NERACOOS DMAC Plan

June 12, 2017

1 Introduction

The Integrated Ocean Observing System (IOOS) is a national network of Regional Associations working to provide ocean observing data and information to a wide variety of end users. As the Regional Association (RA) for the Northeast, NERACOOS has a primary goal to aggregate and provide access to ocean observing data from the region. NERACOOS supports a robust Data Management And Communications (DMAC) program to support this function.

The IOOS Program Office and key data management representatives from all of the regional associations work closely to develop, test and implement standards and recommendations for managing and distributing data. NERACOOS' DMAC team, led by GMRI, has participated in these efforts since the outset of IOOS. The NERACOOS DMAC team has developed and continues to evolve a robust and efficient data management system that leverages the suite of standards and best practices adopted and recommended by IOOS and the DMAC community to provide a centralized source for regional data access.

2 DMAC Management Structure

The NERACOOS DMAC management structure includes the NERACOOS Executive Director, NERACOOS Outreach and Communications Specialist, Gulf of Maine Research Institute (GMRI) DMAC Lead and the GMRI Data Manager. The NERACOOS Executive Director (Ru Morrison) provides strategic guidance and oversight for DMAC activities. NERACOOS Communications and Outreach Specialist (Tom Shyka) works directly with the GMRI DMAC team, provides oversight for DMAC activities, coordinates communications between NERACOOS and the GMRI DMAC lead, and helps coordinate activities of the funded NERACOOS data providers.

The GMRI DMAC Team is lead by co-PI Riley Young Morse who manages all NERACOOS DMAC activities and oversees the GMRI Data Manager. Additionally, the DMAC Lead contributes to proposal development and general NERACOOS data management reporting requirements. The DMAC Lead is the main point of contact for all technical data-related questions and is an expert in the design and development of data products. The DMAC Lead submits quarterly reports to NERACOOS Executive Director for DMAC operations and website services. The DMAC lead also participates in a semi-annual performance review conducted by NERACOOS (See section 1.3.3.4 of the NERACOOS SOP).

The GMRI Data Manager (Eric Bridger) oversees the NERACOOS data management system; the ingestion of data from funded data providers and other data sources, the management of data within the data system, data and system back-up, development and maintenance of data access services. The Data Manager is the main POC with the DMAC staff from the U.S. IOOS Program Office and with the NERACOOS funded data providers. The Data manager participates in IOOS data standards development and testing activities and supports NERACOOS funded data providers to ensure that IOOS recommended standards are implemented and maintained.

The GMRI Data Manager maintains regular communication with the U.S. IOOS Program Office through in-person meetings, phone calls, webinars, and emails. Additionally, The GMRI DMAC lead and Data Manager often consults outside DMAC expertise from our sister RAs (or RICEs) and other federal partners with DMAC expertise. This frequent communication ensures that NERACOOS is aware of all new practices and protocols, as promulgated by the IOOS Program Office, and understands how to implement them. The GMRI data manager, DMAC lead, NERACOOS staff and the funded data providers participate in monthly DMAC calls to discuss new protocols and evaluate implementation of protocols recommended by the U.S. IOOS office.

Whenever the U.S. IOOS Program Office provides guidance on data management protocols, NERACOOS will respond within 1 month with an assessment of the relevance of such guidance to our DMAC procedures and if appropriate, an estimate of the time it will take us, given resources and capacity, to reach compliance. Once the data management lead receives the recommended protocol, they will take the necessary steps towards its implementation in a reasonable and timely manner. Implementation of new services is only limited by personnel time and expertise. If feasible, NERACOOS will implement new protocols in one year.

The GMRI DMAC team (Team Lead and Data Manager) are responsible for all aspects of the NERACOOS Data Management System as described above as well as the design operations and maintenance of the NERACOOS website and products and services provided within the website. The GMRI DMAC team works directly with NERACOOS funded data providers to support and ensure the implementation of data standards and the maintenance and of data feeds. The GMRI DMAC team works closely with the NERACOOS outreach and communications specialist and NERACOOS stakeholders to ensure that products and services meet well-defined end user requirements.

3 Data Sources

The NERACOOS Data Management System (DMS) aggregates observational data streams from providers distributed throughout the region. Provider types are generally organized into NERACOOS funded (see Section 6 for details), federal or state assets, and private research assets and are outlined in Appendix I: NERACOOS Data Source Table. It

should be noted that the classification is based on the management of the assets, as a mix of funding from the sources mentioned above supports some assets. Observational data are organized into type (real-time, near real-time and historic) and are described below. NERACOOS also integrates model data from funded partners and federal sources. Model data are not fully described in the data source assessment but are listed in the data source table.

The NERACOOS DMS was deployed in 2009 leveraging the data management processes and data access products developed by the Gulf of Maine Ocean Observing System (GoMOOS) between 2000-2008. The core system, still in place, had been optimized for the primary data provider, the University of Maine (UMaine). The system has evolved to include additional funded data providers and non-NERACOOS funded data providers in the region.

3.1 Observational Data

3.1.1 Real-time data

Real-time data are largely provided through buoys and sensors that are remotely deployed and autonomously report on regular intervals. These data are typically telemetered from the sensors directly to the data provider where automated QC is applied and real-time data files are produced and distributed. From there, the real-time data streams are then ingested and processed by NERACOOS DMS at regular intervals ranging from 10 minutes to hourly.

3.1.2 Near real-time data

Near real-time data are data that have a lag in availability, usually due to processing or preliminary analysis. The data that fall into this category include MODIS and AVHRR satellite imagery data of the Northeast region. These data are processed daily by UMaine and available through NERACOOS on a several day lag. The daily and monthly summary climatology data from the NERACOOS buoys are also processed on a lag and are made available the following day (daily files) or the first day of the next month (monthly files).

3.1.3 Historical data

Historical data include the full time series from buoy deployments. As buoys and sensors are recovered during redeployment, the data are processed and updated for QA/QC, then provided to NERACOOS as the official historical data record for the deployment.

3.1.4 Model data

Model data include forecast products from federal as well as NERACOOS funded partners. These data are accessed at regular intervals as they are available and are typically integrated into several products available through the data portal as time series data as well as imagery.

3.2 Provider Type

3.2.1 NERACOOS Funded

NERACOOS fully and partially funds the collection of a significant amount of data in the region. The funded data streams undergo the full data management process and must meet basic standards requirements for metadata, file format and QA/QC. The NERACOOS DMAC team works closely with data providers to ensure they have the tools and guidance needed to produce regular, quality real time and historic data. Details of the NERACOOS funded data providers processes are provided in Section 6.

3.2.2 Federal and State

NERACOOS also ingests and displays data from sources not funded by NERACOOS but readily available and relevant to regional partners. These include assets from federal and state providers. While these data are not required by NERACOOS DMAC to meet the requirements of our funded data streams, most of the federal data are made available through standard web services. These data undergo QA/QC procedures required by the Federal or State agency that collects the data. Providers of data in this category include NOAA NDBC, NOAA NEFSC, NOAA CO-OPS, NOAA NERRS, CDIP and USGS.

3.2.3 Private Research

Data in this category may include streams from academic institutions, private industry, and NGOs. These data are not always continuous and generally aren't available in standard services like NERACOOS or federal assets. They can also tend to be shorter-term deployments due to limited project funding. Examples of these data include WHOI's ESP harmful algal bloom data, University of Maine EPSCoR buoys, CINAR TEMPESTS buoys (a 3 year storm intensity modeling and observation study). NERACOOS endeavors to make these data available to the region and works with the individual researchers to accommodate data as time and funding allow. Most of these data are ingested into our ERDDAP data system and are not included in data products.

Additional detail about the various NERACOOS data sources can be found in the Data Source Table (Appendix I).

4 Quality Control

Delivering high quality data is a key tenant of the NERACOOS mission. The quality assurance procedures that NERACOOS and its funded data providers have implemented help ensure that data delivered by NERACOOS is as accurate and precise as possible.

