Florida Coastal Water Quality Monitoring Data Assessment, Access and Training

By:

Katie Petrinec
Guana Tolomato Matanzas National Estuarine Research Reserve

Jeremy Cothran and Dwayne E. Porter
Arnold School of Public Health, University of South Carolina
and
NERRS Centralized Data Management Office

Date Submitted:

31 October 2018

Executive Summary

The Florida Department of Environmental Protection (FLDEP) Florida Coastal Office (FCO) manages 41 special aquatic areas established as preserves to be maintained in essentially natural conditions for the benefit of future generations. In coordination with NOAA, FCO also manages three reserves which are part of the National Oceanic and Atmospheric Administration (NOAA) National Estuarine Reserve System (NERRS). The FLDEP/FCO Florida Aquatic Preserves Program (FLAPP) strives to identify stakeholders, determine their environmental observing needs, and provide data and/or data products that meet their requirements. Stakeholders are defined as those within the FLDEP and external to the FLDEP that require, need access to, and benefit from the data, products and services provided by the FLAPP. The FLAPP recognizes that stakeholder engagement does not end with the delivery of data or products; the development and nurturing of lasting and mutually beneficial partnerships are also critical with these stakeholders. The FLAPP must maintain an open dialogue with stakeholders and assure that data, products, and services evolve as stakeholder needs change. The overall purpose of this project is to define a process to assess 65 years of water quality data collected by the Florida Aquatic Preserve Program (FLAPP) and provide recommendations for the establishment of a system-wide FLAPP water quality monitoring program consistent with the NERRS System-Wide Monitoring Program (SWMP).

In order to most fully support the needs and responsibilities of their stakeholders as related to issues of water and water quality, the FLAPP should design, implement, support and maintain a system-wide water quality monitoring program including the following elements:

- A centralized data management office through the establishment of a data coordinator.
- Infrastructure and personnel at Aquatic Preserve offices necessary for reliable water quality data collection, maintenance, management, documentation, access and archival.
- Ability to access a variety of data types (e.g. satellite, HF radar, glider, buoy) and, as appropriate, integrate these data with FLAPP-collected data.
- A relational database structure that allows for rapid and customizable data queries to meet specific user needs.
- Quality assurance and quality control (QA/QC) protocols that identify spurious data and/or data anomalies prior to access and archival.
- A metadata strategy that ensures that all aspects of data collection, management and QA/QC are documented.
- A data portal providing public access to data, metadata and derived products.
- Web-based tools, including map-based products that enable spatial visualization of data and derived information, and tailored alerts that notify the end user when user-specified thresholds are exceeded.
- Staff training and development of a community of data management experts in the region.

Introduction

A challenging opportunity faced in the Southeast US and in particular the State of Florida (FL) is the integration of multiple local, sub-regional, regional, and national coastal and ocean observing systems that span inshore waters, estuaries, and nearshore and open waters of the Atlantic Ocean and Gulf of Mexico. The true value of these observing systems is only achieved when the data and derived information are used to advance science, improve informed decision making, and also repurposed to meet applied needs that include stakeholder decision making. Stakeholders are defined as those within the FLDEP and external to the FLDEP that require, need access to, and benefit from the data, products and services provided by the FLAPP. The experiences (successes and/or failures) of these diverse systems, both individually and amassed can be used to identify ongoing challenges and issues related to end-user requirements. Moreover, they serve as lessons learned when developing strategies to meet current and future user needs. The purpose of this Deliverable 1.2 is to integrate a current understanding of the state of coastal and ocean monitoring and observing system efforts in the Southeast with an assessment of water quality monitoring efforts, needs, expertise and resources and present straightforward recommendations for the future.

The Florida Department of Environmental Protection (FLDEP) Florida Coastal Office (FCO) manages 41 special aquatic areas established as preserves to be maintained in essentially natural conditions for the benefit of future generations. In coordination with NOAA, FCO also manages three reserves which are part of the National Oceanic and Atmospheric Administration (NOAA) National Estuarine Reserve System (NERRS). The FLDEP/FCO Florida Aquatic Preserves Program (FLAPP) strives to identify stakeholders, determine their environmental observing needs, and provide data and/or data products that meet their requirements. Stakeholders are defined as those within the FLDEP and external to the FLDEP that require, need access to, and benefit from the data, products and services provided by the FLAPP. The FLAPP recognizes that stakeholder engagement does not end with the delivery of data or products; the development and nurturing of lasting and mutually beneficial partnerships are also critical with these stakeholders. The FLAPP must maintain an open dialogue with stakeholders and assure that data, products, and services evolve as stakeholder needs change. The overall purpose of this project is to define a process to assess 65 years of water quality data collected by the Florida Aquatic Preserve Program (FLAPP) and provide recommendations for the establishment of a system-wide FLAPP water quality monitoring program consistent with the NERRS System-Wide Monitoring Program (SWMP).

