Project Report (FWRI Grant #5473):

Florida's Coral Reef Water Quality Data Compilation, Analysis, and Decision Support
Year 3

Principal Investigator(s) and Project Partners:

- Karen Bohnsack, Florida Keys National Marine Sanctuary
  Karen.Bohnsack@noaa.gov
  305-304-2196
- Andy Bruckner, Ph.D., Florida Keys National Marine Sanctuary
  Andy.Bruckner@noaa.gov
  305-304-3670
- Alexandra Fine, University of Miami Cooperative Institute for Marine and Atmospheric Studies, NOAA Atlantic Oceanographic and Meteorological Laboratory
  alexandra.fine@noaa.gov
  347-280-4339
- Lauren Gentry, Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute
  Lauren.Gentry@myfwc.com
  727-896-8626
- Christopher Kelble, Ph.D., NOAA Atlantic Oceanographic and Meteorological Laboratory chris.kelble@noaa.gov
  786-247-6072
- David Kochan, Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute
  David.Kochan@myfwc.com
  727-896-8626
- Lucas McEachron, Ph.D., Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute
  (Corresponding P.I.)
  Lucas.McEachron@myfwc.com
  727-896-8626
- Kelly Montenero, University of Miami Cooperative Institute for Marine and Atmospheric Studies, NOAA Atlantic Oceanographic and Meteorological Laboratory
  kelly.montenero@noaa.gov
  305-361-4386
- Frank Muller-Karger, Ph.D., University of South Florida
  carib@usf.edu
  727-776-3759
- Tylar Murray, Ph.D., University of South Florida
  email@tylar.info
  813-400-0523
- Dan Otis, Ph.D., University of South Florida
  dotis@usf.edu
  727-424-0991
- Omar Ramzy, University of Miami Cooperative Institute for Marine and Atmospheric Studies, NOAA Atlantic Oceanographic and Meteorological Laboratory
  omar.ramzy@noaa.gov

Project Period:	**Project Cost:**
July 1st, 2022 - June 9th, 2023	$77,000

Florida's Coral Reef Water Quality Data Compilation, Analysis, and Decision Support Year 3

Table of Contents

Summary of Deliverables

- The unified water quality monitoring database was updated with 2022 data from each program and each parameter had their rate of change trends calculated with these new data. The unified database can be found here.
- Two new water quality monitoring programs were added to the database ;the City of Miami Beach Water Monitoring and FDEP Biscayne Bay Aquatic Preserve monitoring met the criteria for inclusion and were incorporated into trend analyses and maps.
- A third monitoring program, Biscayne Bay Aquatic Preserve Continuous Water Quality Monitoring, was added to the database but was not included in trend analyses because of a mismatch in temporal sampling (continuous vs monthly/quarterly).
- Two historical water quality datasets, Florida Keys Nearshore Water Quality Monitoring and Biscayne Bay Water Watch, were added to the database. These programs met most database inclusion criteria other than sampling currently and length of sampling. These data were not included in trend analyses. The historical US EPA Environmental Monitoring Assessment Program was not included due to high temporal variability in sampling.
- Map layers and web apps were updated to depict water quality data and trends that include the new 2022 data. Monitoring survey locations on the map were filtered to only depict sites which have ongoing sampling in included programs.
- Existing water quality databases, such as WIN or SEACAR, were identified as the best solution for hosting the unified water quality data. However, existing digital infrastructure is currently unsuitable and requires updates or modifications to host the dataset and automated tools necessary for a unified water quality monitoring network.
- The WIN database lacks much of the functionality needed to streamline and automate the upload, download, and processing of water quality data. Several problems were identified by data providers and the research team:
    - Data providers identified that using WIN is difficult and time-consuming initially, but after data pipelines are established the process is easier.
    - WIN has specific requirements for included analytes, so using alternative methods would have missing data.
    - Research team members had difficulties determining the exact program names required to download monitoring data.
    - Programs had data missing prior to 2017 due to incompatibility of historical data.
- Discussions were held with FDEP staff to identify gaps in sampling and determine potential solutions:
    - Spatial gaps were identified by overlaying the sampling map over the Unified Reef Map. The largest gaps were observed in the northern ECA, Dry Tortugas, and Marquesas, with smaller gaps off Miami Beach and in southern Broward County.
    - The biggest analyte gap was Total Nitrogen in the northern ECA but it can be calculated from existing data.
    - Requesting Broward County, City of Miami Beach, and BBAP monitor Orthophosphates instead of Total Phosphorus would improve compatibility across Florida’s Coral Reef.

