Dataset: Refining Techniques for High-Frequency Monitoring of Chlorophyll in the NERRS

This document provides detailed information about 6 datasets that were generated through a 2020-2021 catalyst project titled *Refining Techniques for High-Frequency Monitoring of Chlorophyll in the NERRS*. This document also provides information about the project. The project was supported by the National Estuarine Research Reserve System (NERRS) Science Collaborative, which is funded by the National Oceanic and Atmospheric Administration. All Science Collaborative supported projects that collect new data adhere to federal data sharing and archiving requirements.

6 related datasets are described in this document:

1. Combined field- and lab-based comparison data collected by participating reserves
2. Monitoring data using different methods from three reserves
3. Interference test results from seven reserves
4. Metadata for sampling details at participating reserves, e.g, site location, protocols
5. Chlorophyll comparison data files from individual reserves
6. Raw files from participating reserves

Data Access and Archival:
The datasets and products from this project have been archived in several ways:

- More information and links to other products generated by project are available through the Science Collaborative: https://nerrssciencecollaborative.org/project/Dix20
- This project and associated data are listed in a national metadata catalog, InPort, link available on the Science Collaborative project page
- All R code developed during the project for data analysis and visualization is publically accessible via www.github.com/skdunnigan/chla-nerrs-catalyst

About the Associated Project

Project title: Refining Techniques for High-Frequency Monitoring of Chlorophyll in the NERRS

Name of reserve(s) involved in the project: Elkhorn Slough, Grand Bay, Great Bay, Guana Tolomato Matanzas), He'eia, Lake Superior, Mission-Aransas, North Inlet-Winyah Bay, Old Woman Creek), Padilla Bay, Sapelo Island, and Wells.

Project Period: December 2020 to November 2021

Project lead and contact information:
Nikki Dix, Guana Tolomato Matanzas NERR, 904-380-8613, nikki.dix@dep.state.fl.us
Erik Smith, North Inlet Winyah Bay NERR, erik@baruch.sc.edu

Silas Tanner, GTM NERR, Silas.Tanner@dep.state.fl.us

Purpose:
Concentrations of the photosynthetic pigment chlorophyll a are used as a proxy for phytoplankton biomass by estuarine scientists and managers to study eutrophication, food web dynamics, and harmful algal blooms. Traditionally, chlorophyll has been measured by filtering a water sample and extracting pigments from the filter in a laboratory, which is the current practice employed by the NERRS in monthly grab samples; however, monthly measurements are not sufficient for tracking plankton dynamics, which fluctuate hourly. Recent sensor technology development allows high-frequency, in situ measurement of chlorophyll on the same YSI EXO sondes used in the NERRS System-Wide Monitoring Program (SWMP). While in situ measurements are related to extracted measurements, there are variations in the environment that cause inconsistencies. Before this project, no tested relationships existed for the EXO sensors, and SWMP practitioners were asking for this information so they can respond to local and national needs for algal bloom research.

Abstract:
Twelve biogeochemically diverse reserves participated in a one -year study to standardize protocols for calibration and data management using both chlorophyll methods (extractions and in situ sensor readings) at various frequencies. Given technician experiences with preliminary sensor deployments and results from previous studies, one objective of this project was to identify possible sensor interferences and develop standardized empirical correction procedures. Another objective of this project was to test how well extracted CHL-A (µg/L) could be predicted from the suite of YSI EXO sensors. While there is value in the YSI EXO Total Algae (TAL) sensor's relative fluorescence units (RFU) for short-term or single-site chlorophyll measurements, reliable predictions of extracted CHL-A (µg/L) would provide a standard measure to compare among sites and with historic data.

About the Project Dataset(s)

1. project-data-all

General description of data:
Data file that contains the combined field- and lab-based comparison data collected by all participating reserves: Elkhorn Slough (ELK), Grand Bay (GND), Great Bay (GRB), Guana Tolomato Matanzas (GTM), He'eia (HEE), Lake Superior (LKS), Mission-Aransas (MAR), North Inlet-Winyah Bay (NIW), Old Woman Creek (OWC), Padilla Bay (PDB), Sapelo Island (SAP), and Wells (WEL).

