Framework Recommendations for a Statewide ASBS Information Management System

Lisa Hazard, Jennifer Bowen, Eric Terrill
Coastal Observing R&D Center
Scripps Institution of Oceanography
La Jolla, CA 92093-0213

1.0 Information Management Overview

In recent years there has been an awakening to the need for integrated information management systems to provide efficiency in assessing and managing regulatory programs. The statewide network of Areas of Special Biological Significance (ASBS) is one example in which a robust and persistent data system is required. A large amount and wide variety of data have been, and will be, collected in the watershed and ASBS through both regulatory permitting requirements and ancillary data collection efforts required to assess ASBS performance. Currently, these datasets are relatively isolated and unavailable to a wide range of users. Information management systems are needed for integration and public data dissemination so that interrelated biological-physical-chemical processes present in the watershed and marine environment can be assessed. These data requirements span both regulatory and non-regulatory based data collection efforts.

The information management system developed for this Plan, and described below, was designed to meet the following project needs:

- Data collection and storage
- Analysis and evaluation by the professional, policy making and regulatory community to assess the performance of the ASBS
- Data availability to the general scientific community
- Dissemination to the public for outreach and stewardship

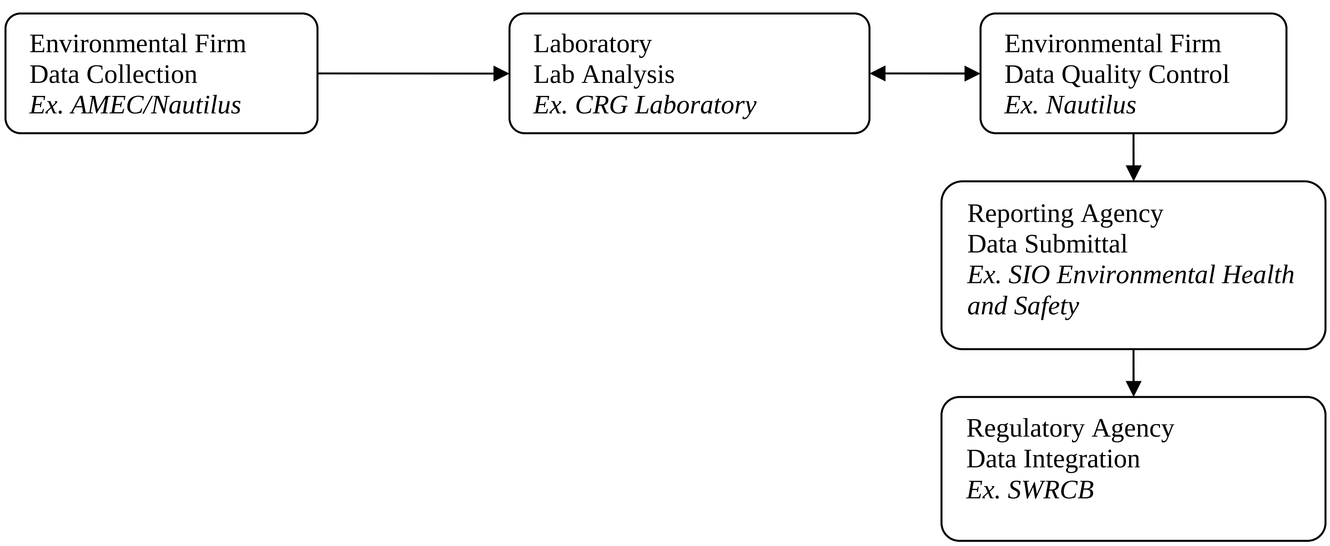
A distinction is made between *information management* and *data management*. Data management consists primarily of the “back-end” system (or network of systems) for data collection, ingestion, storage, archival, and retrieval. A robust data management system should consist of a tested and reliable method of acquiring data, a scalable and accessible method of storing and retrieving data, and a secure and replicated method of archiving data. Information management is the process by which the data becomes useful to decision makers. It includes the mechanisms for utilizing the data, optimized methods for disseminating the data, and the generation and presentation of useful products that can be used for research and decision making. Information management facilitates the transition from content (data) to knowledge.

The goal of the ASBS information management system is to establish the infrastructure needs and generate a conceptual design required for long term assessment of ASBS performance and related management decisions. The infrastructure needs to meet both the needs of the regulatory data collection as well as incorporate monitoring activities, scientific studies, and observations that are required for enhanced ecosystem assessment and ASBS management, yet may be outside of the present regulatory framework.