  - Adding turbidity to the Walton Smith cruises would expand turbidity monitoring throughout the Florida Keys.
- Data providers were given recommendations to improve data compatibility and interoperability. Simple changes to reported parameter names and units were suggested to simplify data processing. Managers were informed of the spatial gaps identified by the research team and FDEP staff and possible sampling improvements and solutions.

Relevant Programs

93 water quality monitoring programs in Florida were filtered against five criteria (Figure 1):

1) Sampled within South Florida;
2) Sampled at least four water quality parameters of interest (i.e., Chlorophyll-a, temperature, salinity, nitrate (NO3), nitrite (No2), nitrate+nitrite (NOx), soluble reactive phosphorus (PO4), silica (Si), turbidity, total nitrogen (TN), and total phosphorus (TP));
3) Contains unique sampling data (i.e., not a derivative dataset from another program, or uploaded to multiple portals);
4) Is still actively sampling;
5) Has at least five years of data.

Two new programs were identified as meeting all criteria and added to the five programs included in the unified water quality monitoring database in previous funding years. The City of Miami Beach Water Monitoring and FDEP Biscayne Bay Aquatic Preserve Monitoring each reached the five year data threshold. Five of the 93 programs identified in previous funding years that met all of these criteria were the NOAA AOML-Walton Smith South Florida ecosystem restoration cruises (AOML-Walton Smith); the Southeast Environmental Research Center water quality monitoring network (SERC); Miami-Dade County's Department of Environmental Resources Management (DERM); the Broward County water quality monitoring program (Broward County); and FDEP’s coral ecosystem conservation area water quality assessment program (DEP-ECA). Additionally, two historical datasets meeting every criteria except active sampling were added to the database but not included in trend analyses; Florida Keys Nearshore Water Quality Monitoring and Biscayne Bay Water Watch were included in the unified water equality monitoring database. The complete matrix comparing water quality programs can be found here.

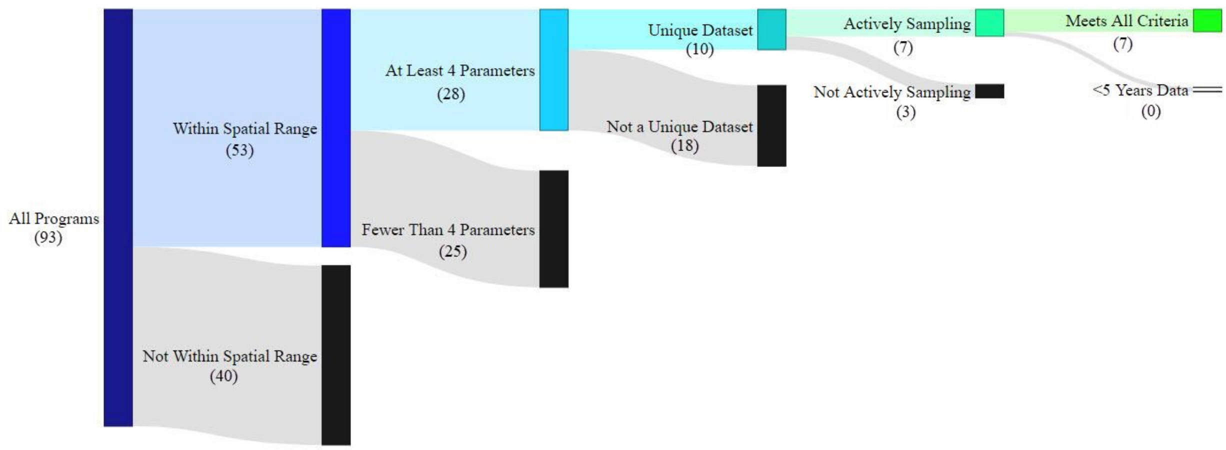

Figure 1. Sankey chart of water quality monitoring programs relevant to South Florida. Seven programs met all criteria considered.