FLAPP Water Quality Monitoring Data Assessment

In August 2018 the CDMO and the Data Coordinator interviewed eight Aquatic Preserve offices to assess historical and current FLAPP water quality monitoring activities. An outcome of the interview process and data review includes assessing the quality of data, identifying needs, providing insight into potential causes of data loss, and providing recommendations to improve data collections.

The Florida Aquatic Preserves were equipped with YSI 6-Series Multiparameter data sondes in 2004 for continuous water quality monitoring. Their water quality stations were established using protocols outlined in the NERRS SWMP Standard Operating Procedures and instrument exchanges made at the water quality stations vary with most exchanges occurring between two-week and

monthly intervals. All calibration, post-deployment readings, and reconditioning of the data sondes and sensors were performed in accordance with methods outlined by either YSI's Operating and Service Manual, the NERR SWMP YSI 6-Series Multi-Parameter Water Quality Monitoring Procedure SOP, or the FLDEP's Division of Environmental Assessment and Restoration (DEAR) Specific Field-Testing Procedures FT 1100 through FT 3000 or a combination of the three.

Historic data ingested into the data portal were collected using both 30-minute (Big Bend Seagrasses Aquatic Preserves (2004 - 2008), Central Panhandle Aquatic Preserves, East Central Aquatic Preserves (office closed), Northeast Aquatic Preserves) and 15-minute intervals (Estero Bay Aquatic Preserve, Charlotte Harbor Aquatic Preserves, Tampa Bay Aquatic Preserves, Northwest Aquatic Preserves, and Big Bend Seagrasses Aquatic Preserves (began in 2009-current)). Data were evaluated for missing, suspect, and/or anomalous readings and metadata were maintained by FLAPP staff for the following offices:

- Central Panhandle Aquatic Preserves
- Northwest Aquatic Preserves
- Big Bend Seagrasses Aquatic Preserves
- Charlotte Harbor Aquatic Preserves
- Estero Bay Aquatic Preserve

Station data imported to a common relational database were mostly in CSV or Excel format files. The main problems with importing the data were:

- accounting for differences in the types of parameters measured across stations or the same station at different time periods
- differences related to how parameters were column titled or ordered in the file header
- differences in how many header lines were listed before the listing of measurements or the line number which contained the primary header titles
- for excel, where multiple spreadsheets were used, differences related to how the spreadsheets were titled, ordered or organized
- trash or truncated field values or filler values, such as a decimal (.) value which were difficult for import scripts to ignore or interpret.

Some files also required manual review to truncate pre- and post- deployment values that should not be considered part of the dataset. Providing and following a recommended data collection format should reduce time spent manually correcting issues that cause an automated import of data files to fail.

Metadata was created by the Data Coordinator as data were investigated for validity based on weather and field observations, quality control checks, graphs, and instrument diagnostics for the following stations:

- Tampa Bay Aquatic Preserves
- Northeast Aquatic Preserves
- East Central Aquatic Preserves

All water quality data are undergoing a two-step (primary and secondary) Quality Assurance / Quality Control (QA/QC) process as outlined in the *NERRS CDMO Data Management Manual*. The primary QA/QC process was performed by the CDMO and is an automated process that applies standardized flags to data that are outside sensor specifications as determined by YSI, the instrument manufacturer. Yearly data files were exported from the project database and sent to the Data Coordinator. The secondary QA/QC process is currently ongoing and is being performed by the Data Coordinator. Yearly data files are being evaluated, and standardized flags and codes as defined by the NERRS CDMO Data Management Manual were applied to flag columns contained within the secondary QA/QC data files using a macro distributed by the CDMO. Missing and anomalous data are flagged and coded using yearly metadata documents when applicable. For offices that did not submit metadata documents, the Data Coordinator performed the secondary QA/QC review by investigating data for validity and compiling review notes for inclusion in metadata documents. Data that have completed the secondary QA/QC process are sent to the CDMO and are uploaded to the FLAPP database and available through the FLAPP data portal.

Although data collected by the FLAPP appear to adhere to protocols (NERRS SWMP and/or FLDEP DEAR) and efforts have been made to ensure the accuracy of the data; historic data were collected using varying methods. All sites experienced missing data with some having experienced larger data gaps then others. In most occurrences missing data was a result of instrument or sensor malfunctions, sensors not being deployed, or maintenance/calibration of equipment. Large gaps in data were attributed to the type of instrument used, repair/replacement of a sampling stations, or lost files.

