**Metadata[[1]](#footnote-1) for Peters et al.**

**Table 1.** Description of the dataset.

|  |  |
| --- | --- |
| **Title of dataset** | *Data for “Mesozooplankton grazing exerts a significant control on cyanobacterial harmful algal bloom (cHAB) growth in a large eutrophic lake"* |
| **URL of dataset** | *Provide the URL of the data repository for the dataset. For the manuscript review stage, it is acceptable to say that this is forthcoming upon decision at the first review stage* |
| **Abstract** | *This study examined mesozooplankton grazing rates and preferences during a cyanobacterial harmful algal bloom (cHAB) in western Lake Erie. Water and zooplankton were collected from one site in western Lake Erie during August and September 2021. Grazing assays were performed and measurements for chl-a, major phytoplankton taxa (Fluoroprobe) and pico- and nanoplankton were collected. Data was analyzed to determine grazing rates and selectivity on the measured food sources. Ambient zooplankton and phytoplankton composition and biomass at the collection site was estimated from a long-term monitoring program (Lake Erie Plankton Abundance Study-- LEPAS).* |
| **Keywords** | zooplankton, grazing experiment, cyanobacteria, nanoplankton, picoplankton, Lake Erie, The Ohio State University Aquatic Ecology Laboratory |
| **Lead author for the dataset** | *Lyndsie Michele Collis* |
| **Title and position of lead author** | *Research Biologist* |
| **Organization and address of lead author** | *State University of New York Research Foundation (SUNY RF)* |
|  |
| *SUNY Oswego* |
|  |
| *United States Geological Survey* |
|  |
| *17 Lake Street* |
|  |
| *Oswego, NY 13126* |
|  |
| **Email address of lead author** | [lyndsie.collis@gmail.com](mailto:lyndsie.collis@gmail.com) |
| **Additional authors or contributors to the dataset** | *D.I. Peters, M.D. Shaw, Z.J. Slagle, H.A. Vanderploeg, J.M. Hood* |
| **Organization associated with the data** | *The Ohio State University* |
| **Funding** | Federal Aid in Sport Fish Restoration Program (F-69-P, Fish Management in Ohio) administered jointly by the United States Fish and Wildlife Service and the ODNR-DOW (project FADX09 to JMH and S. Ludsin). |
|  |
| United States Environmental Protection Agency through an assistance agreement with the Ohio Environmental Protection Agency (OSU-FDERIE19 and OSU-FDERIE22 to JMH) |
|  |
| Ohio Water Resources Research Institute 104(b) Program Grant |
| **License** | [CCBY – requires attribution](https://creativecommons.org/licenses/by/4.0/) |
| **Geographic location – verbal description** | *Western basin of Lake Erie, USA* |
| **Geographic coverage bounding coordinates** | *41.76800 N, -83.30116 W* |
| **Time frame - Begin date** | *August 2021* |
| **Time frame - End date** | *September 2021* |
| **General study design** | Live zooplankton and Lake Erie water used in the experiments, as well as zooplankton and phytoplankton data for analysis were provided by the Lake Erie Plankton Abundance Study (LEPAS), which has monitored plankton and physicochemical variables biweekly at eight sites in the western basin of Lake Erie during May–September since 1995. |
|  |
| Mesozooplankton grazing experiments were conducted via grazer-gradient assays at The Ohio State University's Aquatic Ecology Laboratory. |
| **Methods description** | The methods for zooplankton and phytoplankton biomass quantification are described in “00c\_LEPAS Methods\_20221118.pdf”, which is in the data repository. |
|  |
| Mesozooplankton grazing assays were conducted according to Lehman and Sandgren 1985 and Hambright et al., 2007, during which mesozooplankton were incubated at increasing densities in 200 µm filtered lake water containing phytoplankton and microzooplankton at ambient densities. For a full description of the methods, see "Methods" section of the manuscript. |
| **Laboratory, field, or other analytical methods** | Methods to obtain zooplankton and phytoplankton biomass are described extensively in “00c\_LEPAS Methods\_20221118.pdf” which is available in the data repository. Chl-a measurements were obtained via acetone extraction and measurement on a Trilogy Fluorometer. Phytoplankton taxa during the experiments were measured via Fluoroprobe measurements and validated via light microscopy. Pico- and nanoplankton biomass was estimated via epifluorescence microscophy (Hadas and Berman 1998). |
| **Taxonomic species or groups** | *Zooplankton: Bosmina longirostris, Daphnia mendotae, Diaphanosoma sp., Eubosmina coregoni, Leptodiaptomus siciloides, Mesocyclops edax; Phytoplankton: Aulacoseira sp., Carteria sp., Chlamydomonas sp., Chroomonas sp., Cryptomonas sp., Cyclotella sp., Eudorina sp., Microcystis sp., Rhodomonas sp.* |
| **Quality control** | Plankton data are validated (i.e., checked for completeness, correctness, and errors) at several steps. 1) When samples are brought back to the laboratory, we check to ensure that the sample is preserved and the label is intact and accurate. Data entered into the sample inventory is also checked for completeness. 2) When the count data is transferred from a paper datasheet to an electronic datasheet, the information is checked for anomalous data or sample information, and for entry errors. 3) When data is entered into the LEPAS database and biomass and densities are calculated, the data is checked for completeness, duplicates, missing information, and to make sure it corresponds to previous data. |
|  |
| Five percent of samples are recounted for QC. Recounts are conducted by a different research assistant and involve recounting a subsample without looking at the data from the first sample. |
|  |
| Prior to analyses, all data were screened for reasonableness and outliers. |
| **Additional information** | *N/A* |

