIndividuals |  |
| Eulimnogammarus\_capreolus | Endemic amphipod species | NumerOfIndividuals |  |
| Eulimnogammarus\_cruentes | Endemic amphipod species | NumerOfIndividuals |  |
| Eulimnogammarus\_cyaneus | Endemic amphipod species | NumerOfIndividuals |  |
| Eulimnogammarus\_grandimanus | Endemic amphipod species | NumerOfIndividuals |  |
| Eulimnogammarus\_juveniles | Endemic amphipod genus. Identifying to species introduced risk of misclassification. | NumerOfIndividuals |  |
| Eulimnogammarus\_maackii | Endemic amphipod species | NumerOfIndividuals |  |
| Eulimnogammarus\_marituji | Endemic amphipod species | NumerOfIndividuals |  |
| Eulimnogammarus\_verucossus | Endemic amphipod species | NumerOfIndividuals |  |
| Eulimnogammarus\_viridis\_viridis | Endemic amphipod species | NumerOfIndividuals |  |
| Eulimnogammarus\_vittatus | Endemic amphipod species | NumerOfIndividuals |  |
| Flatworms | Not identified beyond order. | NumerOfIndividuals |  |
| Leeches | Not identified beyond order, although 12 endemic species do exist. | NumerOfIndividuals |  |
| Maackia | Mollusc genus, most of which are endemic | NumerOfIndividuals |  |
| Pallasea\_brandtia\_brandita | Endemic amphipod species | NumerOfIndividuals |  |
| Pallasea\_brandtii\_tenera | Endemic amphipod species | NumerOfIndividuals |  |
| Pallasea\_cancelloides | Endemic amphipod species | NumerOfIndividuals |  |
| Pallasea\_cancellus | Endemic amphipod species | NumerOfIndividuals |  |
| Pallasea\_viridis | Endemic amphipod species | NumerOfIndividuals |  |
| Planorbidae | Mollusc genus, most of which are endemic | NumerOfIndividuals |  |
| Poekilogammarus\_crassimus | Endemic amphipod species | NumerOfIndividuals |  |
| Poekilogammarus\_ephippiatus | Endemic amphipod species | NumerOfIndividuals |  |
| Poekilogammarus\_juveniles | Endemic amphipod genus. Identifying to species introduced risk of misclassification. | NumerOfIndividuals |  |
| Poekilogammarus\_megonychus\_perpolitus | Endemic amphipod species | NumerOfIndividuals |  |
| Poekilogammarus\_pictus | Endemic amphipod species | NumerOfIndividuals |  |
| Valvatidae | Mollusc genus, most of which are endemic | NumerOfIndividuals |  |

**Table name:** site-associated metadata (metadata.csv)

**Table description:** This .csv contains metadata for each of the sampled locations. A full description of the methods can be found in the associated manuscript (Meyer et al., XXXX).

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or  code explanation or date format | Missing value code |
| site | Locational identifier for each sampling site |  |  |
| year | Year sampling occurred | Year |  |
| month | Month sampling occurred | Month |  |
| day | Day sampling occurred | Day |  |
| time | Time sampling occurred | Hours:Minutes | NA |
| lat | Latitude of sampling location | DecimalDegrees |  |
| long | Longitude of sampling location | DecimalDegrees |  |
| site\_description | Researcher description of sampling location |  |  |
| distance\_to\_shore\_m | Distance from in situ sampled location to the shoreline | Meters | NA |
| depth\_m | Depth at in situ sampling location | Meters |  |
| air\_temp\_celsius | Temperature of air at sampling location | Celsius |  |
| surface\_temp\_celsius | Temperature of water’s surface at sampling location | Celsius | NA |
| mid\_temp\_celsius | Temperature of water midway between surface and bottom at sampling location | Celsius | NA |
| bottom\_temp\_celsius | Temperature of water near sediment at sampling location | Celsius | NA |
| comments | Notes in the field describing sampling conditions |  |  |
| shore\_photo | Whether or not photos of the shoreline were taken | YesOrNo |  |
| substrate\_photo | Whether or not photos of the substrate were taken | YesOrNo |  |
| sponges | Whether or not sponges were present at a sampling location | YesOrNo |  |
| brandtia | Whether or not brandtia (endemic amphipod) was present at a sampling location | YesOrNo |  |

**Table name:** Inverse-distance-weighted population (distance\_weighted\_population\_metrics.csv)

**Table description:** This .csv contains populaton data for each of the sampled locations. Although the majority of sites do not contain adjacent developments, we calculated inverse-distance-weighted population for each location based on neighboring settlements. A full description of the methods can be found in the associated manuscript (Meyer et al., XXXX).