This document details recommendations created by the Coastal Observing R&D Center at Scripps Institution of Oceanography (SIO) to establish a comprehensive statewide framework of an ASBS informational management system, and pilot activities undertaken to implement a foundation for that system. The report is organized in the following manner. Section 2.0 reviews the existing data management systems that were identified as having relevance to the ASBS network. Section 3.0 Summarizes the ASBS functionality of the SWAMP data system. Section 4.0 Identifies and discusses those critical variables that are required to assess the processes which influence ASBS performance, yet are outside of the present regulatory framework. Section 5.0 Provides a discussion for requirements, limitations, and technical trade-offs of data display functions. Section 6.0 Provides recommendations for next steps.

2.0 Data Systems Relevant to an ASBS

An assessment of water quality data systems was conducted through a series of meetings with personnel involved in water quality data handing and through online research. To illustrate the complexities and challenges involved in managing these data, it is useful to summarize a typical chain of custody for sample data collected in the La Jolla ASBS. Sampling and analysis is contracted out by most agencies required to report on water quality. The samples are often collected by an environmental engineering firm then sent to a laboratory for analysis. The lab reports results back to the engineering firm for quality control, who in turn reports final results to the contracting agency. Agencies are responsible for reporting measurements to the State Water Resource Control Board (SWRCB). SIO Environmental Health and Safety (EH&S) utilize AMEC Consulting Firm for seawater sampling and MACTEC Engineering and Consulting, Inc. for storm water monitoring. The City of San Diego contracts to Weston Solutions, Inc. for all monitoring efforts. Reporting mechanisms and methods vary between all companies. The primary method of data storage is in spreadsheets (.xls, .csv) and primary method of data transfer is through email or paper reports. In terms of data scalability, query and web display, these practices are not sustainable. No standardization exists between the internal methods of data storage. Below is an example data flow chart, notice there is currently no public display or dissemination.



There have been several Regional and State efforts directed at establishing a standard protocol for water quality measurements. Relevant programs include CIWQS and SWAMP.

California Integrated Water Quality System Project (CIWQS)
This project is defined as:

> "a new computer system for the State and Regional Water Quality Control Boards to track information about places of environmental interest, manage permits and other orders, track inspections, and manage violations and enforcement activities. CIWQS also includes an electronic Self Monitoring Report (eSMR) tool for submission of monitoring reports via an internet web site. CIWQS is part of an overall effort to integrate several disparate legacy systems, compile water quality data, standardize permits, automate processes, and to make data more accessible to State Water Board staff, dischargers, the public, and the U.S. Environmental Protection Agency." (http://www.swrcb.ca.gov/ciwqs/index.html)

The electronic Self Monitoring Report (eSMR) is currently tailored for National Pollutant Discharge Elimination System (NPDES) Permit dischargers. The system is tuned for limited reporting requirements, and eventually plans to include Sanitary Sewer Overflow (SSO) and Storm Water Annual Reporting Module (SWARM) permitting. Because this system is streamlined for single sample type (water), it cannot currently accommodate other data types or batch sampling without significant modification and, therefore, can not be recommended as a comprehensive data management system. However, these specific applications do play an important role in data management and have been implemented within the Southern California Coastal Ocean Observing System (SCCOOS).

CIWQS summary: streamlined, parameter specific data system
Pros: efficient, simple, easily implemented
Cons: limited in scope, non scalable, not easily adapted for ASBS requirements

The Southern California Coastal Ocean Observation System (SCCOOS)
SCCOOS was established by a consortium of research organizations that extends from Northern Baja California in Mexico to Morro Bay at the southern edge of central California, and aims to streamline, coordinate, and further develop individual institutional efforts by creating an integrated, multidisciplinary coastal observatory in the Bight of Southern California to provide data and information primarily for the benefit of society.

SCCOOS aims to integrate a broad suite of observations to include but not limited to: surface currents, satellite imagery, wave conditions and forecasts, meteorological conditions and forecasts, water quality, ocean temperature, salinity, chlorophyll, and density in the form of products and raw data. The SCCOOS data management team has developed a number of innovative data interfaces and products, leveraging google maps to provide localized, zoomable, and navigable interactive display of data. This effort allows scientists, decision makers, and the public access to products that will provide a scientific basis for research, management, and improved uses of the ocean environment.