**Table 2.** Data dictionary: description of the variables (i.e., columns) in EACH dataset.

***Data used to conduct these analyses are provided, however, The Lake Erie Plankton Abundance Study (LEPAS) database is not publicly available due to the sponsors restrictions, although data is available by request.***

***Only input data files are described in Table 2-- descriptions of data generated from the R scripts are not included, but the R scripts and output data files are included in the data submission.***

Dataset filename: *00\_ChlFluoro\_WBLE\_2021\_SampleInfo.csv*

Dataset description: Sample information for the experiments—used to calculate mesozooplankton grazing rates.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| InitialTime | Time of measurement at the beginning of the experiment |  |  | character; mm/dd/yyyy H:M (EST) |  |
| FinalTime | Time of measurement at the end of the experiment |  |  | character; mm/dd/yyyy H:M (EST) |  |
| Time | Experiment event |  | EarlyAug = Early August; EarlySep = Early September | character |  |
| Treatment | Zooplankton treatment |  | 0 = No zooplankton; 1 = Low zooplankton; 2 = Medium zooplankton; 4 = High zooplankton | integer |  |
| Trial | Experiment replicate |  |  | character |  |
| Density\_numL | Density of zooplankton in the treatment | Number of individuals per liter |  | integer |  |

Dataset filename: *01\_2021\_WBLE\_ZP\_grazingZNR\_chl\_wide.csv*

Dataset description: Chl-a measurements from the zooplankton grazing experiments. R script used to generate these data is available upon request.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Event | Sampling event for each experiment |  |  | character |  |
| SampleDate | Date of each experiment |  |  | character; YYYY-mm-dd |  |
| Type | Type of experiment run |  | MESO = mesozooplankton grazing assay; MICRO = microzooplankton grazing assay | character |  |
| Treatment | Zooplankton treatment |  | 0 = No zooplankton; 1 = Low zooplankton; 2 = Medium zooplankton; 4 = High zooplankton | numeric |  |
| Rep | Experiment replicate |  |  | integer |  |
| ZP\_Dens\_num\_L | Denisty of zooplankton in the treatment | Number of individuals per liter |  | numeric |  |
| Initial\_chl\_ugL | Chl-a value in the bottle at the beginning of the experiment | µg / L |  | numeric |  |
| Final\_chl\_ugL | Chl-a value in the bottle at the end of the experiment | µg / L |  | numeric |  |

Dataset filename: *WBLE\_2021\_ZPgrazingexps\_ZPdens.csv*

Dataset description: The average mesozooplankton density in each experiment- calculated from the average of the mesozooplankton densities in each of the mesozooplankton treatments.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Time | Experiment event |  | EarlyAug = Early August; EarlySep = Early September | character |  |
| ZP\_dens\_exp\_num\_L | Denisty of zooplankton in the treatment | Number of individuals per liter |  | numeric |  |