As described above in the Data Source section, NERACOOS ingests data from various federal agencies and these data streams have undergone QA conducted by those federal agencies. NERACOOS also funds a significant amount of data collection by several non-federal partners, which are primarily academic institutions. The Principal Investigators (PI) and the teams that are selected to operate sensors for NERACOOS are

recognized experts in oceanographic data collection. NERACOOS funding agreements require that funded data providers implement QA procedures on all data streams and require that the relevant QARTOD tests be implemented on real-time data streams by May 31, 2018.

4.1 Sensor Calibration and Operation

The proper calibration and operation of sensors is a key component to data quality assurance. The NERACOOS funded data partners follow manufacturer's operational procedures to ensure proper calibration and operation of sensors.

4.2 Data Provider QA Processes

NERACOOS funded data providers are responsible for conducting quality assurance on data that is delivered to NERACOOS. The quality assurance procedures implemented by the funded data partners are described for each data provider in Section 6. The majority of funded partner data streams undergo a robust QA process that includes many of the required QARTOD tests as well as other QA tests and reviews. All non-federal sourced data served by NERACOOS undergoes at least a minimum level of QA.

4.3 QARTOD

NERACOOS and its funded data partners are in the process of a phased implementation of required QARTOD tests on real-time data stream. All data providers are performing some of the required QARTOD tests on the majority of the data streams and are working on implementing the required tests on all real-time data streams. NERACOOS anticipates that implementation of required QARTOD test will be completed by May 2018.

4.4 NERACOOS DAC QA Procedures

Real time data streams undergo automated QA by the data providers to catch missing or out of range readings. These readings are flagged with ‘non-zero’ codes in the feeds that are ingested by the NERACOOS DMS and have different meanings as described in the QA/QC procedures and in the metadata. Historical files served by data providers often contain additional post-deployment QA/QC flags. For real-time data feeds, readings with non-zero QA/QC flags are not displayed in any NERACOOS data products. Historic data available through download tools (e.g. ERDDAP, THREDDS) contain the full data set, including readings with non-zero QA/QC flags, and flag values are described in the metadata record.

5 Data Management System

The IOOS Program Office and key data management representatives from all of the regional associations work closely to develop, test and implement standards and recommendations for managing and distributing data. NERACOOS’ DMAC team, led by GMRI, has participated in these efforts since the outset of IOOS. The NERACOOS DMAC team has developed and continues to evolve a robust and efficient data management

system that leverages the suite of standards adopted and recommended by IOOS and the DMAC community to provide a centralized source for regional data access.

The NERACOOS Data Management System (NERACOOS DMS) is a framework for aggregation, interoperability, discovery and dissemination of observing data (gridded and observational) from the region and is the backbone of the NERACOOS Data Assembly Center (DAC). The GMRI team leverages over fifteen years of open-source software development and data management expertise directly on IOOS initiatives in the region. The resulting NERACOOS DMS is based on the suite of standards and best practices developed, tested and recommended by IOOS and the greater IOOS RA DMAC community. The NERACOOS DMS provides the mechanisms and protocols for the full data lifecycle by integrating, aggregating, and distributing data through a centralized access portal (NERACOOS.org). Following the IOOS DMAC Guiding Principles, the DMS provides the core capacity that connects and integrates observations and forecasts, making quality-controlled data discoverable and accessible to stakeholders through a wide variety of information products. The NERACOOS DMAC team continues to work closely with the IOOS Program Office and DMAC community to implement new requirements and protocols developed to further enhance access to data.

Aggregating and providing central access to data streams from the NERACOOS funded sensors and models was the primary consideration in the development of the NERACOOS DMS. The DMAC team works closely with NERACOOS data providers to ensure that the highest QA/QC data (historical and real-time) are available in standardized formats for all partners and accessible for products and services through the NERACOOS Data Portal. A secondary goal is leveraging the NERACOOS DMS to include external data streams from state, federal and other independent providers (e.g. academic institutions, private entities, etc.) and make accessible through the same products and services. Interest in adding new data streams has been steadily on the rise and the NERACOOS DMS has enabled this to be a much more efficient process than in the past.

The NERACOOS DMS is entirely cloud based using the Amazon Elastic Compute Cloud service (Amazon EC2). The NERACOOS DMS server instances are deployed on a virtual private cloud behind a private security group for maximum cost efficiency and security. The system is backed up nightly, preserving the data, software and configuration of the machine setup for guaranteed disaster recovery. This cloud-based system has now been in place for over 5 years, and the flexibility has enabled several hardware and software upgrades to accommodate the ongoing evolution of the NERACOOS DMS. At this time, the NERACOOS DMAC team has been able to implement the DMS; following best practices and fully or partially adopts most of the guiding principles put forward by NOAA and IOOS (Figure 1).

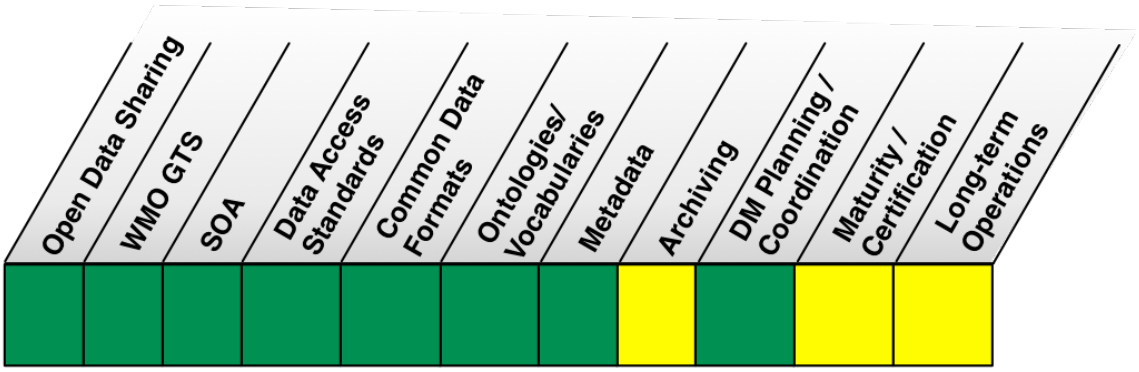


Figure 1: Status of Implementation of DMAC Subsystem Guiding Principles. Green = fully adopted, yellow = partially adopted.

This document describes the NERACOOS DMS in detail and is organized into the following sections:

1. System Architecture
2. Data Requirements
3. Data Access and Services
4. Hardware and Software Configuration

5.1 System Architecture

The NERACOOS DMS is comprised of four major components that manage the acquisition, integration, discovery and dissemination of data from the region. Using interoperability best practices and IOOS recommended standards; data are ingested via several paths to accommodate data providers with differing levels of technology capacity and to maintain legacy data flow during the conversion to the new system. Once ingested into the NERACOOS DMS, data are stored in one of several ways based on data types. The data are immediately made available via standards-based services for access.

Data are acquired directly from the data providers through a Service Oriented Architecture (SOA) approach. By enabling data providers to expose data through standards-based web services (e.g. direct access to data files through ftp or web services or connection to local Thematic Real-time Environmental Distributed Data Services - THREDDS), redundancy and replication is reduced, therefore providing the highest quality data to end-users. The NERACOOS DMAC team works closely with the data providers to produce data in compliant format and adopt data and quality control standards and protocols. As a result, the data managed and curated through the DMS are interoperable and aggregated in region-wide products, though the data are served from distributed systems. All of the available web services have been registered with the IOOS Catalog since its inception, and are also exposed through search-engine friendly Web Accessible Folders (WAF). For data providers lacking bandwidth or capacity to serve data reliably or in compliant formats, the DMS has the capacity to ingest and store these data, and serve through standardized formats. The protocols and services of the

NERACOOS DMS provide a roadmap to integrate new data providers quickly and efficiently (Figure 2) [1].