Task 1: Identify a database management solution to host programmatic updates.

Deliverable: Description of database management solution and data upload form

A simple, straightforward approach for data upload by providers should be developed. This could consist of a checklist form for providers to submit with water quality data and associated metadata in a plain-text file format. This format is similar in structure to what they already provide; data uploads should occur via a simple interface similar to Google Drive, the Github online editor, or a web form. An example form can be found here:
http://fgbnms-dashboard.marine.usf.edu:5000
This example is a simple form where the user selects a filetype and then selects a file for upload. Newly uploaded data files can run through a QC check and load into a public file-based data store. Any QC check scripts should be published on github as python basic functions, providing maximum transparency and adaptability for the community of data providers. This simple html form is also accessible via POST request, so submission could be easily automated and more advanced submission interfaces can be easily built on top of this foundation.

Databases that provide more performant access to the data based on user needs should be spun up as needed. We recommend an ERDDAP as a database solution (https://coastwatch.pfeg.noaa.gov/erddap/index.html), which is a data server that provides a simple, consistent way to download subsets of both tabular and raster scientific datasets in common file formats and make graphs and maps. ERDDAP may be utilized to provide an advanced data Application Programming Interface (API), allowing for easy subsetting and filetype transformations based on the needs of end users. ERDDAP provides Graphical User Interface (GUI) for drafting REpresentational State Transfer (REST) requests; this design accommodates the full needs of a GUI-only user but also promotes development of technical capacity for advanced API usage. The GUI provides guidance on spatiotemporal subsetting of

data files and provides easy access to dozens of common data file formats (csv, nc, mat, png, etc).
Unfortunately, there is currently no way for an ERDDAP server to automatically retrieve data from the (Watershed Information Network (WIN) database via the use of an Application Program Interface (API) or other method (see Task 3 for additional WIN limitations). One solution is to replace the WIN database with an ERDDAP server entirely to 1) continue to provide a single point of entry for data providers, and 2) enable consistency and collaboration with the broader scientific community. Alternatively, FDEP could approach the original WIN developers with a custom programming request to enable ERDDAP compatibility. The latter option may be preferred because providers have already adapted to uploading data to WIN.

Figures 2 and 3 show an example tabular dataset hosted on an ERDDAP server. Users can create graphs of the data (Figure 2) and also download the data in a format of their choice (Figure 3).
http://www.neracoos.org/erddap/tabledap/A01_optics_s_all.graph?time%2Cchlorophyll&time%3E=2015-05-10T14%3A36%3A44Z&.draw=lines&.color=0x000000&.bgColor=0xffffffff