Dataset filename: *00\_LEPAS\_WBLE\_2021\_36873\_FieldExp\_BMS\_ugCL.csv*

Dataset description: Mesozooplankton biomass from the field site (36-873) during the experiments and average density and biomass of mesozooplankton in the experimental treatments.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Date | Date of experiment |  |  | character; YYYY-mm-dd |  |
| Site | Site water was collected for experiment |  |  | character |  |
| Zp\_dens | Mesozooplankton density at site | Number of individuals per liter |  | numeric |  |
| Zp\_bms | Mesozooplankton biomass at site | µg dry weight per liter |  | numeric |  |
| Zp\_bms\_ugC\_L | Mesozooplankton biomass at site, in carbon units | µg Carbon per liter |  | numeric |  |
| Month2 | Month experiment was conducted at |  |  | character | NA |
| Time | Experiment event |  | EarlyAug = Early August; EarlySep = Early September | character | NA |
| Ind\_wt\_ugC | Average weight of an individual zooplankton duirng each experiment | µg Carbon per liter |  | numeric |  |
| ZP\_dens\_exp\_num\_L | Denisty of zooplankton in the treatment | Number of individuals per liter |  | numeric | NA |
| Zp\_bms\_exp\_ugC\_L | Biomass of zooplankton in the treatment | µg Carbon per liter |  | numeric | NA |

Dataset filenames: *NanoEAugCount.csv, NanoESepCount.csv, PicoEAugCount.csv, PicoESepCount.csv, NanoPico\_InitialCollection\_Count.csv (Note: all of these datasets have the same format)*

Dataset description: Raw data of picoplankton (0.2 - < 2µm) and nanoplankton (2-20 µm) counts from the August and September experiments. Pico- and nanoplankton counts were conducted by Dan Peters via epifluorescence microscopy.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Name | Sample ID of particle measured: format is "ZPtreatment\_Replicate\_Initial/Final\_Meso\_Event\_VolFiltered\_FilterUsed\_Dapi/Chl\_ImageNumber |  | 0 = No zooplankton; 1 = Low zooplankton; 2 = Medium zooplankton; 4 = High zooplankton | character |  |
| Object | Particle number |  |  | integer |  |
| Area | Area of the object | µm2 |  | numeric |  |
| EqDiameter | A size feature derived from the area. It determines the diameter of a circle with the same area as the measured object. | µm |  | numeric |  |
| Perimeter | Total boundary length of the object. Calculated from four projections in the directions 0, 45, 90, and 135 degrees using Crofton's formula. | µm |  | numeric |  |
| Length | Length is a derived feature appropriate for elongated or thin structures. As based on the rod model, length is calculated as: Perimeter + sqrt(Perimeter^2 - 16 x Area) / 4 | µm |  | numeric |  |
| Width | Width is a derived feature appropriate for elongated or thin structures. It is based on the rod model and calculated as: Width = Area / Length | µm |  | numeric |  |
| MaxFeret | Maximal value of the set of Feret's diameters. | µm |  | numeric |  |
| MinFeret | Minimal value of the set of Feret's diameters. | µm |  | numeric |  |
| Circularity | Circularity equals 1 only for circles-- all other shpaes are characterized by a value smaller than 1. It is a derived shape measure, calculated as: 4 x pi x Area / Perimeter^2 |  |  | numeric |  |
| Elongation | Characterizes the object shape. It is calculated as: MaxFeret / MinFeret |  |  | numeric |  |
| MeanIntensity | Value is derivd from the intensity histogram. It is the arithmetic mean of the pixel intensities. | lux |  | numeric |  |
| SumIntensity | Sum of intensity in every pixel of the object | lux |  | integer |  |
| Unit | Measurement unit of length/width | µm |  | character |  |
| Channel | Filter used during imaging |  |  | character |  |

Dataset filename: *00\_NanoPico\_WBLE\_2021\_SampleInfo.csv*

Dataset description: Sample information from the grazing experiments and epifluorescence microscopy counts, used to calculate biovolume.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Time | Experiment event |  | EarlyAug = Early August; EarlySep = Early September | character |  |
| Size | Type of count, either "Nano" or "Pico" |  |  | character |  |
| Treatment | Zooplankton treatment |  | 0 = No zooplankton; 1 = Low zooplankton; 2 = Medium zooplankton; 4 = High zooplankton | numeric |  |
| Trial | Experiment replicate |  |  | character |  |
| Volume\_mL | Amount of sample filtered for the microscopy count | mL |  | integer |  |