More about the data:

- Raw spreadsheet
  Any columns with "f_" are the quality assurance/quality control code applied to the data column of the same name.

UNID	Unique identifier for row of information
qaqc code	Quality Assurance/Quality Control Code applied to entire entry
qaqc_notes	Notes to accompany the code applied in previous column
isco_deployment_no	A way to classify ISCO deployments, this is also a running number but it is associated with the actual deployment of the ISCO and not the individual samples collected by the ISCO. In the example template ISCO deployment '1' includes samples '1-11'.
sample_no	A sample is defined as the collected water used to fill one 10L tank
reserve_code	Three letter reserve_code as established by the NERRS CDMO
station_code	Acronym for the SWMP water quality station as submitted to the NERRS CDMO (ex: gtmpcwq)
datetime_collected	Date and time the water sample was collected. Time in 24hr

datetime_filtered	Date and time the sample batch were filtered. Time in 24hr
datetime_extracted	Date and time the sample batch were extracted. Time in 24hr
datetime_analyzed	Date and time the sample batch were analyzed. Time in 24hr
chla_ugL	Extracted chlorophyll a value in micrograms per liter
chla_RFU	Extraced chlorophyll a value in RFU, if applicable
remarks	Any remarks associated with the extraction
depth	Depth data (m) from EXO2 datasonde
level	Level data (m) from EXO2 datasonde
fdom_qsu	Fluorescent dissolved organic matter (qsu) from EXO2 datasonde

fdom_rfu	Fluorescent dissolved organic matter (rfy) from EXO2 datasonde
do_pct	Percent saturation dissolved oxygen (%) from EXO2 datasonde
do_mgl	Dissolved oxygen (mg/L) from EXO2 datasonde
sal	Salinity (psu) from EXO2 datasonde
sp_cond	Specific Conductivity (mS/cm) from EXO2 datasonde
turb	Turbidity (NTU) from EXO2 datasonde
p_h	pH from EXO2 datasonde
temp	Temperature (celsius) from EXO2 datasonde
chlorophyll_rfu	Chlorophyll in fluorescence (rfu) from EXO2 datasonde
chl_fluor	Chlorophyll (µg/L) from EXO2 datasonde
exo_filename	Name of accompanying raw EXO2 datafile that can be found in the reserve-raw-files
sample_code	Code used to identify row value
method	Isco (from diel sampler method) and tank (from methods in the lab with a tank)
field_notes	Any descriptive field notes by sampling technician to be included with data.
rep	Sample replicate number (Method replicates are denoted by a decimal point used to indicate the replicate. Example: 1.1, 1.2 would be method replicates 1 and 2 of sample replicate 1.)
sonde_start_time_24hr	Time EXO sonde was deployed in tank
sonde_end_time_24hr	Time EXO sonde was removed from tank
additional_notes	Notes from Project Lead at reserve-level

- qaqc spreadsheet - contains all the same columns of information as the raw spreadsheet, but has had rejected data removed for analysis (“f_” columns containing <-3> codes)