Targeted architectures are currently in use within the Southern California Coastal Observing System (SCCOOS) data management system. For example, SCCOOS ingests data from Environmental Health agencies throughout Southern California for public display of water quality data. The data are saved in a simple relational database. Although, the SCCOOS data management team did transform the excel files into a UNIX/LINUX-based MySQL database for measurement number scalability, the database is tuned specifically for water quality measurements alone and does not contain controlled vocabularies or outside observational fields. SCCOOS partnered with Southern California Coastal Water Research Project (SCCWRP) Information Systems Manager, Larry Cooper, to create a transfer mechanism and format based on the CIWQS NPDES data format for bacteria data from bight wide Environmental Health Agencies. These observations can be found online at: http://www.sccoos.org/data/waterquality/ Participating agencies include:

- Santa Barbara County; Environmental Health Services
- Ventura County, Department of Environmental Health
- Los Angeles County, Department of Public Health, Environmental Health
- City of Long Beach, Health & Human Services
- Orange County, Health Care Agency, Water Quality Department
- San Diego County, Department of Environmental Health.

Agencies are able to submit measurements to SCCOOS data managers located at Scripps through an enabled macro on their reporting spreadsheet. Data is emailed to a specified water quality address at SCCOOS where the attachment is then parsed into a database through an automated process. By simplifying the reporting process and automating data ingestion, SCCOOS data managers are able to maintain the data flow independent of user intervention. This system is sufficient for storing time series of water quality measurements of total coliforms, fecal coliforms, and enterococci at given station locations. However, as an ASBS data management system, this standalone water quality system is limited in scope and scalability, yet it demonstrates the capability for integrating spatially dispersed, mandated data sets into a unified system that has the potential for integration with other variables of interest.

SCCOOS summary: regional observational data system
Pros: includes non-regulatory observational parameters and limited regulatory data
Cons: not fully developed for integrated ASBS regulatory and non-regulatory parameters

Surface Water Ambient Monitoring Program (SWAMP)
SWAMP is defined as:

> "SWAMP is a statewide monitoring effort designed to assess the conditions of surface waters throughout the state of California. The program is administered by the State Water Board. Responsibility for implementation of monitoring activities resides with the nine Regional Water Quality Control Boards that have jurisdiction over their specific geographical areas of the state. Monitoring is conducted in SWAMP through the Department of Fish and Game and U.S. Geological Survey master contracts and local Regional Boards monitoring contracts. SWAMP also hopes to capture monitoring information collected under other State and Regional Board Programs such as the State's TMDL

(Total Maximum Daily Load), Nonpoint Source, and Watershed Project Support programs. SWAMP does not conduct effluent or discharge monitoring, which is covered under National Pollutant Discharge Elimination System permits and Waste Discharge Requirements."
(http://www.waterboards.ca.gov/swamp/index.html)

The SWAMP data management system is more comprehensive than CIWQS, including lookup tables for varying sample types, preparation methods, and collection methods. The SWAMP data structure also includes documented templates for lab entry and backend storage relationships. Due to the complexity of the SIO permit requirements, the limited scope of CIWQS, and the growing use of SWAMP throughout the state, data managers determined integration of the SWAMP structured system would be a preferred method for data storage, retrieval, and display of ASBS regulatory data.

SWAMP summary: comparatively matured water quality monitoring data system
Pros: comprehensive, collaborative, and becoming a standard
Cons: complex and still under development, developed for focused site

3.0 SWAMP Data System Details

The SWAMP data system is examined in more detail as it was identified to be the system which most closely matches the regulatory needs for the ASBS.

SWAMP is comprised of several modules. The backend or database contains approximately 38 tables each with dependencies or relationships with other tables. Lookup tables consist of static information which can relate to various measurements. Examples of such a table include agency information (name, address, contact, email, telephone, etc.); station information (name, address, latitude, longitude, county, water body type, etc.); and analyte description (name, number, group, description, etc.). Lookup tables do not always describe a physical location, contact, or parameter but could also reference units, qualifiers, or codes. Lookup tables limit field vocabulary avoiding data entry errors such as misspellings, capitalization differences, or invalid data types. They also optimize database functionality and size as each data result does not have to include full station, agency, sample type information reducing duplicity within the database. The backend contains the complete set of data with relationships between tables joining related information. For example, the results table would contain a station ID along with time of observation and observation value, the station ID field would be joined to the station ID field in the station ID lookup table. The station ID lookup table would contain information regarding that particular station as previously described. Entry into the data system requires input from data collectors taking measurements and field observations. Within the SWAMP system, data entry to the backend is facilitated through excel spreadsheets.