Dataset filename: *00\_ZP\_grazing\_field\_chl\_WBLE\_2021.csv*

Dataset description: Chl-a measurements collected from the field site during the zooplankton grazing experiments.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Time | Experiment event |  | EarlyAug = Early August; EarlySep = Early September | character |  |
| insitu\_chla\_ugL | Chlorophyll-a (Chl-a) measurement at the field site during the grazing experiments, fluorescence units | ug chl-a / L |  | numeric |  |
| insitu\_chla\_ugC\_L | Chlorophyll-a (Chl-a) measurement at the field site during the grazing experiments, Carbon units | ug Carbon / L |  | numeric |  |

Dataset filename: *03\_fluoroprobe\_exp\_phytogrps\_WBLE2021.csv*

Dataset description: Fluoroprobe data from the zooplankton grazing experiments, which measured the biomass of Cyanobacteria, Cryptophyta, Bacillariophyta, and Chlorophyta at the beginning and end of the experiments. R script used to generate these data is available upon request.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Event | Experiment event |  |  | character |  |
| SampleDate | Date of experiment |  |  | character; YYYY-mm-dd |  |
| Type | Type of grazing experiment |  | MESO = mesozooplankton; MICRO = microzooplankton | character |  |
| Treatment | Zooplankton treatment |  | 0 = No zooplankton; 1 = Low zooplankton; 2 = Medium zooplankton; 4 = High zooplankton | numeric |  |
| Rep | Experiment replicate |  |  | integer |  |
| ZP\_Dens\_num\_L | Mesozooplankton density in the treatment | Number of individuals per liter |  | integer | NA |
| Phyto\_grp | Phytoplankton taxa measured |  |  | character |  |
| Final\_chl\_ugL | Chl-a measured at the end of the experiment | µg / L |  | numeric | NA |
| Initial\_chl\_ugL | Chl-a measured at the beginning of the experiment | µg / L |  | numeric | NA |

Dataset filename: *00\_fluoroprobe\_EarlyAugSept\_field\_WBLE2021.csv*

Dataset description: Fluoroprobe measurements collected from the field site during the zooplankton grazing experiments.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Time | Experiment event |  | EarlyAug = Early August; EarlySep = Early September | character |  |
| Treatment | Lake Erie water treatment for Fluoroprobe measurements: Either "whole" water or sieved through a 200 µm mesh |  |  | character |  |
| Phyto\_grp | Phytoplankton taxa measured |  |  | character |  |
| Avg\_ugL | Chl-a measured at the end of the experiment | µg / L |  | numeric |  |
| C\_conv | Carbon:Chl-a conversion ratio, from Yacobi and Zohary (2010). |  |  | integer | NA |

Dataset filename: *00\_fluoroprobe\_Cconversions.csv*

Dataset description: Carbon conversion factors for the Fluoroprobe groups (Cyanobacteria, Cryptophyta/Bacillariophyta, Chlorophyta), from Yacobi and Zohary (2010).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Phyto\_grp | Phytoplankton taxa measured |  |  | character |  |
| C\_conv | Carbon:Chl-a conversion ratio, from Yacobi and Zohary (2010). |  |  | integer | NA |

Dataset filename: *00\_lake\_erie\_habs\_field\_sampling\_results\_2012\_2018\_v2.csv*

Dataset description: Physical, chemical, and biological water quality data collected from a small boat in western Lake Erie, Great Lakes from 2012-05-15 to 2018-10-09. Accessed via https://www.ncei.noaa.gov/archive/accession/0187718 on 10/01/2020.

Dataset filename: *00\_lake\_erie\_habs\_field\_sampling\_results\_2019.csv*

Dataset description: Physical, chemical, and biological water quality data collected from a small boat in western Lake Erie, Great Lakes from 2019-04-30 to 2019-10-07. Accessed via https://www.ncei.noaa.gov/archive/accession/0209116 on 10/01/2020.

Metadata for these datasets can be found via the above links.