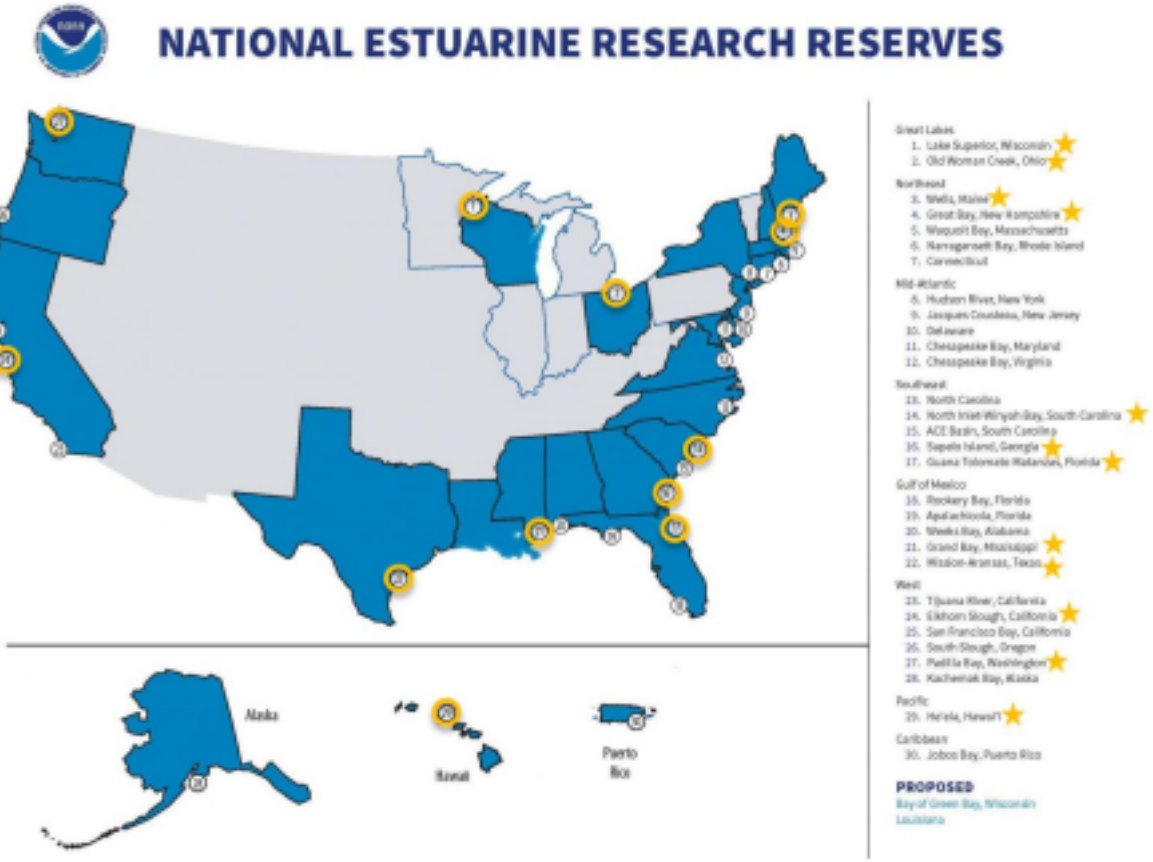
Data collection period: December, 2020 to October, 2021

Geographic extent:

1. LKS -92.295,-91.935,46.73,46.59, Lake Superior, WI
2. OWC: 41.22570, -82.30530 Old Woman Creek, OH
3. WEL -70.6,-70.535,43.35,43.26 Wells, ME
4. GRB -70.97,-70.82,43.18,42.975 Great Bay, NH
5. NIW -79.285,-79.15,33.375,33.204 North Inlet-Winyah Bay, SC
6. SAP -81.31,-81.25,31.5,31.375 Sapelo Island, GA
7. GTM: 29.66694, -81.2575 Guana Tolomato Matanzas, FL
8. GND -88.48,-88.39,30.43,30.31 Grand Bay, MS
9. MAR -97.3,-96.75,28.32,27.82 Mission Aransas, TX
10. ELK: 36.81806, -121.73940 Elkhorn Slough, CA
11. PDB -122.562,-122.455,48.58,48.45 Padilla Bay, WA
12. HEE -157.81,-157.77,21.42,21.45 Heeia, HI

File format:
Excel (.xlsx) workbook files containing spreadsheets with data.

Maps and schematics for data collection:



Research reserves that participated in the study: Lake Superior (LKS), Old Woman Creek (OWC), Wells (WEL), Great Bay (GRB), North Inlet-Winyah Bay (NIW), Sapelo Island (SAP), Guana Tolomato Matanzas (GTM), Grand Bay (GND), Mission-Aransas (MAR), Elkhorn Slough (ELK), Padilla Bay (PDB), and He'eia (HEE).