Because the SWAMP data system handles several different data types and is quite complex, data entry into the system is not trivial for an individual lab handling one data type. Templates have, therefore, been generated for simplification and ease of data entry. Templates for chemical, toxicity, and station entry are developed for data entry. Each template consists of a Results worksheet including all of the fields necessary to describe that particular parameter. Subsequent worksheets within the excel file consist of lookup tables with related ID's for field entry within the main Results tab. There are actually very few fields within the Results worksheet that are not correlated to a lookup table containing a controlled vocabulary. Again, this helps facilitate ambiguity and variability between data entry personnel. Consistency between observation reporting is essential for future analysis and comparison across parameters, time periods, analytes, etc. For a given data type, the templates contain full relationships and input fields. Completed templates must then be ingested into the backend database. Ingestion can be automated through programmed parsing scripts. The scripts will read template files, strip out values and load into the appropriate tables within the backend. Once new values are entered into the system, they will be queried and displayed on the website. The following is a graphical display of bacteriological data ingestion from lab submittal to web display presented in steps.

1.) Submit SWAMP compatible template to information management electronically

LabSampleID	StationCode	EventType	SampleDate	SampleTime	SampleTypeCode	SampleReplicate	DepthSampleCollection	DepthUnit	ProjectID	Season	AgencyCode	FailureReason	SampleComments	Preparation	PreparationDate	DigestExtractMethod	DigestExtractDate	LabBatch	AnalysisDate	SampleID
1077	000SIO003	WaterChem	2005-10-04	7:55 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-10-04		1950-01-01	CityLab_00078_W_BAC	2005-10-04	1026
1078	000SIO003	WaterChem	2005-10-13	10:16 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-10-13		1950-01-01	CityLab_00079_W_BAC	2005-10-13	1027
1079	000SIO003	WaterChem	2005-10-20	10:40 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-10-20		1950-01-01	CityLab_00080_W_BAC	2005-10-20	1028
1080	000SIO003	WaterChem	2005-10-26	10:20 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-10-26		1950-01-01	CityLab_00081_W_BAC	2005-10-26	1029
1081	000SIO003	WaterChem	2005-11-02	8:20 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-02		1950-01-01	CityLab_00082_W_BAC	2005-11-02	1030
1082	000SIO003	WaterChem	2005-11-09	8:10 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-09		1950-01-01	CityLab_00083_W_BAC	2005-11-09	1031
1083	000SIO003	WaterChem	2005-11-16	8:50 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-16		1950-01-01	CityLab_00084_W_BAC	2005-11-16	1032
1084	000SIO003	WaterChem	2005-11-21	7:56 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-21		1950-01-01	CityLab_00085_W_BAC	2005-11-21	1033
1085	000SIO003	WaterChem	2005-11-30	8:25 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-30		1950-01-01	CityLab_00086_W_BAC	2005-11-30	1034
1086	000SIO003	WaterChem	2005-12-07	8:40 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-12-07		1950-01-01	CityLab_00087_W_BAC	2005-12-07	1035
1087	000SIO003	WaterChem	2005-12-14	8:40 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-12-14		1950-01-01	CityLab_00088_W_BAC	2005-12-14	1036
1088	000SIO003	WaterChem	2005-12-21	9:06 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-12-21		1950-01-01	CityLab_00089_W_BAC	2005-12-21	1037
1089	000SIO003	WaterChem	2005-12-28	8:50 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-12-28		1950-01-01	CityLab_00090_W_BAC	2005-12-28	1038