Considerations for Establishment of a FLAPP System-wide Water Quality Monitoring Program

System-wide Network Design:

Data and information management (DIM) is a necessary critical component of an observing system. It is the repository of information gathered by and in support of the observing system and it is the pathway for communication to engagement. The need to interface with all elements of the observing system can make it difficult to delineate data management, and for this reason it is important to clearly define the breadth and scope of end users and their relevant needs as related to data and derived information products.

- Developing and sustaining healthy stakeholder relationships requires a commitment from both parties; the data and services provider and the end user. It is incumbent on the FLAPP to ensure that all components of the observing system (i.e. observing, data management and communication, modeling, outreach) actively engage in determining end-user requirements and ensuring that these requirements are met. For example, the DIM team must be cognizant of the web “sophistication” of the intended user(s) and develop products and tools appropriate for that audience. Conversely, stakeholders need to be proactive, responsive and realistic in conveying their requirements and needs. Both parties must recognize that successful relationships are not built upon an “if you build it they will come” mentality. Experience has shown that the most

successful partnerships have resulted from an iterative developmental process informed by trial and error.

- A successful example of this iterative process was the development of the Marine Weather Portal for the NOAA National Weather Service (NWS). Southeast Coastal Ocean Observing System Regional Association (SECOORA; www.secoora.org) members worked with local NWS Weather Forecast Offices (WFO) in the Southeast and Gulf of Mexico to create the portal (http://secoora.org/data/marineweatherportal). These WFOs wanted a "one-stop-shop" for mariners and fishermen to access marine weather, hazards, and real-time data. This website was designed to meet the needs of NWS forecasters as well as mariners and therefore required repeated testing by NWS staff and focus groups comprised of mariners and fishermen. Consequently, it took several years to reach the point where all parties were satisfied with the final product.
- An important lesson learned is that successful product development requires a commitment of time and effort from all parties and a realization that these efforts may extend beyond the life of a typical grant cycle. There are additional roles for agency partners and end users as well. These entities need to continue to advocate for regional data, products, and services, as well as the groups that provide this valuable information, especially when these data, products and services support their individual agency missions. Moving forward requires improved communication amongst all parties, alignment of FLAPP initiatives down to the sub-regional and local level, and greater investment from non-FLAPP partners.
- Stakeholders need to advocate for local data, products, and services, and for those who provide this information, as they support the agency missions. Further, partners should actively engage in the development of plans, as appropriate, to inform observational gap analyses and to assist
- In addition to the current suite of data providing partners, explore additional data provider partnerships to include federal, state and local regulatory and public health entities including in situ and remotely sensed data, model output, and data and model-derived products.

Data Quality:

The quality of the data ingested by and accessible from the FLAPP data portal is critical to the success of the effort. As part of the FLAPP efforts to provide access to high quality monitoring data, FLAPP should:

- Continue to partner with and take advantage of established estuarine and riverine environmental monitoring programs in the region; in particular the NOAA National Estuarine Research Reserve's System-wide Monitoring Program (SWMP) and associated Centralized Data Management Office (CMDO) provide access to a wealth of expertise, tools, documentation and training in support of long-term environmental monitoring of biotic, abiotic, nutrient and meteorological conditions.
- Establish, document and adhere to a standardized system of data quality assurance and quality control that helps to identify spurious data and/or interesting data anomalies for data collected under the auspice of the FLAPP.

    - The CDMO provides a series of tools to support QA/QC. The FLAPP should review these available tools and adopt as appropriate.
- Require appropriate metadata to be developed and maintained for all data available via the FLAPP data portal (See comments under *Data Access and Archival* for more recommendations on metadata.)
- Develop and publish a strong data disclaimer!

Data Access and Archival:

The ultimate purpose of amassing large volumes of data is to be able to optimize their use for various assessment, management or analytical purposes. Thus, the FLAPP data portal should serve as a central organized platform from which data can be exported via a variety of formats and services. At the national level, recent focus has been on identification of various data applications and designing the services that provide new applications and data utility. Numerous groups have developed open source code for a variety of formats that are periodically delivered to users and for data-based "services," which are custom-designed for particular types of users.

Depending on the needs of FLAPP end users, further possible formats and services include the following:

- Google Maps/Earth, KML – allows mapping and sharing of KML/KMZ data via Google Maps/Earth
- Time-series graphs – individual sensor/observation graphs
- GeoRSS – platform-specific GeoRSS feeds which provide hourly updates for RSS type readers
- GeoJSON – platform-specific GeoJSON feeds which provide hourly updates for Javascript oriented browser functionality
- HTML tables – platform-specific HTML tables which can be dynamically pulled for AJAX style applications doing webpage HTML content swapping
- CSV – CSV for Excel or general import,
- Shapefile - Shapefile for GIS-style applications
- Open Geospatial Consortium (OGC) web services such as WMS (Web Mapping Service), WFS (Web Feature Service) and SOS (Sensor Observing Service)
- Quality-control and event notification – the archival collection of data allows for self-referential quality-control and trending of data and event notification when observations exceed known historical, local parameters.
- FLAPP should consider developing a data sharing agreement to include the publication of federally-approved metadata to be provided with all available data.
    - The NOAA is recommending the ISO 19115-2 metadata standard which requires nearly complete FGDC-format metadata.
    - While the adoption of standard formats and protocols for data collection should be encouraged, providing complete metadata would not require data providers to adhere to a data collection standard.

