**Data Processing Workflow for Flume Experiments**

**(High Level Overview)**

*Note: for a technical description of critical value descriptions and script function, refer to Biogeomorphic Flume/Code/* [*readme.txt*](https://drive.google.com/open?id=15d1vbtHc_BCW6GVShvhgQ8w6UBYlF-aZ) *and* [*Flume Parameter Calculations*](https://drive.google.com/open?id=1ZedYJQsXx7dSClzeSQ7QIN7DNFF_F8Hg)

**Objective:** to obtain estimates for critical values, including particle removal rate due to direct capture, particle removal rate due to settling and effective capture efficiency. See [flume parameter estimates](https://docs.google.com/spreadsheets/d/13O139CTIYFpqPZT72yevfwmSKKIRY_8eavzKa2MGWuk/edit?usp=sharing) for a full list of the values to calculate for each experiment.

A typical flume experiment collects the following forms of data:

1. LISST data (time series of particle size distributions)
2. Peristaltic pump data (time series of mass concentrations of sediment)
3. Sediment trap data (mass of sediment settled over the course of the experiment)
4. Vectrino data (time series of velocity point measurements)

Overview of data processing steps

**1) Process LISST data.**

1. Refer to Biogeomorphic Flume / SOP for Flume Experiments section 4.2.2: Offloading LISST Data for instructions on obtaining .asc formatted data for the following steps.
2. LISST data in .asc format does not need to be manually cleaned; the scripts can all be run (in any order) using the raw .mat datafile exported by the LISST software as the input data file.
3. Run all scripts located in [Biogeomorphic Flume/Code](https://drive.google.com/open?id=1ZKrcZYTQV1RxpQw-QAhj6OVxtas5iAnf) and save all output plots as images. See table below for details on which scripts provide which outputs.
4. Put images in drive under [Biogeomorphic Flume/Experiment plots](https://drive.google.com/open?id=1KrJ9F9VAadtqD4Go6tLGLxjFiSxM21ly) and record the parameter estimates in [flume parameter estimates](https://docs.google.com/spreadsheets/d/13O139CTIYFpqPZT72yevfwmSKKIRY_8eavzKa2MGWuk/edit?usp=sharing)

**LISST data scripts and outputs**

|  |  |  |  |
| --- | --- | --- | --- |
| **Script** | **Output Parameter(s)** | **Corresponding variable name(s) in script** | **Output Figure(s)** |
| LISST\_processing.R | n/a | n/a | Cumulative Density Curve |
| bin\_visualization.R | n/a | n/a | Volume Concentration by Bin |
| k\_calculation.R | k (total particle removal rate, bin-specific), effective capture efficiency | K\_exp, ece | n/a |
| k\_calculation\_all\_bins.R | k (total particle removal rate, mean across all bins), effective capture efficiency | K\_exp, ece | n/a |
| particle\_size\_diagnostics.R | n/a | n/a | Weighted Average Particle Size, Empirical CDF, Total Particle Concentration |
| volume\_weighted\_mean\_particle\_size.m | Mean volume concentration, volume-weighted mean particle diameter | Mean\_vol\_conc, volume\_weighted\_particle\_d | n/a |

**2) Process peristaltic pump and sediment trap data.**

1. This can only be done after all pump samples and all sediment traps have been filtered and weighed. The procedure for filtering pumped samples is in [SOP for Flume Experiments](https://docs.google.com/document/d/1NOGzMRS78mayP7-_7fn5sZgcszuzQ5H8rPIo4Mu0DUs/edit?usp=sharing).
2. Use the scripts in [Biogeomorphic Flume/Code](https://drive.google.com/open?id=1ZKrcZYTQV1RxpQw-QAhj6OVxtas5iAnf) and the associated readme file, which notes formatting conventions (data must be fed to the scripts as .CSV files)
3. All pump/trap data is located in [Biogeomorphic Flume/Flume Parameter Estimates](https://docs.google.com/spreadsheets/d/1_V4tcwvWRQOZSJCkzGPVfZoaz5Ro9Crnd1bn8nHgq2E/edit?usp=sharing) and formatted .CSV data is [here](https://drive.google.com/open?id=1LUAM520Qrx5yKSbdy1-0qK3JZ6-WLfPg).

|  |  |  |  |
| --- | --- | --- | --- |
| **Script** | **Output Parameter(s)** | **Corresponding variable name(s) in script** | **Output Figure(s)** |
| peristaltic\_pump\_plotter.R | n/a | n/a | Mass concentration time series |
| peristaltic\_pump\_k\_estimation.R | k | k\_final | Mass concentration time series by point location with fitted exponential model |

**3)** Run Two-Station Method Calculation

1. See Laurel’s theoretical calculations [here](https://drive.google.com/open?id=1l-jeBzO3rFvOW2KKmNczI9ceQgn33Eyq), which are implemented in Matlab in a script called two\_station\_method.m. Running it requires both peristaltic pump and LISST data.
2. The script’s output is a time series of concentration removed by direct capture in CSV format.

**4)** Process Vectrino data.

*Note: Vectrino data is not currently used to obtain any critical parameter estimates; its main purpose thus far in the flume experiments has been to gauge flow velocity at an upstream point location.*

1. Vectrino data needs to have a threshold despiking algorithm applied to the raw data; this is called “Vectrino\_processing\_lowSNR\_RA\_v5\_VPro\_021418.m.” This script outputs a separate, processed data file of the same name with the suffix “\_processed”, and that file should be used for all analyses and applications.

* The relevant scripts for processing vectrino data for flume applications are also in [Biogeomorphic Flume/Code](https://drive.google.com/open?id=1ZKrcZYTQV1RxpQw-QAhj6OVxtas5iAnf). For processing vectrino data for other applications, see full code repository at <https://gitlab.com/ckeating/vectrino-data-analysis>