1090	000SIO003	WaterChem	2005-10-04	7:55 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-10-04		1950-01-01	CityLab_00091_W_BAC	2005-10-04	1026
1091	000SIO003	WaterChem	2005-10-13	10:16 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-10-13		1950-01-01	CityLab_00092_W_BAC	2005-10-13	1027
1092	000SIO003	WaterChem	2005-10-20	10:40 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-10-20		1950-01-01	CityLab_00093_W_BAC	2005-10-20	1028
1093	000SIO003	WaterChem	2005-10-26	10:20 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-10-26		1950-01-01	CityLab_00094_W_BAC	2005-10-26	1029
1094	000SIO003	WaterChem	2005-11-02	8:20 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-02		1950-01-01	CityLab_00095_W_BAC	2005-11-02	1030
1095	000SIO003	WaterChem	2005-11-09	8:10 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-09		1950-01-01	CityLab_00096_W_BAC	2005-11-09	1031
1096	000SIO003	WaterChem	2005-11-16	8:50 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-16		1950-01-01	CityLab_00097_W_BAC	2005-11-16	1032
1097	000SIO003	WaterChem	2005-11-21	7:56 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-21		1950-01-01	CityLab_00098_W_BAC	2005-11-21	1033
1098	000SIO003	WaterChem	2005-11-30	8:25 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-30		1950-01-01	CityLab_00099_W_BAC	2005-11-30	1034
1099	000SIO003	WaterChem	2005-12-07	8:40 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-12-07		1950-01-01	CityLab_00100_W_BAC	2005-12-07	1035
1100	000SIO003	WaterChem	2005-12-14	8:40 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-12-14		1950-01-01	CityLab_00101_W_BAC	2005-12-14	1036
1101	000SIO003	WaterChem	2005-12-21	9:06 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-12-21		1950-01-01	CityLab_00102_W_BAC	2005-12-21	1037
1102	000SIO003	WaterChem	2005-12-28	8:50 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-12-28		1950-01-01	CityLab_00103_W_BAC	2005-12-28	1038
1103	000SIO003	WaterChem	2005-10-04	7:55 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-10-04		1950-01-01	CityLab_00104_W_BAC	2005-10-04	1026
1104	000SIO003	WaterChem	2005-10-13	10:16 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-10-13		1950-01-01	CityLab_00105_W_BAC	2005-10-13	1027
1105	000SIO003	WaterChem	2005-10-20	10:40 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-10-20		1950-01-01	CityLab_00106_W_BAC	2005-10-20	1028
1106	000SIO003	WaterChem	2005-10-26	10:20 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-10-26		1950-01-01	CityLab_00107_W_BAC	2005-10-26	1029
1107	000SIO003	WaterChem	2005-11-02	8:20 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-02		1950-01-01	CityLab_00108_W_BAC	2005-11-02	1030
1108	000SIO003	WaterChem	2005-11-09	8:10 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-09		1950-01-01	CityLab_00109_W_BAC	2005-11-09	1031
1109	000SIO003	WaterChem	2005-11-16	8:50 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-16		1950-01-01	CityLab_00110_W_BAC	2005-11-16	1032
1110	000SIO003	WaterChem	2005-11-21	7:56 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-21		1950-01-01	CityLab_00111_W_BAC	2005-11-21	1033
1111	000SIO003	WaterChem	2005-11-30	8:25 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-11-30		1950-01-01	CityLab_00112_W_BAC	2005-11-30	1034
1112	000SIO003	WaterChem	2005-12-07	8:40 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-12-07		1950-01-01	CityLab_00113_W_BAC	2005-12-07	1035
1113	000SIO003	WaterChem	2005-12-14	8:40 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-12-14		1950-01-01	CityLab_00114_W_BAC	2005-12-14	1036
1114	000SIO003	WaterChem	2005-12-21	9:06 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-12-21		1950-01-01	CityLab_00115_W_BAC	2005-12-21	1037
1115	000SIO003	WaterChem	2005-12-28	8:50 AM	Grab	1	0	m	06SI9000	Fall	CityLab			LabFiltered	2005-12-28		1950-01-01	CityLab_00116_W_BAC	2005-12-28	1038