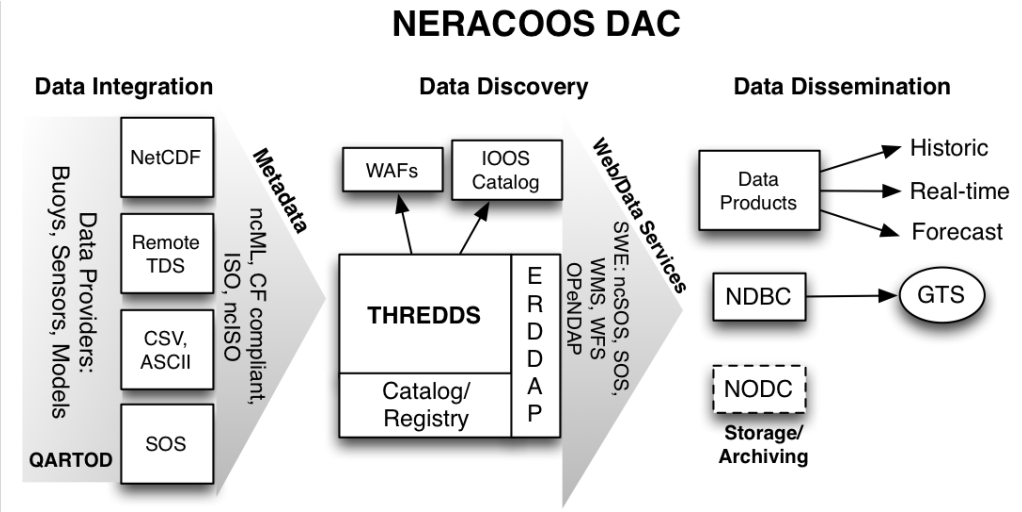


Figure 2: NERACOOS DAC schematic. The data managed and curated through the DMS are interoperable and aggregated in region-wide products, though the data are served from distributed systems. The protocols and services of the DMS provide a roadmap to integrate new data providers quickly and efficiently.

---

[1] See DMF and Data Management Plan 2014 at http://neracoos.org/certification2016

5.2 Data Requirements

Many NERACOOS funded partners are currently capable of storing and serving data in standard compliant formats (e.g. netCDF) that can be ingested by the NERACOOS DMS using automated protocols and services. Guidance is provided to NERACOOS data providers so they can output their data in compliant formats to avoid duplication from additional storage where not necessary. For data providers lacking bandwidth or capacity to serve data reliably or in compliant formats, the DMS has the capacity to harvest, store and serve these data in standardized output formats.

The basic requirement for funded data providers to contribute data to the NERACOOS DMS is that that data are available in netCDF or via THREDDS connection and include appropriate metadata. For data providers not currently serving netCDF, efforts were undertaken to understand capacity and potential for serving netCDF (i.e. Python scripts to translate text to netCDF) and provide support to produce the necessary files. Metadata records are reviewed with data providers and guidance given on how to update records with missing or incomplete information.

5.2.1 Data Types

NetCDF serves as the basic data format supported for data output. Most observation data providers already use netCDF as their data storage format, or are capable of producing netCDF files, and the software implemented in the framework is capable of

processing data in this format.

NetCDF CF 1.6, the most recent release of the Climate Forecast convention, is required either in the netCDF headers or via TDS ncML plugin. The protocol to update files with ncML has been well documented and is typically set up once per provider. Updates aren't necessary unless major changes happen with the data collection protocols. Middleware is available to enable data providers to update metadata through ncML (Signell 2014). Additionally, Unidata's Common Data Model, in particular the Discrete Sampling Geometries for Point Time Series, was an important resource in mapping data from the formats used by data providers to be compliant with CF 1.6.

5.2.2 Metadata

The NERACOOS framework has and will continue to leverage existing work at national level with regard to cataloging and metadata and vocabulary standards. We currently utilize the ISO 19115 Metadata Standard and ncML to make data CF and ISO compliant. The data available through NERACOOS are catalogued locally using a Web Accessible Folder that provides up to date status information to the IOOS catalog crawler.

Additional vocabularies and standards to describe data utilize OGC SWE (SOS, SensorML metadata), CF, MMI, and discrete sampling geometries described in the UCAR netCDF Common Data Model and Java Library (http://www.unidata.ucar.edu/software/thredds/current/netcdf-java).

5.3 Data Access and Services

Data are made accessible from the NERACOOS DMS through machine-machine services and human readable products and tools described in Section 7. The data access and services through NERACOOS utilize IOOS recommended and Open Geospatial Consortium standards. Figure 3 illustrates the wide array of data sources, ingest services, storage and access services that are available through the NERACOOS DMS. Specific details about how each NERACOOS data provider flows data to NERACOOS DMS are outlined in section 6.