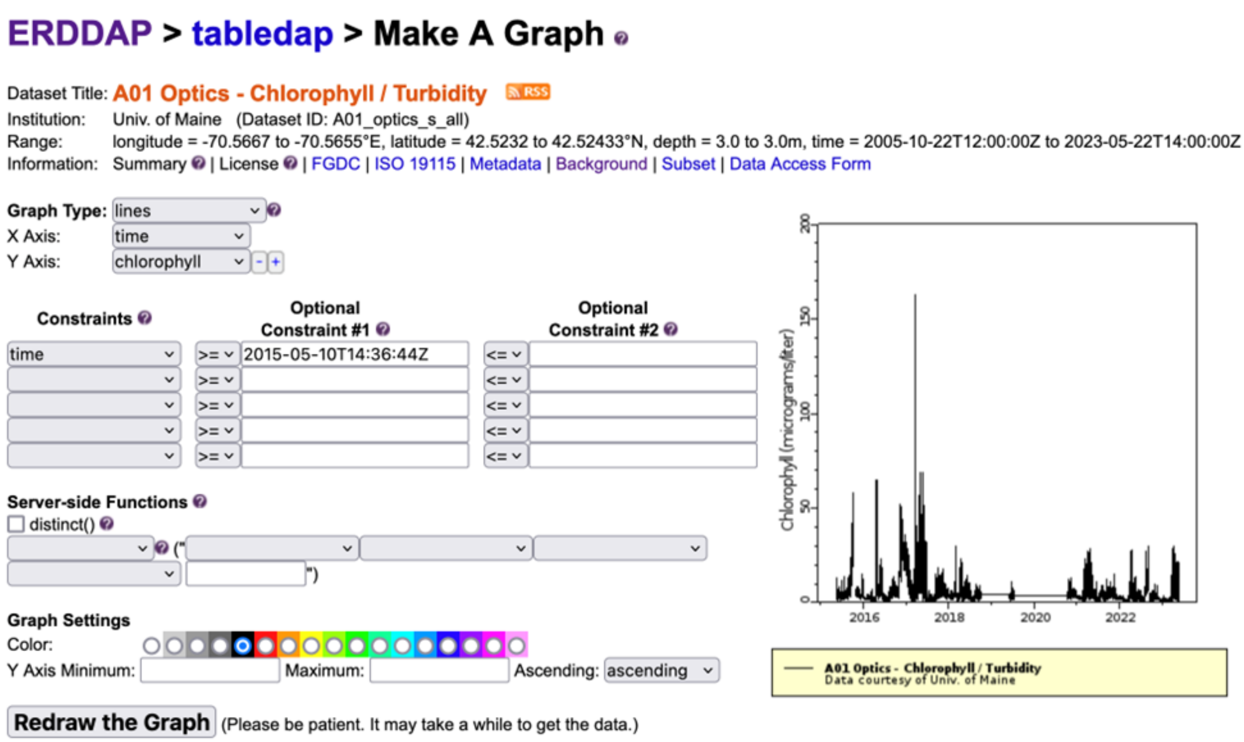

Figure 2. Example ERDDAP server graph interface with tabular data similar to what could be generated by this product. Here, users can create a graph of their parameter of choice (chlorophyll or turbidity) and set x and y axis options; none of these capabilities are available in WIN without significant manual processing.

Users can also select parameters, create subsets and download data in one of several common formats using the ERRDAP "Data Access Form" as shown in Figure 3.

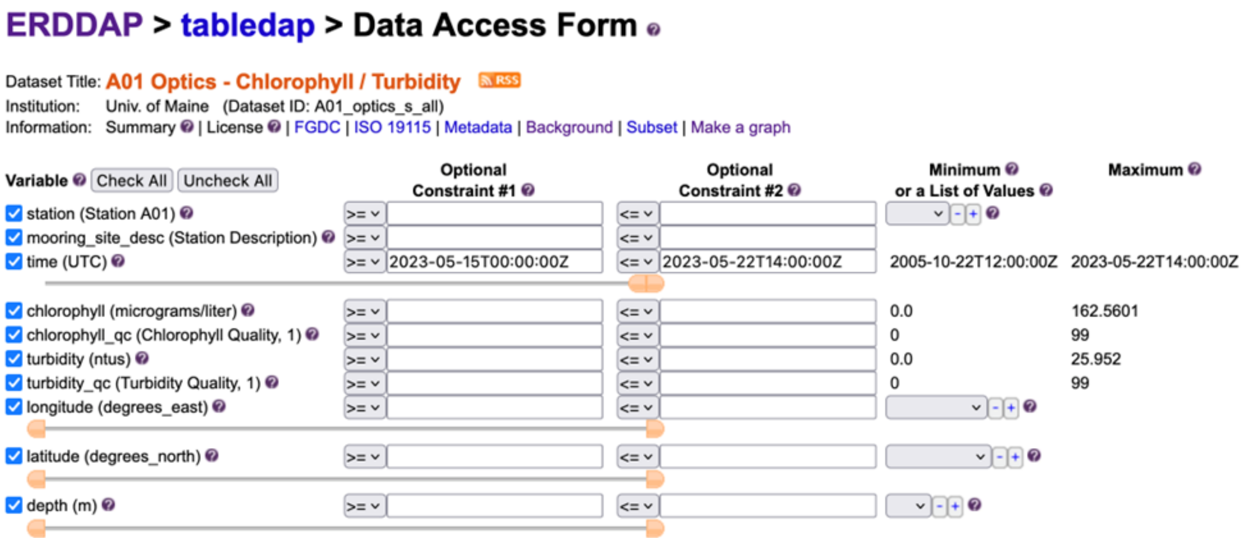

Figure 3. Example ERDDAP Data Access Form for the example dataset shown above in Figure 2. Here, the user can select from one of two parameters (chlorophyll or turbidity) and also extract latitude/longitude and depth information.