ResultsExample | FieldObservations | FieldObsLookUp | StationsExample | LabBatchExample | Preparation | DigestExtractL

Figure 1. Results table from SWAMP compatible template

2.) Template is parsed and ingested into backend relational database

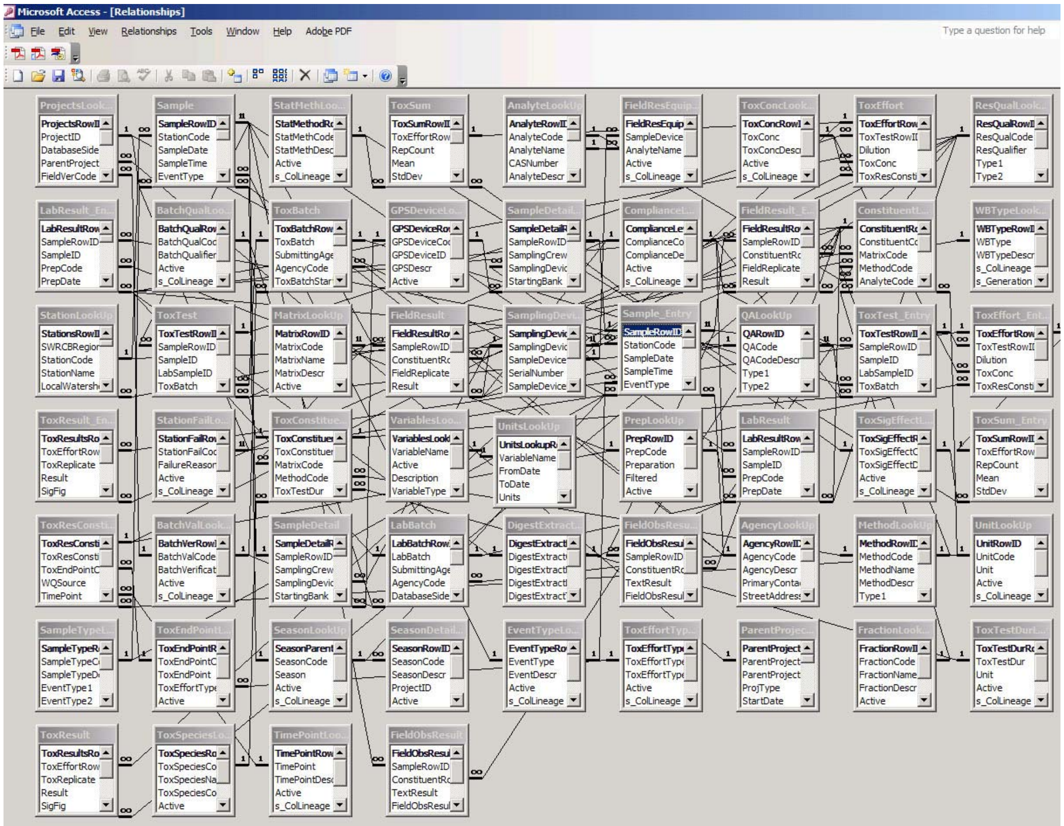


Figure 2. Relationship view of database taken from Access (operational database is in MySQL)

3.) Query database and Display to Web

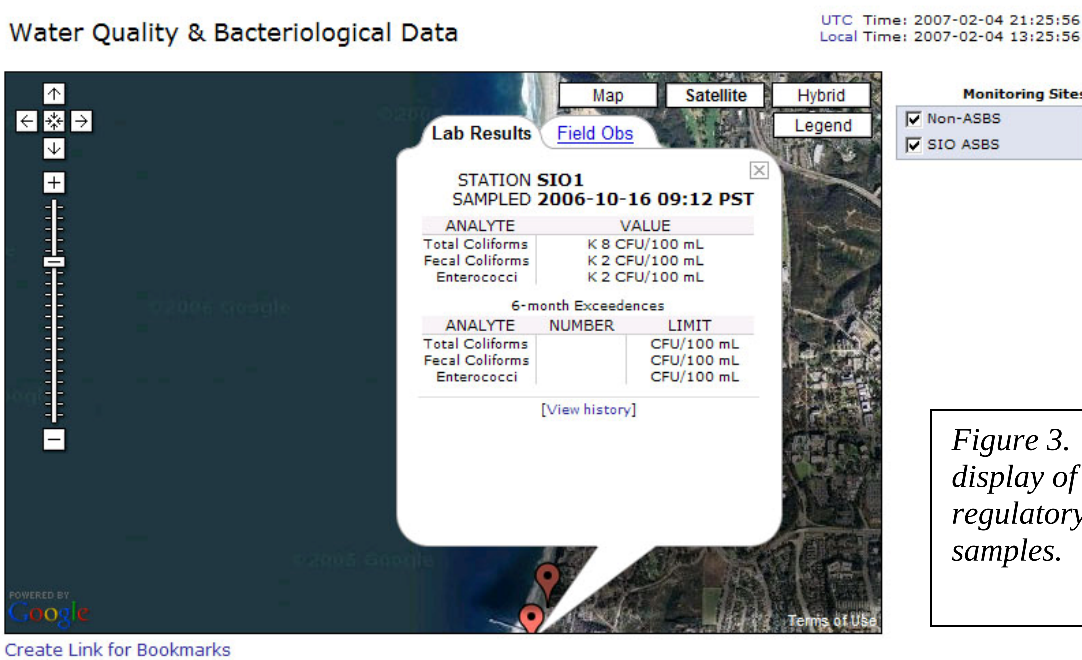


Figure 3. Online display of ASBS regulatory bacteria samples.