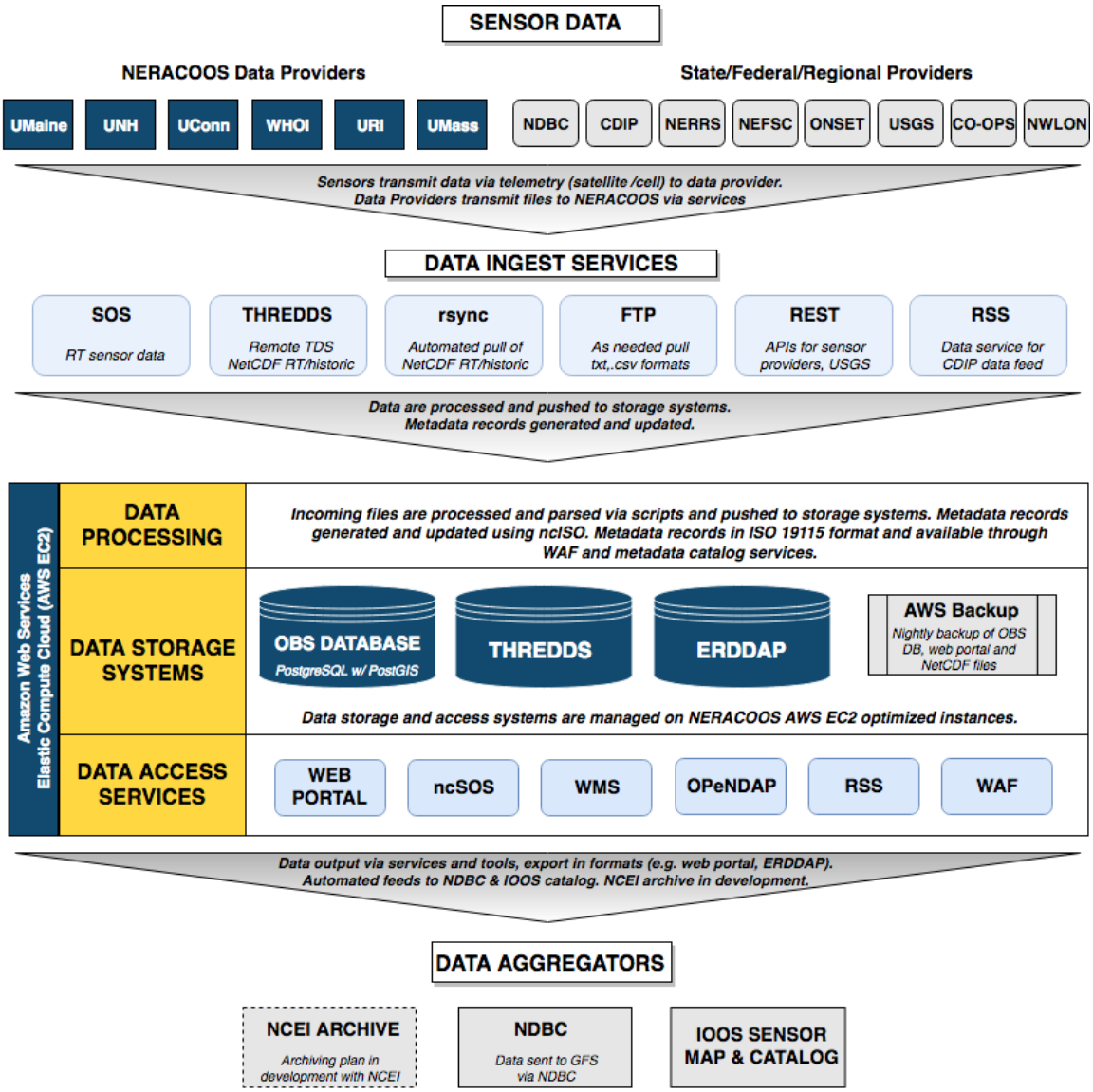


Figure 3: NERACOOS Data Management System Schematic

5.3.1 THREDDS (TDS) Thematic Real-Time Environmental Distributed Data Services

A THREDDS data Server (TDS) is installed on the NERACOOS cloud environment. Data are integrated into TDS via direct TDS to TDS communication with data partners hosting their own remote TDS (UMass Dartmouth and UConn) or by accessing netCDF files via remote web-accessible directories (UMaine and UNH) or as local files on stored on NERACOOS servers. The TDS plugin architecture utilizes a suite of standards that handle in-situ observations and models with similar protocols.

TDS is utilized to serve the NERACOOS provider data and metadata in IOOS DMAC compliant formats (ISO 19115,WMS, SOS, OPeNDAP, or subsets of netCDF). This distributed approach gives data providers control over the data and metadata and avoids replication. Post-recovery files that have been QA/QCd can then replace provisional real-time as they become available. The configuration of TDS/ERDDAP

enables it to recalibrate for new data automatically. A new SOS service is created for single parameter requests or time-series requests for data products. TDS can make collections of single files (i.e. all deployments of one buoy that may be in individual files). Various TDS plugins have been installed to accomplish this (ncML, ncISO, ncWMS, ncSOS, and OPeNDAP)

5.3.2 OPeNDAP - Open source Project for a Network Data Access Protocol

While OPeNDAP protocols are implemented within the TDS, NERACOOS continues to maintain a legacy OPeNDAP server and is running Hyrax version 1.8.

5.3.3 ERDDAP - Environmental Research Division's Data Access Program (NOAA NMFS SWFSC)

An ERDDAP server, a human readable alternative to TDS, is used as front end for speedier data access and provides output in a variety of usable formats. ERDDAP uses netCDF as a data source, but other formats such as .csv have been processed successfully for small, discrete historic data files. ERDDAP serves as a front-end tool for advanced users, but is also an effective backend providing data as a service for new products and tools. ERDDAP produces output files in common outputs: .html table, ESRI .asc and .csv, Google Earth .kml, OPeNDAP binary, .mat, .nc, ODV .txt, .csv, .tsv, .json, and .xhtml.

5.4 Hardware and Software Configuration

The NERACOOS Data Management System currently resides on three discrete server instances within a private Amazon Elastic Compute Cloud (Amazon EC2). The scalable instances were selected and optimized for the particular load required. These include instances dedicated to database, webserver and java processes. All three of the server instances are backed by Solid State Storage Devices (SSD) and are running a Linux Ubuntu OS.

The database server instance is optimized for compute power and is used to store and manage data for ingest and delivery of data streams that drive web-delivered products on the NERACOOS website, as well as the database that drives the Drupal Content Management System. The server is configured to run PostgreSQL with PostGIS and MySQL database capability.

The webserver server instance is optimized for a general-purpose mix of compute and memory allocation. This instance serves the code that runs the Drupal Content Management System and the individual data-driven applications and products that make up the NERACOOS.org website. The server is configured for PHP and JavaScript frameworks.

The Java instance is home to the data storage components of the DMS and is optimized for the memory load required to run the software for TDS, ERDDAP, Hyrax OPeNDAP and GeoServer software systems. The server is configured to run Tomcat and Java based software packages.

By standing up discrete instances optimized for the tasks, the risk of a resource-intensive process impacting the rest of the system is greatly reduced. This system is also very scalable enabling addition of new instances or adding capacity to existing instances.

The current configuration is outlined in the table below:

Table 1: NERACOOS hardware and software configuration

Database Instance

AWS Instance type/OS	Detail	Application environment/ software stack	Features
C3.xlarge Ubuntu 14.04 LTS	4 vCPU 7.5 GB memory, 2 x 40 SSD storage (GB)	PostgreSQL with PostGIS and MySQL	Compute optimized: High Frequency Intel Xeon E5-2680 v2 (Ivy Bridge) Processors, enhanced networking, SSD-backed instance storage

Webserver Instance

AWS Instance type/OS	Detail	Application environment/ software stack	Features
C3.large Ubuntu 14.04 LTS	2 cpu, 3.75 GB memory, 2 x 16 SSD storage (GB)	PHP, JavaScript	Compute optimized: High Frequency Intel Xeon E5-2680 v2 (Ivy Bridge) Processors, enhanced networking, SSD-backed instance storage

Java Instance

AWS Instance type/OS	Detail	Application environment/ software stack	Features
M3.large Ubuntu 14.04 LTS	2 vCPU, 7.5 GB memory, 1 x 32 SSD storage (GB)	Tomcat/Java for ODPDX, THREDDS, ERDDAP, GeoServer	General purpose/optimized: High Frequency Intel Xeon E5-2670 v2 (Ivy Bridge) Processors, SSD-backed instance storage

6 NERACOOS Data Streams

Overview

NERACOOS funds a significant amount of data collection in the Northeast region. While we strive for consistency and efficiency, each funded data provider has slightly different internal procedures and systems for data acquisition, processing, QA/QC and managing data. In this section we provide detail on each NERACOOS funded data provider who currently contributes data to the system. The table below (Table 2) provides and overview of the funded data provider, supported assets, and current and future data flow protocols. More detail for each data provider is described in detail in this section.

NERACOOS is working with all funded data providers to produce CF 1.6 netCDF, which will enable the IOOS compliant ncSOS services directly from the NERACOOS TDS. For

some providers, particularly UMaine, this has been a fairly complex upgrade due to the differences in discrete sampling geometry conventions for multiple depth and time variables at one station.

Table 2: NERACOOS data provider's data flow overview.

Data Provider	Sensors / Assets	Present Data Flow	Planned Data Flow
University of Maine	7 Buoys in Gulf of Maine	UMaine produces data in netCDF 1.4 and transmits to FTP site, NERACOOS automatically retrieves via rsync and inserts into OBS DB, converts to 1.6 and updates THREDDS catalog and adds data to ERDDAP.	UMaine will produce netCDF 1.6 for RT and historic data, files will be transmitted to FTP site. NERACOOS will retrieve via rync and add data directly into THREDDS and ERDDAP. UMaine is considering installing THREDDS server which NERACOOS will point to for catalog updates and adding data to ERDDAP.
	3 HF Radar Stations	UMaine sends data to HF RADAR DAC, NERACOOS acquires data via DAC feed and application	No change
	Satellite	UMaine acquires satellite data from NOAA and NASA data centers, processes data and provides satellite imagery to NERACOOS.	No change
University of Connecticut (UConn)	3 Buoys in LIS (Western LIS, Execution Rocks, and Central LIS)	UConn sends data to NDBC, NERACOOS acquires data from NDBC and inserts into OBS DB	UConn will produce netCDF 1.6 files and make available via on-site THREDDS. NERACOOS will acquires data from remote THREDDS, adds to ERDDAP.
University of New Hampshire (UNH)	3 Buoys/ 1 Shore station (Great Bay Buoy, CML station, CO2 buoy)	Legacy SOS installed on local data system, NERACOOS acquires RT data from all three stations and inserts into OBS DB.	Great Bay Buoy, CML, CO2: UNH will produces netCDF 1.6 and make available via on-site THREDDS. NERACOOS will acquire data from THREDDS and add to ERDDAP. UNH has recently installed on-site THREDDS server, NERACOOS has been testing data and access protocols to make remote THREDDS content available through catalog and ERDDAP server.