In terms of data visualization, the majority of frontend development can be done using R; more specifically: the Quarto, Shiny, and flex dashboard libraries. The Tampa Bay Estuary Program has developed a water quality data dashboard based on these libraries (https://tbep.org/water-quality-dashboard/), which will serve as a template for the frontend user interface to be developed for this project over the 2023-2024 funding cycle.

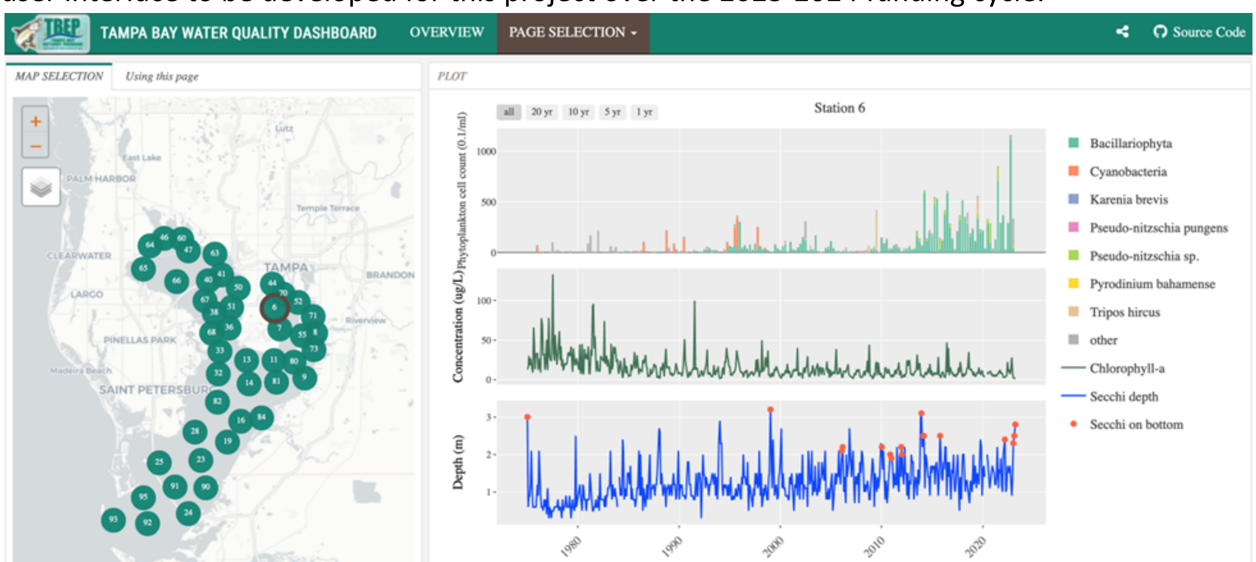

Figure 4. Example water quality data dashboard developed by the Tampa Bay Estuary Program to display time series of water quality parameters in Tampa Bay (https://tbep.org/water-quality-dashboard/). The plots at right display data from a user-selected location (Station 6 in Hillsborough Bay in this example). Locations are selected by clicking on the map at left.

The example in Figure 4 shows “Site Trends” of several water quality parameters in Tampa Bay for a location selected by the user. Once a location is selected by the user, a set of time series of parameters of interest are displayed. The source code for this data dashboard is publicly available and will be tailored for use in South Florida and along Florida’s Coral Reef to cover the water quality data in the unified database compiled as part of this project.

Task 2: Identify specific locations on Florida’s Coral Reef that could be sampled differently between programs, conduct outreach to individual programs to enhance interoperability, and address any remaining statistical issues pertaining to combining programmatic data.