2. case-studies

General description of data:
Combined YSI EXO2 datasonde data with extracted chlorophyll from the System-Wide Monitoring Program (SWMP) grab and diel samples from three reserves: Elkhorn Slough (ELK), Guana Tolomato Matanzas (GTM), and Old Woman Creek (OWC).

More about the data:
YSI EXO2 raw data (fluorescent dissolved organic matter, chlorophyll fluorescence), extracted chlorophyll concentrations, and other fluorescence values associated with laboratory extraction.

Data collection period: December, 2020 to October, 2021
ELK: January, 2021 to September, 2021
GTM: January, 2020 to December, 2020
OWC: December, 2020 to November, 2021

Geographic extent:
ELK: 36.81806, -121.73940
GTM: 29.66694, -81.2575
OWC: 41.22570, -82.30530

File format:
Excel (.xlsx) workbook files containing spreadsheets with data.

Maps and schematics for data collection:
See ELK, GTM, and OWC in map included in section 1 above.

3. interference -data

General description of data:
Data and metadata from interferences testing on fluorescent dissolved organic matter (fdom), turbidity (turb), and temperature (temp) from participating reserves: Guana Tolomato Matanzas (GTM), Lake Superior (LKS), North Inlet -Winyah Bay (NIW), Padilla Bay (PDB), He'eia (HEE), Mission Aransas (MAR), and Old Woman Creek (OWC).

More about the data:

Please review reserve-specific metadata documents (detailed in Section 4) for more specific details.

Data collection period: July, 2021 to September, 2021

Geographic extent:
See GTM, LKS, NIW, PDB, HEE, MAR, and OWC in section 1 above and Table 3 in the Recommendations Report

File format:
Excel (.xlsx) workbook files containing spreadsheets with data. All files start with type of interference 'fdom', 'turb', or 'temp'. fDOM data file names contain metadata: 'std-' refers to the standard used in the experiment, mostly surrogates distilled from local reserve water, indicated with three letter reserve code; `ambient-` indicates the source of ambient water surogate is applied to; `DI` is a deionized water trial; ending numeric refers to the trial. fdom_std_owc_ambient-ocw_DI_1.xlsx contains spreadsheets for ambient and deionized water trials.
Metadata file for interference testing is a .docx

Maps and schematics for data collection:
See section 1 above.

4. metadata

General description of data:
Metadata files for participating reserves in the project: Elkhorn Slough (ELK), Grand Bay (GND), Great Bay (GRB), Guana Tolomato Matanzas (GTM), He'eia (HEE), Lake Superior (LKS), Mission Aransas (MAR), Old Woman Creek (OWC), Padilla Bay (PDB), Sapelo Island (SAP), and Wells (WEL). Metadata includes contact information for project leads at each participating reserve, project objectives, quality assurance/quality control protocols, entry verification, methods, site location and information, collection period, parameter titles and description for chlorophyll data files, sensor information, limits of detection for chlorophyll analysis, laboratory methods, and relevant references.

More about the data:
Metadata includes contact information for project leads at each participating reserve, project objectives, quality assurance/quality control protocols, entry verification, methods, site location and information, collection period, parameter titles and description for chlorophyll data files, sensor information, limits of detection for chlorophyll analysis, laboratory methods, and relevant references.

Data collection period: November, 2020 to December, 2021

Geographic extent:
See section 1 above and Table 3 in the Recommendations Report

File format:
Word (.docx) document

5. reserve-chla-data

General description of data:
Individual reserve data files for the main project: Elkhorn Slough (ELK), Grand Bay (GND), Great Bay (GRB), Guana Tolomato Matanzas (GTM), He'eia (HEE), Lake Superior (LKS), Mission-Aransas (MAR), North Inlet-Winyah Bay (NIW), Old Woman Creek (OWC), Padilla Bay (PDB), Sapelo Island (SAP), and Wells (WEL).

More about the data:
Please review reserve-specific metadata documents (detailed in Section 4) for more specific details.