	1 CDIP	CDIP buoy telemeters data to CDIP Data Center; NERACOOS acquires RT data from RSS feed produced by UCSD.	No change
University of Rhode Island (URI)	Stations of the Narragansett Bay Fixed Site Monitoring Network	Legacy SOS installed on local data system at URI, NERACOOS acquires RT data from three stations and inserts into OBS DB	URI has already installed on-site THREDDS server and netCDF files from these buoys are available. NERACOOS has pointed THREDDS to URI remote THREDDS to populate catalog and will be making data available via ERDDAP.
Woods Hole Oceanographic Institution	1 HF Radar station	WHOI sends data to HF RADAR DAC, NERACOOS acquires data via DAC feed and application	No change
Charybdis Group	3 Tide Gauges (Scituate, Gloucester, Hampton)	Data are telemetered to sensor provider data management system. NERACOOS acquires data on demand as .csv.	Data provider will be moving to SOAP based SOS service. NERACOOS will acquire data and make available through THREDDS and ERDDAP.
USGS/Woods Hole Group	1 CDIP Buoy Cape Cod Bay	CDIP buoy telemeters data to CDIP Data Center; NERACOOS acquires RT data from RSS feed produced by UCSD.	No change

6.1 University of Maine Moored Buoy Array

Summary

The University of Maine (UMaine) operates seven oceanographic data buoys for NERACOOS: A01, B01, E01, F01, I01, M01, N01. The standard measurements on these buoys include near-surface winds, air temperature, barometric pressure, visibility, non-directional waves, surface ocean temperature and salinity, ocean temperature and salinity at depth and ocean currents throughout the water column. Additional detail on the parameters measured and specific sensors deployed is available in the NERACOOS Funded Assets table (http://www.neracoos.org/sites/neracoos.org/files/documents/NERACOOS Funded Assets_0.pdf). The data from all instruments are connected to a Campbell Scientific CR1000 data logger in each buoy for collection and transmission via cellphone IP data transfers. Backup transmissions are also received via NOAA's GOES satellite systems. All real-time data are processed and quality-controlled by UMaine and provided to NERACOOS. Data are stored in netCDF files; flat ascii files are generated from the

netCDF files. These buoys are at least serviced annually and some buoys are serviced twice a year.

Data Flow

The data processing system for data generated by the NERACOOS oceanographic buoys operated by UMaine consists of both field components and shore side components. A programmable data logger inside the buoy well timestamps data received from sensors, aggregates it, and transmits the data to UMaine via two methods, a) cellphone or satellite modem, or b) NOAA GOES satellite. This results in two streams of largely redundant data, designated "realtime" and "goes-realtime" respectively.

The two streams of buoy telemetry received at UMaine are processed by a near-real time processing system that has been in operation since approximately 2001, and extensively modified since then. The system comprises a UNIX-based application server and a UNIX-based web server, and is driven by a number of unix shell scripts, python scripts, and MATLAB scripts. Data is stored and archived in netCDF format, following CF and COARDS conventions in effect when the system was built. Additional metadata for netCDF files and directory structure is accessed via MySQL databases.

NERACOOS receives the provisional "realtime" buoy data from UMaine through an automated process that aggregates data and generates netCDF files that are pushed to an FTP site. On the NERACOOS side, a script runs every 10 minutes that checks for new data packages on the FTP server and acquires files via rsync. NetCDF files go through two processes on the NERACOOS side: legacy and new. The legacy processes uses scripts to unpack and extract data and metadata and insert into the Observations Database (OBS DB), which supports a set of operational data products. The newer processes extract metadata from the netCDF files, converts names to CF 1.6 accepted standard, and produces new netCDF CF 1.6 files, which are added to the NERACOOS THREDDS and ERDDAP catalogs. The netCDF files are aggregated by sensor packages and deployment. The reason for redundancy is to maintain functionality of legacy/operational data products as we transition the data system to the netCDF/THREDDS/ERDDAP model. Once all data products are getting data from the ERDDAP services, the legacy OBS DB and associated processes will not be needed. UMaine is also working to produce CF 1.6 files, which will eliminate the conversion step by GMRI. The netCDF files UMaine is developing will also be NCEI 2.0 compliant, which is required format for archiving. We anticipate that the system for generating CF 1.6 files will be completed in December 2017.

UMaine also generates historic data files after a buoy is recovered and data are downloaded from the data loggers. Data recovered from data loggers are reconciled with the real-time data files at UMaine and new files are generated that represent a more complete record of that deployment with a higher level of QA/QC. These netCDF files are pushed out to the FTP site into a subdirectory for historic data. The NERACOOS rsync process checks these directories as well, and any new packages are acquired and

added to the THREDDS/ERDDAP catalog using the same format described above. In the catalog, they are identified as historic to differentiate from the real-time provisional files.

Quality Control

All data variables reported to NERACOOS in real time netCDF files have an ancillary QC flag. A non-zero flag represents an invalid data value for that record. Malformed data and invalid data structures from instrumentation (similar to "Syntax Test" in QARTOD manuals) are likely to be rejected at the data logger onboard the buoy by embedded software. Shore side, initial quality control checks tailored to the separate data streams (similar to "Timing/Gap" and "Syntax" checks in QARTOD manuals) are performed on all of the cellphone and GOES satellite telemetry. After this point in the processing, the same processing routines are used for both streams.

A range check ("Gross Range Test" per QARTOD manual) is carried out for all reported data variables processed through the system. The "valid_range" netcdf variable attribute is used for this for this purpose at an early stage in processing. QC flags for data values falling outside the "valid_range" attribute are set to the value "1". Valid ranges for data variables are set by template prior to buoy testing and deployment. The ranges are determined by buoy personnel based on instrument limitations, empirical data, and knowledge of site environments -- most are tighter than the maximum "sensor span" for a given data variable.

A "Location Check" is also performed for each reported data variable in the system – as GPS positions are updated, the distance from the updated position to the nominal deployment position is calculated and if this is larger than a watch circle radius, alarms immediately notify data management personnel. Data from a confirmed position outside the watch circle are not reported and a QC flag is set to a non-zero value.

In addition to the real-time automated QC checks, data is reviewed by data management personnel on a daily basis using a variety of diagnostic tools including comparison with neighboring platforms (similar to QARTOD "Neighbor Tests"). If data from a sensor is found to be erratic or suspect in any way, the QC flags for associated variables are set to the non-zero value of "2" until the data can be assessed manually for validity. Additional automated and operator-supervised QC is carried out on when sensors with onboard storage are recovered post-deployment.

Collaborating scientists provide chlorophyll sensors deployed on UMaine buoys. The collaborating scientists maintain the sensors according to manufacturer specifications and have sensors serviced by the manufacturer. The scientist will deliver and mount the sensors on the buoy and provide UMaine with calibration coefficients prior to deployment. UMaine runs basic automated QC (range and location checks) on the real-time data streams. The collaborating scientists also monitor the data and inform UMaine if they detect any issues.

Automated QC checks for the in-situ current measurements (RDI ADCPs) are limited and are targeted for improvement. Part of this limitation is due to restricted bandwidth for telemetry that does not allow for sending back extensive diagnostic data in real-time (i.e. echo intensities, correlation magnitudes, and vertical and error velocities) for acoustic samples.

Automated QC checks for in-situ surface wave measurements (Summit accelerometer, UMaine Directional Wave Module) are also targeted for improvement.

Status of required QC tests for the data variables listed in the QARTOD manuals are listed in the following table. Data that fail a test are flagged with a non-zero QC flag.