Deliverable: Summary notes of meetings with five, minimum, water quality monitoring programs

The water quality research team held multiple meetings with FDEP staff and water quality program managers to discuss potential programmatic changes to improve water quality monitoring and sampling across Florida’s Coral Reef and streamline the data aggregation process. The May 19, 2023 meeting was held with representatives from all 7 water quality monitoring programs.

a. *September 16, 2022. FDEP Water Quality Project Meeting*

The research team held a meeting with FDEP representatives on September 16, 2022 in Miami, FL to identify opportunities to improve sampling spatial coverage and discuss potential solutions for storing the unified water quality monitoring dataset. We overlaid the results of the spatial analyses conducted during the 21/22 funding year on the hardbottom reefs mapped by the Florida Unified Reef Map to visualize gaps specifically on Florida’s Coral Reef. Using the spatial coverage of the total phosphorus, the analyte with the smallest calculated range, we identified several areas in Southeast Florida and the Florida Keys with inadequate sampling. The northern extent of the ECA had the largest continuous gaps in sampling on hardbottom (Figure 5), with smaller gaps present throughout Broward and Miami-Dade counties. Regular sampling gaps were seen in the Florida Keys, but the biggest needs were identified in the Marquesas and Dry Tortugas; however, the challenge of increasing sampling frequency and coverage in such remote areas may require consideration of different techniques such as continuous sampling with autonomous sensors.

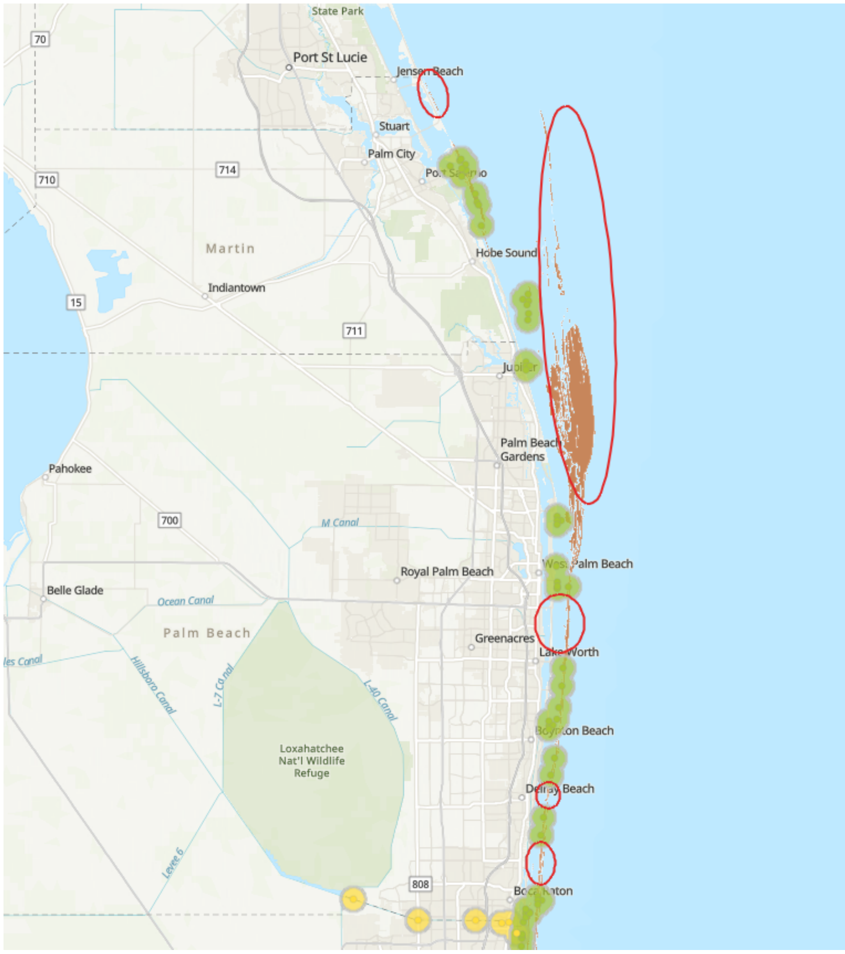

Figure 5. Sampling gaps in the northern ECA that are identified by overlaying the results of the gap analysis on the Unified Reef Map hardbottom and coral reef layer. The green circles are sampling sites with a 1.7 km buffer representing the most conservative spatial overlap of measured water quality analytes.