    - One-off applications could continue to be constructed for providers not able to meet data sharing standards in the event FLAPP has resources to allocate to those providers.
    - FLAPP should consider taking advantage of regional, state and national data access and archival initiatives by NOAA's National Centers for Environmental Information (NCEI; https://www.ncei.noaa.gov/), and the Southeast Coastal Ocean Regional Association (SECOORA; www.secoora.org).
- Funding decisions should be tied, in part, to clearly identified user-needs and should require that funding recipients document stakeholder use of their data, product or service. This documentation should include stakeholder input that provides metrics, if possible, on how observing data, data products, and/or services have impacted their agency or individual performance.

Draft Recommendations and Other Considerations

The overall purpose of this project is to define a process to assess 65 years of water quality data collected by the Florida Aquatic Preserve Program (FLAPP) and provide recommendations for the establishment of a system-wide FLAPP water quality monitoring program consistent with the NERRS System-Wide Monitoring Program (SWMP). In order to most fully support the needs and responsibilities of their stakeholders as related to issues of water and water quality, the FLAPP should design, implement, support and maintain a system-wide water quality monitoring program including the following elements:

- A centralized data management office through the establishment of a data coordinator.
- Infrastructure and personnel at each Preserve necessary for reliable water quality data collection, maintenance, management, documentation, access and archival.
- Ability to access a variety of data types (e.g. satellite, HF radar, glider, buoy) and, as appropriate, integrate these data with FLAPP-collected data.
- A relational database structure that allows for rapid and customizable data queries to meet specific user needs.
    - A prototype relational database containing historical water quality from eight (8) Aquatic Preserves has been developed as part of this effort.
- Quality assurance and quality control (QA/QC) protocols that identify spurious data and/or data anomalies prior to access and archival.
    - As reference above, CDMO provides access to well-documented tools to support QA/QC of SWMP water quality data and the FLAPP is encouraged to utilize these tools as appropriate.
    - NOAA also supports consistent QA/QC procedures through the implementation of the Quality Assurance for Real-Time Oceanographic Data (QARTOD, https://ioos.noaa.gov/project/qartod/) and represents another option as related to consistent QA/QC practices.
- A metadata strategy that ensures that all aspects of data collection, management and QA/QC are documented.
    - As noted above, NOAA is recommending the ISO 19115-2 metadata standard which requires nearly complete FGDC-format metadata.

    - Developed metadata should be uploaded to the NOAA InPort (https://inport.nmfs.noaa.gov/inport/) which is the centralized repository of metadata as required by the NOAA Data and Information Policy and Directive and the Data Documentation Procedural Directive.
- A data portal providing public access to data, metadata and derived products.
    - A prototype data portal has been developed as part of this assessment providing access to the prototype relational database referenced above (http://165.227.254.90/flapp/).
- Web-based tools, as appropriate, including map-based products that enable spatial visualization of data and derived information.
- Staff training and development of a community of data management experts in the region.

Recognizing that funding is often identified as a limiting factor in determining what or what not to implement in support of an environmental monitoring program, additional components for Data Acquisition and Management consideration include:

- Establish a Data Management Committee (DMC) to provide guidance, oversight and support for the DIM team. The DMC should be comprised of a membership with relevant expertise in environmental observing, data management, product development, and local product needs. Committee membership should be revolving and staggered to ensure an influx of fresh ideas while ensuring continuity.
- Establish a training network and annual workshop to engage data providers, data managers, and end users to support program unity, rigor and relevance.
    - Such a training network can be done with the confines of the FLAPP or in collaboration with local programs such as the NOAA NERRS, IOOS Regional Associations, and state NOAA Sea Grant programs.
    - The CDMO hosts an annual Technician Training Workshop (TTW) to provide SWMP technicians and a limited number of external attendees basic to advanced training on all aspects of the SWMP. FLAPP staff have participated in the TTW in recent years. While the CDMO is providing FLAPP staff access to online training modules on a variety of monitoring and data management topics, the recommendation is that the FLAPP develop their own in-house TTW equivalent.
- Explore external (e.g. cloud computing) opportunities for system infrastructure (processing, storage, serving, etc.) implementation and maintenance.