In-situ temperature and Salinity		**Temperature**	**Salinity**	**Conductivity**
Test 1	Timing/Gap test	*Yes*	*Yes*	*Yes*
Test 2	Syntax Test	*Yes*	*Yes*	*Yes*
Test 3	Location Test	*Yes*	*Yes*	*Yes*
Test 4	Gross Range	*Yes*	*Yes*	*Yes*
Test 5	Climatology Test	*No*	*No*	*No*
Test 7	Rate of change test	*Yes*	*Yes*	*No*

Wind Data		**Wind Speed**	**Wind Direction**	**Wind Gust**
Test 1	Timing/Gap test	*Yes*	*Yes*	*Yes*
Test 2	Syntax Test	*Yes*	*Yes*	*Yes*
Test 3	Location Test	*Yes*	*Yes*	*Yes*
Test 4	Gross Range	*Yes*	*Yes*	*Yes*
Test 5	Climatology Test	*No*	*No*	*No*

In-situ Surface Wave Data		**Significant Wave Height**	**Dominant Wave Period**	**Avg Wave Direction**
Test 16	Stuck Sensor	*No*	*No*	*No*
Test 17	Operational Frequency Range	*No*	*Yes*	*No*
Test 18	LF Energy	*No*	*No*	*No*
Test 19	Bulk Wave Params	*Yes*	*Yes*	*Yes*
Test 20	Rate of Change	*No*	*No*	*No*

In-situ Current Observations		**Current Speed**	**Current Direction**	**Current U**	**Current V**
Test 2	Check Sum	*No*	*No*	*No*	*No*
Test 3	Sensor Tilt	*No*	*No*	*No*	*No*
Test 4	Speed of Sound	*No*	*No*	*No*	*No*
Test 6	Signal Strength	*No*	*No*	*No*	*No*

Test 10	Current Speed	*Yes*	*Yes*	*Yes*	*Yes*
Test 11	Current Direction	*Yes*	*Yes*	*Yes*	*Yes*
Test 12	Horizontal Velocity	*No*	*No*	*No*	*No*
Test 15	Stuck Sensor	*No*	*No*	*No*	*No*
Test 16	Echo Intensity	*No*	*No*	*No*	*No*

Backup and Archival

The UMaine data management system includes several data backup procedures. As buoy telemetry is received and processed in real-time, netCDF files specific to each buoy deployment are updated in a directory tree on the UMaine application server. Immediately after update of buoy data for a particular buoy, that data is replicated (via rsync) to the web server for distribution.

That data directory on that web server is used by NERACOOS (and CariCOOS) to ingest (via rsync) buoy data.

Once daily the entire directory tree on the application server (which contains historical and post-recovery as well as real-time netCDF data files) is replicated (again, in case any post real-time edits or updates were missed) to the web server for distribution.

Both the application server and the web server are backed up daily by incremental backup and weekly by full backup to rotating hard drives on a backup server located in a different on-campus building at UMaine.

Data archival with NCEI is described in Section 8.3.

6.2 University of Maine High Frequency Radar

Summary

UMaine operates three High Frequency Radar (HFR) stations for NERACOOS. These stations are located on Grand Manaan Island, New Brunswick, Cape St. Mary, Nova Scotia, and Green Island, Maine.

Data Flow

SeaSonde SSRS-100 hardware and software from CODAR Ocean Sensors run autonomously at the three sites. The remote sites acquire HFR data, archive the raw data, generate radial files, and push the radial files to a processing server at UMaine. Raw data is collected and brought back to UMaine during visits to the remove sites. The processing server, an SSDP-100 Central Site Management/Data Combining Station from CODAR Ocean Sensors, applies QA/QC and archives the radials files received from remote stations. The HFR radial files are ingested from the UMaine processing server by HRFNET, the national distributed processing system, where additional QA/QC is carried out and spatially averaged surface current data ("totals files") are generated. HFRNET makes radial and total files available through THREDDS and provides data

visualization products for the aggregated HFR data. The UMaine processing server also generates a set of spatially averaged surface current data ("totals files") for local review.

NERACOOS displays regional HFR data by using a mapping product developed by HFRNet. No HFR data are ingested or stored in the NERACOOS data system.

Quality Control
UMaine relies on manufacturer recommended software settings in the SeaSonde software acquisition software running on remote HFR sites, for acquisition of HFR data and unattended generation and quality control of the radials files prior to ingest by HFRNET. Further filtering and QA/QC of radials files, as well as generation and QA/QC of spatially averaged surface current data is carried out by HFRNET.

Backup and Archival
HFR radials files from UMaine are archived by the HFRNet at UCSD, along with totals files generated by HRFNET from the radial files. UMaine maintains a single set of the RAW HFR data collected from site visits on individual hard drives at the Orono campus. UMaine also maintains an archive of radial files received from HFR stations on the processing server at the Orono campus.

6.3 University of Maine Satellite Imagery

Summary
The Satellite Oceanography Data Laboratory (SODL) at UMaine provides the NERACOOS Satellite data products. UMaine collects, processes, QA/QCs, and archives the satellite data and delivers high-resolution data products to NERACOOS. Coverage includes the Gulf of Maine and Long Island Sound. The primary sources of satellite data for the NERACOOS stream are: NOAA AVHRR and MODIS (Aqua and Terra) which SODL uses to provide fully processed sea surface temperature (AVHRR) and ocean color and sea surface temperature (MODIS) products multiple times each day at 1 km resolution.

Data Flow
UMaine SODL acquires and processes multiple satellite data streams to produce NERACOOS satellite data products. Two sources of satellite data are collected by the UMaine system for the NERACOOS data products:

1. MODIS (Moderate Resolution Imaging Spectroradiometer) is an instrument aboard NASA's EOS Aqua and Terra satellites. MODIS Aqua and Terra views the entire Earth's surface every 1 to 2 days, acquiring data in 36 spectral bands (channels) that are used to calculate ocean surface temperature and biogeochemical products (SODL produces and serves a surface chlorophyll a product).
2. AVHRR (Advanced Very High Resolution Radiometer) aboard the NOAA POES satellite system is a 5-channel radiometer providing measurements used to calculate sea surface temperature. Currently, NOAA maintains two such

operational satellites, providing at a minimum 4 overpasses per day (depending on orbital geometry). Each satellite covers the Earth surface every 1-2 days.

Each orbit of NOAA AVHRR SST data and NASA MODIS color/SST data that include coverage of the Gulf of Maine are received, archived, and processed into geophysical products (4-6/day AVHRR, 1-2/day MODIS) at 1km resolution.

The SODL is linked via wide bandwidth Internet connection to the main US east coast NOAA satellite data server facility to receive the raw, unprocessed, NOAA HRPT telemetry stream that contains the AVHRR data in near real time (http://www.class.ngdc.noaa.gov/saa/products/welcome). The HRPT data acquired by the SODL originate from the Wallops Island VA reception facility.
For the MODIS data stream, the SODL server system interrogates the NASA Goddard Space Flight Center MODIS satellite reception and archive system where SODL has a subscription (https://earthdata.nasa.gov/earth-observation-data/near-real-time/download-nrt-data/modis-nrt).

Data Processing and QA/QC
MODIS: MODIS raw swath data are processed to fully cloud-masked and atmospherically corrected bio-geophysical ocean products using SODL-developed scripts that call subroutines from NASA's SEADAS satellite data processing software (https://seadas.gsfc.nasa.gov). Ancillary data files and updated coefficients required for the processing are acquired in real time from additional NASA data servers within the SEADAS processing system. Both the chlorophyll a and the SST MODIS products produced for NERACOOS utilize the standard global NASA ocean product algorithms implemented within SEADAS and are updated as NASA updates their products.

AVHRR
AVHRR raw data are processed to SST products using SeaSpace's proprietary TERASCAN software (https://www.seaspace.com). For this processing, each scene is manually viewed, checked and adjusted by a trained SODL operator to ensure the image geometry is correct. Two products are created, 1) a "Standard cloud-masked" product that uses information from within the scene itself and coefficients developed by SODL and implemented within the SeaSpace TERASCAN software, and 2) an "Enhanced cloud masked" product that uses the Standard Cloud mask product plus additional information from SST data over the previous several days implemented within SODL-written scripts modified from protocols published by Cayula and Cornillon, 1995 (Journal of Atmospheric and Oceanic Technology, 12: 821-829).