After examining the spatial coverage of sampling locations, the research team and FDEP discussed the possibility of requesting individual monitoring programs to include additional analytes that are lacking in specific areas. The largest analyte gap was total nitrogen in the northern ECA, which can be calculated and reported by FDEP without extra sampling effort. Both total phosphorus and orthophosphates showed gaps on the FCR, and FDEP suggested that requesting Broward County, City of Miami Beach, and BBAP to monitor orthophosphates instead of total phosphorus would improve continuity. Finally, adding turbidity to the Walton Smith cruises would expand turbidity monitoring throughout the Florida Keys.

FDEP also requested the inclusion of historical datasets into the unified water quality monitoring database where possible. The research team identified two programs, Florida Keys Nearshore Water Quality Monitoring and Biscayne Bay Water Watch, that meet all the non-temporal inclusion criteria. We also added to the database the ongoing BBAP continuous water quality sampling program to provide higher temporal resolution data and provided a template for future continuous sampling programs. Neither the historic datasets nor continuous sampling were included in trend analyses as the data lacked interoperability, but merging these datasets will provide important information for researchers.

Finally, research team staff and FDEP discussed potential solutions for storage and access of the unified water quality monitoring dataset. FDEP staff expressed a desire to avoid creating a new web data portal and instead preferred the option of using an existing database to host the unified data. As such, the research team shifted focus away from creating a new ERDDAP server and began scoping the potential of adapting WIN or SEACAR as a possible hosting solution.

b. *May 19, 2023. Meeting of Water Quality Program Managers in South Florida.*

The research team held a virtual outreach meeting with FDEP and water quality monitoring program managers to discuss improving data interoperability, streamlining data aggregation and processing, and identifying potential ways to improve the usability of the WIN database for both uploading and downloading monitoring data. Representatives from each of the 7 water quality programs attended and were given project background and updates on the temporal (trend) and spatial (gap) analyses conducted in the 21/22 FY. Program managers and data providers were asked about their challenges when using the WIN database and opportunities to streamline the data upload process. The goal of these discussions was to limit the need for the research team to manually contact individual programs for data updates and instead create a pipeline that uses existing databases to streamline the process. Specifics about the challenges of using WIN as a database solution are detailed in Task 3.

Participants further discussed data interoperability and potential solutions for filling gaps. While WIN does not accept calculated analytes, reporting total nitrogen and others in the database with post-processing would provide important information for database users. Additionally, WIN only accepts data from NELAC certified labs and methods, but there are opportunities to use alternative methods of analysis, remote sensing, or continuous sampling to increase data coverage of the FCR. Data collected by other methods may not be interoperable with the unified water quality database but there is a need for increased spatial and temporal coverage. Representatives from the FDEP ECA water quality sampling program asked for more detailed information about sampling gaps in the ECA

Task 3: Determine where the database management solution will sit for long-term sustainability, and/or if there is an existing solution that meets regional management needs.

Deliverable: Report on long-term solutions and any existing solutions

A database management solution is useful when there are existing web services for distributing data, but the solution can also access data that is stored locally (i.e., sent via email). Each dataset requires a small set of code to tell the solution how to access the dataset. After the data is online, end users can request the data in various ways and in various formats. While relatively simple, the solution requires some maintenance and monitoring as data change, as users manually upload data, and/or as end-user needs change.

The current database solution in place for data providers for this project and many others, is the Watershed Information Network (WIN) database, which is administered by FDEP. The majority of the programs that provide data for the unified database developed as part of this project upload their data to WIN. Based on our discussion with data providers and the project team on May 19th, 2023, the WIN database in its current form largely meets the needs of data providers in terms of uploading of data. Data providers upload data to WIN quarterly after QC. Data providers are also able to update their data in the WIN database from time to time based on the subsequent discovery of data quality issues with previously submitted data. Data providers stated that the initial connection to the database, selection of the correct parameters and data uploads can be daunting for new users. According to data providers, this process often requires the help of WIN data coordinators who are part of FDEP. However, once the initial connection to the database is established and these issues are worked out, data uploads to the database are relatively straightforward.