4.0 Non-regulatory data required to assess ASBS performance

There are several different types of data that must be collected to thoroughly and effectively manage the ASBS. These types include regulatory ecosystem management data, and supplemental environmental data. The data management team recommends adoption and modification of the SWAMP (Surface Water Ambient Monitoring Program) backend for regulatory ASBS data. Ecosystem and an expanded environmental data set will require a separate data system. Following ecosystem management and supplemental environmental data recommendations, those data systems will require design and development within the Implementation Grant. Background on recommended SWAMP configuration system details can be found at http://www.cordc.ucsd.edu/projects/asbs/

Designated Areas of Special Biological Significance (ASBS) along the coast of California exist in a complex coastal regime subject to ever-changing land-sea-atmospheric interactions. As a result, when evaluating the performance and behavior of an ASBS, it will be important to understand the physical environment both within and surrounding its environs. This regional description of the time-varying coastal environmental processes relevant to the ASBS will be critical to understanding ecosystem changes within the ASBS. A critical assessment question coastal zone managers face will be:

1. How to link trends and changes in the monitoring data to the management decisions made within the ASBS.
2. Assessing whether the observed changes are a result of climate/natural variability, or if external, anthropogenic influences are impacting the ASBS.

Attributes and processes deemed relevant to assessing the ASBS include:

- Local meteorological conditions including regional precipitation which influence the watershed bordering the ASBS.
- Location and size of nearby wetlands, the state of the wetland (entrance open/closed) and volumes of freshwater exchange with the coast.
- Time records of the flow of freshwater from streams, and rivers nearby the ASBS. The mouths of many of these sources may be lagoons and wetlands, or discharges outside of the ASBS.
- The ocean circulation both within and surround the ASBS. Observations of the circulation will allow estimates of the residence time of discharges within the ASBS, as well as provide insight on how non-ASBS discharges may influence the ASBS. Regional observations of the physical circulation can be used for tracking the fate and transport of both pollutants within the ASBS, but also provide data to assist with understanding biological connectivity within and across the ASBS boundaries. Coastal circulation will also influence the flux of nutrients into the ASBS through internal tidal surges and regional upwelling events.
- Sediment composition and transport processes within the ASBS are needed both for understanding transport of pollutants with the ASBS, and for assessing changes in migrating sediment levels which can influence the biota.
- Coastal ecosystems are sensitive to the temperature of the ocean. In addition, changes in temperature can also co-vary with nutrient levels. Ocean temperatures and stratification should be monitored continuously to assess changes on both tidal, daily, seasonal, and climate scales.

- The optical properties of water within and surrounding the ASBS should be monitored as underwater ecosystems depend on sunlight. The optical properties may covary with both biological productivity and concentrations of fine sediments.

A draft list of variables relevant to understanding these processes are provided below. Specifics related to the density of observations at individual ASBS sights must be determined after an appropriate assessment has been conducted. It is recommended that these observations be made near-continuously when technically feasible.

- Ocean stratification measurements (temperature and salinity)
- Ocean surface current maps both internal and external to the ASBS boundary
- Ocean current profiles (observations at depth)
- Ocean salinity
- Wave height and direction, modeled surfzone currents
- Bathymetry maps (repeated observation to document changes)
- Bottom type, grain size and substrate
- Local meteorology to describe local precipitation and upwelling favorable winds
- Flow rates for local estuaries and freshwater discharges (natural or otherwise) nearby the ASBS
- Time records of nutrients (nitrates, phosphates, silicates) within the ASBS

The SWAMP data system is designed and tuned for ASBS regulatory water quality data. Ecosystem and supporting environmental data necessary for full ASBS assessment must be integrated with the regulatory water quality, chemistry, toxicity, and field observations. SWAMP, however, is not the complete storage mechanism for these other data types. Each data type may require its own database or be saved as a series of time series files with supporting metadata. The joining components for integrating and comparing across multiple variables and/or data sets are location (latitude and longitude), elevation, and time. Those variables allow cross referencing in space and time. A complete ASBS data system must include both regulatory and non-regulatory data including biological, chemical, and physical attributes. The closest data system that meets the data management needs for these essential variables is within the ocean observing framework provided by SCCOOS.

5.0 Information Presentation

While backend data collection, storage, archival, and integration are the foundation of a data system; data display and dissemination also play an integral role in the full information management system. Data transforms to information when it is appropriately processed and presented in a clear, comprehensible manner. Data collected over long time series can show trends and highlight anomalies. Point measurements such as bacteriological samples collected at a given location can be displayed on a timeline putting recent observations into a longer time period perspective. Gridded data such as surface currents are better displayed on a map putting the vectors into context of the region. Data sets such as satellite ocean color and surface currents can be overlaid to help examine a relationship or correlation. In designing integrated products for web display user needs must be addressed.