Backup and Archival
AVHRR SST final image products are archived in TERASCAN data format (TDF, a proprietary format used by the SeaSPACE TERASCAN processing software). These files, however, are accessible / readable as simple binary files. Additionally, JPEG copies of

each processed SST image are produced, posted to the UMaine SODL web site, and archived.

MODIS products are archived as netCDF files.

Both sets of products are archived in a hierarchical directory structure at the UMaine Advanced Computing Group (https://acg.umaine.edu), a professionally monitored, maintained, high performance computational facility with multiple back-up.

6.4 University of Connecticut Moored Buoy Array

Summary

University of Connecticut (UConn) operates three oceanographic data buoys for NERACOOS: Execution Rocks (EXRX), Western Long Island Sound (WLIS), and Central Long Island Sound (CLIS). These moorings measure wind speed and direction, salinity, conductivity, temperature, pressure, currents, and dissolved oxygen near the surface, bottom and mid-depth. The WLIS and CLIS buoys also support wave observations. Additional detail on the parameters measured and specific sensors deployed is available in the NERACOOS Funded Assets table (http://www.neracoos.org/sites/neracoos.org/files/documents/NERACOOS Funded Assets_0.pdf). All buoys are on a yearly service cycle with an approximate turnaround of 3-4 weeks. During this time the buoys are hauled and returned to the Avery Point facility. Mooring hardware is replaced and the buoy hull and tower is repaired and repainted. Instrument sensors are cleaned, calibrated, and repaired if needed, and any software upgrades applied.

Data Flow

The data from each sensor is collected by a Campbell Scientific CR1000 datalogger for storage and transmission via cellphone TCP/IP data transfers. All data received at UConn are automatically entered into a MS-SQL database and quality-controlled. UConn generates and sends data messages to NDBC as part of the real-time data processing stream. NERACOOS retrieves UConn data hourly from NDBC and incorporates into the NERACOOS OBS DB where they are available to NERACOOS data products.

UConn is in the process of developing and testing a system to convert the data to netCDF format and store it locally on a UConn based THREDDS server. The data will then be pulled by NERACOOS from the THREDDS server. Once this is complete, the UConn buoy data will be available through THREDDS and ERDDAP. We anticipate that this system will be complete by October 2017.

Quality Control

The key component of the data collection and dissemination process aboard the metocean buoys is the data logger. The data logger collects, and transmits the data, via cellular communications, back to UConn where it is stored and subsequently accessed

and redistributed by NERACOOS. Nearly all the data are sampled and collected every 15 minutes and published to the web in near real time.

All data are initially received as comma delimited ASCII data files that are then read by automated scripts and stored in a MS-SQL database as provisional data. First order QA/QC protocols are subsequently applied to the data from the provisional database. UConn is developing a system to convert data to netCDF format for distribution.

Quality control checks are applied to the data after insertion into the database and before conversion to netCDF format. The main function of the first check is to find invalid data ranges from the sensors. Invalid data ranges are removed and replaced with a NAN. The data are then scanned for duplicate timestamps and location verification. Duplicate data are removed from the database, data collected outside the defined station (watch) radius are saved but stored in a separate data table. Data are then manually reviewed and compared with any additional data collected from or near the station (i.e., CTD casts, other sampling surveys, etc.) and if determined to be suspect, is flagged.

Backup and Archival
The live data stream collected using the proprietary data logger software and stored as comma delimited ASCII files are backed up to a local external hard drive and onto UConn's Enterprise File Services network drive, located offsite on the Storrs campus, every hour. In addition to the live data stream, data is logged on each instrument and on the data logger. These data are collected when the buoy is hauled or the instruments are retrieved and subject to the same backup procedure. Data on the database server is backed up onto two separate external hard drives once daily.

Data archival with NCEI is described in Section 8.3.

6.5 University of New Hampshire Mooring and Sensor Network

Summary
The University of New Hampshire (UNH) operates several monitoring assets for NERACOOS including the Great Bay buoy, CO2 buoy, the Coastal Marine Lab system (CML) and a CDIP buoy. The Great Bay Buoy is a highly instrumented mooring in Great Bay, NH that is deployed during ice-free months. This buoy measures a standard suite of meteorological and oceanographic parameters including air temperature, winds, water temperature and salinity. The Great Bay buoy also monitors dissolved oxygen, chlorophyll, optical properties, dissolved nutrients and CO2. CML is located at the mouth of the Great Bay Estuary. This station monitors surface meteorology and water quality and CO2. The CO2 buoy is located off the coast of NH and measures surface meteorology, water temperature, salinity, and air and water CO2. The CDIP buoy is located offshore on Jeffrey's Ledge and monitors waves and water temperature. Additional detail on the specific sensors deployed is available in the NERACOOS Funded

Assets table (http://www.neracoos.org/sites/neracoos.org/files/documents/NERACOOS Funded Assets_0.pdf).

Data Flow

Great Bay Buoy

The data from all instruments on the Great Bay buoy are connected to an Axys Watchman500 data logger. Data are telemetered via radio transmission to a shore-based computer at the UNH Jackson Estuarine Laboratory. Real-time data are processed and currently served to NERACOOS as text files. UNH is in the process of implementing a system to convert Great Bay buoy data to netCDF format and post on the UNH THREDDS server.

CML

Data from the CML are continuously logged by a computer in the CML building, and transmitted hourly to a computer system on the UNH Durham campus, where data from each sensor are automatically processed hourly to align data in time and produce hourly means, medians, minimums, maximums, and standard deviations. Data also undergo automated quality control tests (QARTOD tests) when available. After quality control testing, data are transmitted to NERACOOS as text files, and are also provided via THREDDS server in netCDF format.

CO2 Buoy

The data from all instruments are connected to an Axys Watchman500 data logger with the exception of the xCO2 and pH readings, which log directly to the MAPCO2 system and are sent back to PMEL daily. The remaining data are telemetered via radio transmission to a shore-based computer at the Seacoast Science Center and transmitted hourly to a computer system on the UNH Durham campus. Data undergoes automated quality control tests (QARTOD tests) when available. After quality control testing, data are transmitted to NERACOOS as text files, and are also provided via THREDDS server in netCDF format.

NERACOOS currently acquires real-time data from all three of the UNH buoys via an SOS service. Data are acquired and inserted into the OBS DB where they are available to data products. UNH has set up a THREDDS server and is producing CF 1.6 netCDF files for the CML and CO2 buoy data streams. NERACOOS is in the process of testing this service and will be making the switch to acquiring data directly from THREDDS to insert into the OBS DB and ERDDAP. UNH is also transitioning the data processing procedures for the Great Bay buoy to generate netCDF 1.6 compliant real-time and historical QA/QC files that will be accessible directly from the UNH THREDDS server. The NERACOOS THREDDS server is configured to point to UNH's THREDDS server to populate the NERACOOS data catalog. Once this is complete, all of the UNH buoy data will be available through THREDDS and ERDDAP. We anticipate that this transition will be complete by January 2018.

The UNH maintained CDIP buoy data flow is described in Section 6.9.

Quality Control

All CML and CO2 Buoy data variables reported to NERACOOS as netCDF files, that have published QARTOD procedures, include ancillary QC flags corresponding to the appropriate QARTOD procedure. Real time text file data have values flagged with a value other than 1 (i.e. "Suspect" or "Fail") removed. A table of checks performed on appropriate variables is given below.

A "Location Check" is also performed for each reported data variable in the CO2 Buoy system – as GPS positions are updated, the distance from the updated position to the nominal deployment position is calculated and if this is larger than a watch circle radius, alarms immediately notify data management personnel. Data from a confirmed position outside the watch circle are not reported and a QC flag is set to a non-zero value.