The main issues related to the usability of the WIN database concern data retrievals. Several issues were raised by the project team in this regard:

1. Data collected prior to 2017 is not currently in WIN. Data uploads to the WIN database began in 2017 and historical data prior to that are stored in EPA's STOrage and RETrieval (STORET) database. It is unclear if data collected prior to 2017 will ever be migrated to WIN.

2. Some analytes are not named consistently. The use of "Chlorophyll-a" vs. "chlorophyll-a" vs. "Chlorophylla" for example can create issues during the data ingestion process. There is an option to use a "cross-analyte ID", which may alleviate some of these issues.

3. The names of programs that provide data can be non-intuitive, which makes accurate queries of existing data difficult. There is an option to use a WIN Advanced View & Extraction System (WAVES) query to identify and extract the data of interest.

4. There are no calculated values in WIN, only direct measurements of an established list of analytes. For example, there is no way to extract values of total nitrogen from the WIN database, so a user must do those calculations themselves.

5. There is no way to retrieve data from the WIN database in an automated manner via the use of an Application Program Interface (API) or other method that enables compatibility with an ERDDAP server, which is the preferred database management solution. Data retrieval from WIN requires a user to navigate the GUI interface to select the correct data provider programs and parameters.

Overall, several data providers and members of the project team stated that the WIN database is not user-friendly and often requires assistance from WIN data coordinators or other FDEP personnel to be able to use the database. The WIN database seems to function well for data providers for data uploads once an initial connection to the database is established. However, the main issues related to WIN occur when retrieving data to create a unified database, as we have developed in this project, or for use in a visualization or decision support tool. One goal of this project is to streamline the entire workflow to make it as automated as possible. This avoids the need for personnel to manually extract data from WIN and will allow for more frequent updates of the unified database. The current configuration of WIN makes this difficult and labor intensive given the lack of functionality for automated retrievals.

The WIN database is well-established and the current data providers are able to use it for data uploads with minimal issues. We see no need to replace it with another database solution at this time. However, we would like to explore the possibility of adding more functionality to the database in terms of automated data retrievals, the migration of historical data collected prior to 2017 and perhaps more documentation or tutorials to make the database more user-friendly for data retrievals. This will be one of the tasks for year 4 of this project.

Contact List for Data Providers

- **AOML-Walton Smith**: NOAA AOML South Florida Ecosystem Restoration Cruise Data
    - Alexandra Fine (alexandra.fine@noaa.gov)
    - Ian Smith (ian.smith@noaa.gov)
- **SERC**: Florida International University South Florida Estuaries Water Quality Data
    - Dr. Yan Ding (yding@fiu.edu)
    - Dr. Henry Briceno (bricenoh@fiu.edu)
- **DERM**: Water Quality Monitoring Program
    - Omar Abdelrahman (omar.abdelrahman@miamidade.gov)
    - Yin Chen (yin.chen@miamidade.gov)
- **Broward County**: Water Quality Monitoring Program
    - Broward County Environmental Lab (Resilience@broward.org)
    - Lindsey Visser (lvisser@broward.org)
- **DEP-ECA**: Coral Ecosystem Conservation Area Water Quality Assessment
    - Alycia Shatters (alycia.shatters@dep.state.fl.us)
- **Miami Beach:** City of Miami Beach Water Monitoring
    - Elizabeth Wheaton (ElizabethWheaton@miamibeachfl.gov)
- **BBAP:** North Biscayne Bay Seagrass Loss Water Quality Program & Biscayne Bay Aquatic Preserves Continuous Water Quality Monitoring
    - Griffin Alexander (griffin.n.alexander@floridadep.gov)

Selected Meetings, Presentations, and Materials

- September 16, 2022. FDEP Water Quality Project Meeting. Notes

- January 31, 2023. “Florida’s Coral Reef water quality data compilation and analysis”, Gulf of Mexico Alliance Data and Monitoring Team mid-year meeting. Presentation

- May 19, 2023. Meeting of Water Quality Program Managers in South Florida. Notes