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or  code explanation or date format | Missing value code |
| site | Locational identifier for each sampling site |  |  |
| replicate | Replicate for a given location | Numerical values represent a replicate; the letter “C” represents a control. |  |
| fragments | Number of microplastic fragments observed | Number |  |
| fibers | Number of microplastic fibers observed | Number |  |
| beads | Number of microplastic beads observed | Number |  |
| comments | Observer comments while enumerating microplastics |  |  |
| volume\_filtered\_ml | Volume for a given replicate filtered | Milliliters |  |

**Table name:** Nutrient data (nutrients.csv)

**Table description:** This .csv contains nutrient data for each of the associated sampling locations. A full description of the methods can be found in the associated manuscript (Meyer et al., XXXX).

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or  code explanation or date format | Missing value code |
| site | Locational identifier for each sampling site |  |  |
| replicate | Replicate for a given location |  |  |
| nh4\_mg\_dm3 | Ammonium concentration | MilligramsPerCubicDecimeter |  |
| no3\_mg\_dm3 | Nitrate concentration | MilligramsPerCubicDecimeter |  |
| tp\_mg\_dm3 | Total phosphorus concentration | MilligramsPerCubicDecimeter |  |
| tpo43\_mg\_dm3 | Total phosphate concentration | MilligramsPerCubicDecimeter |  |

**Table name:** Periphyton abundance data (periphyton.csv)

**Table description:** This .csv contains periphyton abundance data for each of the sampled littoral sites. A full description of the methods can be found in the associated manuscript (Meyer et al., XXXX).

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or  code explanation or date format | Missing value code |
| site | Locational identifier for each sampling site |  |  |
| replicate | Replicate number for a given sampling site |  |  |
| subsamples\_counted | Number of 10 microliter subsamples counted for a given replicate |  | NA |
| diatom | Number of diatom cells counted for a given replicate | NumberOfCells | NA |
| spirogyra | Number spirogyra cells counted for a given replicate | NumberOfCells | NA |
| spirogyra\_filaments | Number of spirogyra filaments counted for a given replicate | NumberOfFilaments | NA |
| ulothrix | Number of ulothrix cells counted for a given replicate | NumberOfCells | NA |
| ulothrix\_filaments | Number of ulothrix filaments counted for a given replicate | NumberOfFilaments | NA |
| tetrasporales | Number of tetrasporales cells counted for a given replicate | NumberOfCells | NA |
| pediastrum | Number of pediastrum cells counted for a given replicate | NumberOfCells | NA |
| desmidales | Number of desmidales cells counted for a given replicate | NumberOfCells | NA |
| lyngbya | Number of lyngbya cells counted for a given replicate | NumberOfCells | NA |
| comments | Notes from the obsverer |  | NA |

**Table name:** Pharmaceutical and Personal Care Product (PPCP) data (ppcp.csv)

**Table description:** This .csv contains PPCP concentrations for each of the associated sampling locations. A full description of the methods can be found in the associated manuscript (Meyer et al., XXXX).