Dataset filename: *InitialCollection\_2021\_ZPexps\_rawdata4.csv*

Dataset description: Fluoroprobe data collected from the field site during two experiments. One experiment (6/29/21) was conducted during low/no cHABs, and the other experiment (10/4/21) was conducted during the cHAB event, which were used to make Figure S3.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Date\_Time | Date and Time of Fluoroprobe measurement |  |  | character; mm/dd/YYY Hour:Minutes (EST) |  |
| Event | Experiment event |  |  | character |  |
| Sample | Time sample was measured during the experiment |  |  | character |  |
| Type | Lake Erie water treatment for Fluoroprobe measurements: Either "whole" water or sieved through a 200 µm mesh |  |  | character |  |
| Rep | Experiment replicate |  |  | integer |  |
| Chlorophyta | Chl-a measurement of Chlorophyta taxa | µg / L |  | numeric |  |
| Cyanobacteria | Chl-a measurement of Cyanobacteria taxa | µg / L |  | numeric |  |
| Bacillariophyta | Chl-a measurement of Bacillariophyta taxa | µg / L |  | numeric |  |
| Cryptophyta | Chl-a measurement of Cryptophyta taxa | µg / L |  | numeric |  |
| YellowSubst | Chl-a measurement of YellowSubst taxa | µg / L |  | numeric |  |

Dataset filename: *03b\_LEPAS\_phyto\_biomass\_09012024.csv*

Dataset description: Phytoplankton biomass of each genus, collected and enumerated via LEPAS protocols. The lengths, widths, and diameters were calculated from the first 20 lengths of each genus. The script used to generate these data is not available, as it extracts directly from the LEPAS database.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Sample\_ID | LEPAS Sample ID-- format is "Date-Site" |  |  | character |  |
| SampleDate | Date of collection |  |  | character; mm/dd/YYY |  |
| Genus\_type | Taxonomic Genus of the identified phytoplankton |  |  | character |  |
| Measured\_number | Individual measurement number for each genus. |  |  | character |  |
| Length\_um | Length of individual phytoplankton. Measured for elongated taxa | µm |  | numeric | NA |
| Width1\_um | Width of individual phytoplankton. Measured for elongated taxa | µm |  | numeric | NA |
| Width2\_um | NA | µm |  | logical | NA |
| Diameter\_col\_um | For colonial species: Diameter of the entire colony. | µm |  | numeric | NA |
| Diameter\_cell\_um | Diameter of an individual cell. Measured for circular taxa. | µm |  | numeric | NA |
| Size\_cat | Indicates if a phytoplankton individual is larger or smaller than 200 µm. |  |  | character |  |

Dataset filename: *03b\_LEPAS\_phyto\_RawLengths.csv*

Dataset description: Individual lengths of each phytoplankton genus identified via LEPAS protocols. As per protocols, only the first 20 lengths of each genus are measured per sample.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Units** | **Code explanation** | **Data format** | **Missing data code** |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Sample\_ID | LEPAS Sample ID-- format is "Date-Site" |  |  | character |  |
| SampleDate | Date of collection |  |  | character; mm/dd/YYY |  |
| Genus\_type | Taxonomic Genus of the identified phytoplankton |  |  | character |  |
| Measured\_number | Individual measurement number for each genus. |  |  | character |  |
| Length\_um | Length of individual phytoplankton. Measured for elongated taxa | µm |  | numeric | NA |
| Width1\_um | Width of individual phytoplankton. Measured for elongated taxa | µm |  | numeric | NA |
| Width2\_um | NA | µm |  | logical | NA |
| Diameter\_col\_um | For colonial species: Diameter of the entire colony. | µm |  | numeric | NA |
| Diameter\_cell\_um | Diameter of an individual cell. Measured for circular taxa. | µm |  | numeric | NA |
| Size\_cat | Indicates if a phytoplankton individual is larger or smaller than 200 µm. |  |  | character |  |

Dataset filename: *00\_fluoro\_phyto\_bms\_ugCL\_less200um\_2021\_36873.csv*

Dataset description: phytoplankton biomass less than 200 µm, as measured by the Fluoroprobe. Lake Erie water from site 36-873 was sieved through a 200um sieve for the grazing experiments prior to measurements.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *The name of the variable in the dataset; avoid special characters, dashes and spaces* | *A detailed description of the variable* | *Units the variable is measured in* | *If you use codes in your column, please explain each code, such as: LR = Little Rock Lake; A=sample; etc.* | *State exactly how the data are stored; for dates, state how it is formatted, including time zone, etc.* | *If data are missing, indicate how they are stored, such as NULL, NA, blank cell, etc.* |
| Event | Experiment event |  |  | character |  |
| Treatment | Lake Erie water treatment for Fluoroprobe measurements: Water was sieved through a 200 µm mesh. |  |  | character |  |
| SampleDate | Date of Fluoroprobe measurement |  |  | character; mm/dd/YYYY |  |
| Phyto\_grp | Phytoplankton taxa measured, common name. |  |  | character |  |
| C\_ugL | Chl-a measurement of phytoplankton group | µg C / L |  | numeric |  |
| Phyto\_grp2 | Phytoplankton taxa measured, scientific name. |  |  | character |  |