- Tank spreadsheet

  Columns A:N all pertain to the sample collected and extracted for analysis.

sample_no	A sample is defined as the collected water used to fill one 10L tank
reserve_code	Three letter reserve_code as established by the NERRS CDMO
station_code	Acronym for the SWMP water quality station as submitted to the NERRS CDMO (ex: gtmpcwq)
datetime_collected	Date and time the water sample was collected. Time in 24hr
field_notes	Any descriptive field notes by sampling technician to be included with data.
rep	Sample replicate number (Method replicates are denoted by a decimal point used to indicate the replicate. Example: 1.1, 1.2 would be method replicates 1 and 2 of sample replicate 1.)
datetime_filtered	Date and time the sample batch were filtered. Time in 24hr
datetime_extracted	Date and time the sample batch were extracted. Time in 24hr
datetime_analyzed	Date and time the sample batch were analyzed. Time in 24hr
chla_ugL	Extracted chlorophyll a value in micrograms per liter

chla_RFU	Extraced chlorophyll a value in RFU, if applicable
remarks	Any remarks associated with the extraction
sonde_start_time_24hr	Time EXO sonde was deployed in tank
sonde_end_time_24hr	Time EXO sonde was removed from tank

Columns O:AO are the averaged EXO data sonde measurements for the data, while the sonde was in the tank, that comes off the KOR export file.

Column AJ (EXO_filename) is the name associated with the trimmed EXO file for the time in the tank.

- ISCO spreadsheet

Columns A:K all pertain to the sample collected and extracted for analysis.

isco_deployment_no	A way to classify ISCO deployments, this is also a running number but it is associated with the actual deployment of the ISCO and not the individual samples collected by the ISCO. In the example template ISCO deployment '1' includes samples '1- 11'.
sample_no	This will be a running number as the project progresses. A sample is defined as the collected water used to fill one ISCO bottle. (Method replicates were collected and denoted by a decimal point used to indicate the replicate. Therefore, ISCO sample 1.1 would be the first method replicate of sample 1).
reserve_code	Three letter reserve_code as established by the NERRS CDMO

station_code	Acronym for the SWMP water quality station as submitted to the NERRS CDMO (ex: gtmpcwq)
datetime_collected	Date and time the water sample was collected. Time in 24hr
datetime_filtered	Date and time the sample batch were filtered. Time in 24hr
datetime_extracted	Date and time the sample batch were extracted. Time in 24hr

datetime_analyzed	Date and time the sample batch were analyzed. Time in 24hr
chla_ugL	Extracted chlorophyll a value in micrograms per liter
chla_RFU	Extraced chlorophyll a value in RFU
remarks	Any QAQC remarks associated with the extraction

Columns L:AK EXO data sonde measurements that are paired with ISCO timestamps. For more information on the EXO files, see Section 13.

Data collection period: December, 2020 to October, 2021

Geographic extent:
See section 1 above and Table 3 in the Recommendations Report

File format:
Excel (.xlsx) workbook files containing spreadsheets with data, see "More about the data" above for sheet and column specifics.

6. reserve-raw-files

General description of data:
Subfolders that contain all accompanying reserve -specific files. These files were not standardized for the project and were organized and uploaded by the project leads at each participating reserve. For more information on these files, it is best to contact the reserve project lead which can be found in the metadata for each participating reserve: Elkhorn Slough (ELK), Grand Bay (GND), Great Bay (GRB), Guana Tolomato Matanzas (GTM), He'eia (HEE), Lake Superior (LKS), Mission-Aransas (MAR), North Inlet-Winyah Bay (NIW), Old Woman Creek (OWC), Padilla Bay (PDB), Sapelo Island (SAP), and Wells (WEL).

More about the data:
Please review reserve -specific metadata documents (detailed in Section 4) for more specific details.

Data collection period: December, 2020 to October, 2021

Geographic extent:
See section 1 above and Table 3 in the Recommendations Report

File format:
Files can be in .bin, Excel (.xlsx), comma separated (.csv), text (.txt), and Word (.docx) formats.