Visualization tools for ecosystem assessment must be developed further in order to comprehensively analyze the ASBS in context with surrounding areas. The data management community is struggling with those visualization tools. Often times, a GIS tool is used for layering environmental data. Unfortunately, those tools prove to be sluggish and burdensome due to the shear volume of data processing required. Most servers either take a significant amount of time to display the data or cause an error upon retrieval of multiple layers. They also can not display time series of events such as a developing current field or bacteria results over the latest wet period. Standalone software programs such as google earth and fleudermaus provide an excellent visualization tool, but require local access to data sets. Serving capabilities have not yet been fully developed. Solutions to this problem include automated processing of established data products, database indexing for faster data retrieval, multiple processors within data servers, and increasing internet bandwidth. Developmental and technological advancements on these fronts require planning, engineering, and resources.

SCCOOS has developed a number of innovative data interfaces and products, leveraging google maps to provide localized, zoomable, and navigable interactive display of data. Providing data visually online is a powerful tool, enabling academics, decision makers, and the public easy access to public data. Users are able to download data values as well in an ascii tabular format. With manageable data sets, ascii download is often times the preferred method. It's easily understood and ingestible into an alternate analysis software package. Most spatial or Geographic Information System (GIS) data is less suited for such transfer methods and requires alternate formatting for data download. These types of data are far too voluminous for tab or common delimited files. The SIO data management team plans to display regulatory bacteria, toxicity, and chemical analysis data in a similar format to the tiled google display. The data management team has implemented improved data dissemination utilities through the use of recent web based technologies and mapping capabilities. Future data products can be integrated and designed based on user needs assessment and utility.

6.0 Recommendations

Recommendations for future data system development and management include defining changes within, adoption, and implementation of the SWAMP structure; development and design of a data system for ecosystem management; integration of environmental observational data; needs assessment with ASBS science and management community to define optimal data distribution, presentation, and analysis tools; and prototyping implementation of an end-end system in an ASBS to serve as a model for a statewide system.

Although SWAMP in its current form does not fulfill the entire suite of regulations, the system can serve as a building block for a comprehensive and transferable relational data management system for ASBS regulatory data. The SWAMP data management system was chosen over other data management systems because it is more comprehensive, including lookup tables for laboratory contacts, station ID, units, analytes, methods, etc. and the need for statewide compliance and compatibility. The SWAMP system is not a single solution data system for all required ASBS assessment measurement parameters. A fully functional information management system can be considered a system of systems. Much of the ecosystem management and environmental observational data will also need to be saved in formats which give the flexibility needed for examining multidisciplinary processes. Required spatial and temporal cross referencing attributes include latitude, longitude, elevation, and time. These attributes enable efficient data integration, analysis, and visualization.

Future efforts should involve expertise found at Moss Landing Marine Laboratory who are currently expanding the SWAMP data system. The system should consist of a relational database within a unix/linux operating system environment (e.g. mysql, postgres, oracle). For security purposes data management best practices should be stored on a system with backup capabilities. Ideal programming would include some sort of redundant array of independent disks (RAID) and offsite backup utility. Finally, reasonable products and public display is an essential component of the information management system. Management and assessment of the ASBS extends far beyond collection, ingestion, and display of regulatory data. Integration of physical and biological data is necessary for full ecosystem analysis. Web based data presentation and dissemination will allow the interrelationships of these datasets to be examined over space and time. Visualization methods should be leveraged from the SCCOOS information management system for dissemination and display. The future of the ASBS information management system should include an iterative implementation method whereas the system is designed, tested, and then improved based on performance, reliability, and comprehensiveness.

Implementation Goals and Objectives

- Design and implement a robust and scalable data management system for storage, archival, retrieval, dissemination, and display of regulatory data leveraging from the SWAMP data system.
- Determine undefined attributes necessary for realizable ecosystem assessment of ASBS.
- Begin integration and aggregation of biological and physical data based on location (latitude, longitude), time, and elevation leveraging existing data sets within SCCOOS.
- Following needs assessment with ASBS science and management community, display ASBS data in an organized and digestible format easily accessible to scientists, decision makers, and the general public.
- Create an iterative management process for continued improvement and regional integration.