**Table 3.** Scripts/Code used in the analysis.

***A more detailed description of all scripts as well as the input and output files generated is provided in the “Read Me” file entitled (“00a\_ReadMe”).***

|  |  |  |
| --- | --- | --- |
| File name | Description | Scripting language |
| 01\_WBLE\_2021\_NanoPico\_calcBV.R | Calculates autotrophic and heterotrophic biovolume of nano- and picoplankton from zooplankton grazing experiments. | R |
| 02\_WBLE\_2021\_NanoPico\_calcCR\_boot.R | Calculates the mesozooplankton clearance rate for pico- and nanoplankton, as well as the pico- and nanoplankton growth rate via bootstrapping. | R |
| 03\_InitialCollection\_NanoPico\_calcBV.R | Calculates autotrophic and heterotrophic biovolume of nano- and picoplankton from samples collected from the field site during zooplankton grazing experiments. | R |
| 04\_WBLE\_2021\_NanoPico\_calcGR.R | Calculates the mesozooplankton grazing rate, percent food growth consumed, and percent standing stock consumed for pico- and nanoplankton. | R |
| 05\_WBLE\_2021\_chl\_calcCR\_boot.R | Calculates the mesozooplankton clearance rate for total phytoplankton (chl-a), as well as the total phytoplankton (chl-a) growth rate via bootstrapping. | R |
| 06\_WBLE\_2021\_chl\_calcGR.R | Calculates the mesozooplankton grazing rate, percent food growth consumed, and percent standing stock consumed for total phytoplankton (chl-a). | R |
| 07\_WBLE\_2021\_fluoro\_calcCR\_boot.R | Calculates the mesozooplankton clearance rate for different phytoplankton taxa (Cyanobacteria, Cryptophyta/Bacillariophyta, Chlorophyta; measured via the FluoroProbe) via bootstrapping. Also calculates the growth rate of these phytoplankton taxa via bootstrapping. | R |
| 08\_WBLE\_2021\_fluoro\_calcGR.R | Calculates the mesozooplankton grazing rate, percent food growth consumed, and percent standing stock consumed for different phytoplankton taxa (Cyanobacteria, Cryptophyta/Bacillariophyta, Chlorophyta; measured via the FluoroProbe). | R |
| 09\_WBLE\_2021\_ZPexp\_CRGR\_combinedata.R | Combines grazing data calculated in previous scripts into a single data file. | R |
| 10\_WBLE\_2021\_ZPexp\_CRGR\_Figs4to6.R | Makes Figures 4-6 in manuscript. | R |
| 12\_NOAA\_HABcruise\_fieldsampling\_PAR\_cyanobms\_forMS.R | Uses publicly available data from NOAA’s Great Lakes Environmental Research Laboratory (NOAA-GLERL) to create Figure S2. Also calculates Photosynthetically Active Radiation (PAR) in western Lake Erie to determine the PAR level for the grazing experiments. | R |
| 13\_SuppFig3\_fluoroprobetime.R | Creates Supplementary Figure 3—Fluoroprobe data by major group (Cyanobacteria, Cryptophyta, Bacillariophyta, Chlorophyta), measurements are taken every three seconds by the Fluoroprobe. | R |
| 14\_WBLE\_2021\_chla\_field\_Fig2b.R | Creates Figure 2b—Chl-a measurements collected from the field during the mesozooplankton grazing experiments. | R |
| 15a\_LEPAS\_phyto\_lengths\_NanoPicoMS\_Fig2a\_FigS5.R | Calculates phytoplankton biomass <200 µm, collected from the Lake Erie Plankton Abundance Study (LEPAS) for script 15b (which creates Figure 2a). Also creates Figure S5, which shows phytoplankton biomass by genus for phytoplankton less than 200 µm. | R |
| 15b\_LEPASphyto\_fluoro\_less200um\_combine\_graphs.R | Creates Figure 2a—biomass of phytoplankton groups estimated via the Fluoroprobe and light microscopy. | R |

1. *This metadata template was adapted from the L&O Letters metadata template, which* *liberally borrows from a similar document provided by the Environmental Data Initiative.* [↑](#footnote-ref-1)