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or  code explanation or date format | Missing value code |
| site | Locational identifier for each sampling site |  |  |
| paraxanthine | Concentration of paraxanthine, also known as 1,7-dimethylxanthine. Paraxanthine is the main human metabolite of caffeine | NanogramsPerLiter |  |
| acetaminophen | Concentration of acetaminophen, also known as paracetamol. | NanogramsPerLiter |  |
| amphetamine | Concentration of amphetamine. | NanogramsPerLiter |  |
| caffeine | Concentration of caffeine | NanogramsPerLiter |  |
| carbamezapine | Concentration of carbamezapine | NanogramsPerLiter |  |
| cimetidine | Concentration of cimetidine | NanogramsPerLiter |  |
| cotinine | Concentration of cotinine, which is the main human metabolite of nicotine | NanogramsPerLiter |  |
| diphenhydramine | Concentration of diphenhydramine | NanogramsPerLiter |  |
| mda | Concentration of mda | NanogramsPerLiter |  |
| mdma | Concentration of mdma | NanogramsPerLiter |  |
| methamphetamine | Concentration of methamphetamine | NanogramsPerLiter |  |
| morphine | Concentration of morphine | NanogramsPerLiter |  |
| phenazone | Concentration of phenazone | NanogramsPerLiter |  |
| sulfachloropyridazine | Concentration of sulfachloropyridazine | NanogramsPerLiter |  |
| sulfamethoxazole | Concentration of sulfamethoxazole | NanogramsPerLiter |  |
| thiabendazole | Concentration of thiabendazole | NanogramsPerLiter |  |
| trimethoprim | Concentration of trimethoprim | NanogramsPerLiter |  |
| collection\_year | Year sample was collected in the field | Year |  |
| collection\_month | Month sample was collected in the field | Month |  |
| collection\_day | Day sample was collected in the field | Day |  |
| analysis\_year | Year sample was analyzed | Year |  |
| analysis\_month | Month sample was analyzed | Month |  |
| analysis\_day | Day sample was analyzed | Day |  |

**Table name:** Stable Isotopes data (stable\_isotopes.csv)

**Table description:** This .csv contains carbon and nitrogen stable isotope data within periphyton and macroinvertebrate tissue for each of the associated sampling locations. A full description of the methods can be found in the associated manuscript (Meyer et al., XXXX).

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or  code explanation or date format | Missing value code |
| site | Locational identifier for each sampling site |  |  |
| C13 | Carbon 13 percentages | Percent |  |
| N15 | Nitrogen 15 percentages | Percent |  |
| Genus | Amphipod genus or Periphyton |  |  |
| Species | Amphipod species |  | NA |

**Table name:** Total Lipid data (total\_lipid.csv)

**Table description:** This .csv contains lipid gravimetry data for periphyton and benthic macroinvertebrate tissue for each of the associated sampling locations. A full description of the methods can be found in the associated manuscript (Meyer et al., XXXX).

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or  code explanation or date format | Missing value code |
| site | Locational identifier for each sampling site |  |  |
| Genus | Taxonomic Genus for amphipods and Drapa (an endemic alga). Periphyton and snails were not identified genus, and are listed just as “periphyton” or “snails” |  |  |
| Species | Amphipod species |  | NA |
| total\_lipid\_mg\_per\_g | Total amount of lipids in a sample | MilligramsOfLipidPerGramOfTissue |  |
| deviation | Samples were weighed three times and deviation was calculated in measurement | MilligramsOfLipidPerGramOfTissue |  |

## Scripts/code (software)

(List any software scripts/code you would like to archive along with your data. These may include processing scripts you wrote to create, clean, or analyze the data.)

|  |  |  |
| --- | --- | --- |
| File name | Description | Scripting language |
| 00\_disaggregated\_data\_cleaning.R | This script cleans misspellings, incongruencies, and data poor quality from the raw data outputs. Although data are kept disaggregated (i.e., as replicates). This step largely removes data with poor quality and enables data to be interoperable with all other CSVs. | R |
| 01\_data\_cleaning.R | This script aggregates data among replicates for a given sampling location. It outputs aggregated CSVs that are used for successive analytics scripts. | R |
| 02\_sewage\_indicator\_analysis.R | This script takes outputs from script 01 ti relate each sewage indicator with inverse distance weighted population. It outputs a plot with aggregated model results and plotted data for each sewage indicator. | R |
| 03\_community\_composition\_analysis.R | This script takes outputs from script 01 to relate sewage indicators and inverse distance weighted population with changes in benthic periphyton and macroinvertebrate community composition. | R |
| 04\_fatty\_acid\_analysis.R | This script takes outputs from script 01 to analyze benthic algae and macroinvertebrate fatty acid compositions, and then relates those fatty acid profiles with sewage indicators. | R |
| 05\_table\_formatting.R | This script formats descriptive tables using the outputs from script 01. | R |
| 06\_map\_making.R | This script generates a map of the sampled locations. | R |
| panel\_cor\_function.R | This script is sourced to analyze cross-correlation among variables. | R |

1. This document liberally borrows from similar documents at SBC and GCE [↑](#footnote-ref-1)