The CO2 buoy CO2 sensor data is sent to Pacific Marine Environmental Lab (PMEL) to undergo QA/QC. The QA/QC procedure used at PMEL is described in http://www.earth-syst-sci-data.net/6/353/2014/essd-6-353-2014.html

All chlorophyll sensors operated by UNH are returned to wetlabs and recalibrated on the recommended schedule. UNH staff visits the CO2 buoy four times a year to collect discrete chlorophyll samples that are used to validate the chlorophyll sensor operation. The Great Bay NERRS system wide monitoring program collects extensive water quality data from Great Bay monthly and these results are used to validate the chlorophyll sensor on the Great Bay buoy.

In addition to the real-time automated QARTOD QC checks, data are reviewed by data management personnel on a regular basis using a variety of diagnostic tools including comparison with neighboring platforms. Additional automated and operator-supervised QC is carried out on when sensors with onboard storage are recovered post deployment. UNH runs QARTOD tests on the CO2 buoy chlorophyll data and are in the process of implementing the chlorophyll QARTOD tests for the Great Bay buoy.

Status of required QC tests for the data variables listed in the QARTOD manuals are listed in the following table. Data that fail a test are flagged with a non-zero QC flag.

CO2 Buoy

Temperature/salinity

		Temperature	Salinity
Test 1	Gap Test	*yes*	*yes*
Test 2	Syntax Test	*yes*	*yes*
Test 3	Gross Range Test	*yes*	*yes*
Test 4	Climatology Test	*yes*	*yes*

Oxygen		Oxygen Concentration
Test 1	Gap Test	*yes*
Test 2	Syntax Test	*yes*
Test 3	Gross Range Test	*yes*
Test 4	Climatology Test	*yes*

Turbidity		Turbidity
Test 1	Gap Test	*yes*
Test 2	Syntax Test	*yes*
Test 3	Gross Range Test	*yes*
Test 4	Climatology Test	*yes*

Atmospheric data		Wind Speed	Wind Direction	Wind Gust	Air Temperature
Test 1	Gap Test	*yes*	*yes*	*yes*	*yes*
Test 2	Syntax Test	*yes*	*yes*	*yes*	*yes*
Test 3	Gross Range Test	*yes*	*yes*	*yes*	*yes*
Test 4	Climatology Test	*yes*	*yes*	*yes*	*yes*

CML

Temperature/salinity

		Temperature	*Salinity*
Test 1	Gap Test	*yes*	*yes*
Test 2	Syntax Test	*yes*	*yes*
Test 3	Gross Range Test	*yes*	*yes*
Test 4	Climatology Test	*yes*	*yes*

Oxygen		*Oxygen Concentration*
Test 1	Gap Test	*yes*
Test 2	Syntax Test	*yes*
Test 3	Gross Range Test	*yes*
Test 4	Climatology Test	*yes*

Turbidity		*Turbidity*
Test 1	Gap Test	*yes*
Test 2	Syntax Test	*yes*
Test 3	Gross Range Test	*yes*
Test 4	Climatology Test	*yes*

Atmospheric data		*Wind Speed*	*Wind Direction*	*Wind Gust*	*Air Temperature*
Test 1	Gap Test	*yes*	*yes*	*yes*	*yes*
Test 2	Syntax Test	*yes*	*yes*	*yes*	*yes*
Test 3	Gross Range Test	*yes*	*yes*	*yes*	*yes*
Test 4	Climatology Test	*yes*	*yes*	*yes*	*yes*

Backup and Archival
All UNH data are archived on the local UNH computer, an external hard drive, and backed up to a cloud data storage service (Box@UNH).

Data archival with NCEI is described in Section 8.3.

6.6 University of Rhode Island

Summary
NERACOOS provides partial support to the University of Rhode Island (URI) to operate a dock and mooring based water quality sensors and implement data management processes for the dock and mooring stations of the Narragansett Bay Fixed Site Monitoring Network (NBFSMN). The partners of the NBFSMN fund the majority of the data collection and management effort. Data from this network are being integrated into the NERACOOS DMS. Additional detail on the specific sensors deployed is available in the NERACOOS Funded Assets table (http://www.neracoos.org/sites/neracoos.org/files/documents/NERACOOS Funded Assets_0.pdf).

Data Flow
At the present time, real-time data are acquired from two stations in the NBFSMN via an SOS service. Data are acquired and inserted into the OBS DB every hour where they are available to data products. As part of the transition the new data management system, URI has set up a THREDDS server where they are making netCDF files directly available for those two stations and several others. The NERACOOS TDS server is configured to point to URI's TDS and populate the catalog. The NBFSMN buoy data will be available through the NERACOOS ERDDAP by December 2017. The list of stations currently serving surface and bottom data via OPeNDAP and SOS are Mount View, Quonset Point, Cole River, Taunton River and the GSO-Pier.

Quality Control
The NBFSMN has an extensive Quality Assurance Project Plan (QAPP) that documents their QA procedures and was approved by the EPA. Please see NBFSMN QAPP (http://neracoos.org/sites/neracoos.org/files/documents/NBFSMN%20QAPP.pdf)

Backup and Archival
Data from the NBFSMN are backed up on the base station computer at Marine Ecosystems Research Laboratory building within URI-GSO. All NBFSMN data (raw, edited and corrected data sets) are stored and available to the public annually through the RIDEM website (http://www.dem.ri.gov/programs/emergencyresponse/bart/stations.php).

6.7 Charybdis Group Tide Gauges

Summary
NERACOOS support the operation of three tide gauges that are located in Scituate, MA, Gloucester, MA, and Hampton, NH. These gauges monitor coastal water level.The Charybdis Group out of Boston, Massachusetts, maintains the stations.

Data Flow
A programmable Microwave Radar transducer (MRT) is coupled with an integrated data logger/telemetry unit (DLT) to upload the data over the 3G cellular network to a cloud based hosted service and archive (HAS).

NERACOOS currently uses an ftp system to retrieve data in .csv format from a ftp server. The data are currently accessed on demand when requested by the data products. The data access system is moving to a SOAP based SOS service, which will greatly improve the process for acquiring and disseminating the data.

Quality Control
The MRT has an inherit precision and accuracy that is specified by the manufacturer. In accordance with the manufacturer the instrument is sent back to the factory for evaluation and calibration at the prescribed intervals. The manufacturer of the MRT is ultimately responsible for the accuracy of the instrument as long as the maintenance intervals are respected.

Waves and swells affect the measurement of MLLW elevation; such effect is removed by applying an appropriate sampling protocol (NOAA or CG30) as described above. This process also takes care of the occasional presence of flotsam.

The location and placement of the tide gauge is carefully studied to insure the least interference of local currents and other dynamic phenomena. The Platform Of Opportunity (POO) is selected to be a stable platform and its elevation is surveyed at the time of installation and yearly thereafter. If a chance in elevation of the platform is discovered and no specific point in time can be determined for when the shift occurred (such as an extreme episodic event), then the change in elevation is considered gradual and is linearly applied retroactively to the previous year's data; alternatively, the change is applied as a single vertical shift from the date of the identified episodic event, and the data are shifted accordingly from that point onward. Charybdis Group LLC is responsible for the elevation reference of the tide gauge.

For any transmission error from the DLT to the HAS we defer to the Cellular Network and host CRC (Circular Redundancy Checks). At the specified 6 minute acquisition rate, the DLT is capable of holding up to 2 years worth of data. In the event of a cellular network disservice the DLT will be able to locally store the data and later upload to the

HAS asynchronously. The data integrity from the DLT to the HAS is the responsibility of the DLT Manufacturer, since the HAS is a bundled service.

The MRT measures raw distance to water (ullage). As well, the DLT stores raw distance to water. The HAS stores raw distance to water and converted water level with respect to MLLW; the conversion formula is hosted by the HAS.

Backup and Archival
All data is maintained on the HAS on raid 6 and higher disk arrays and periodically archived in accordance to best industry standards.

6.8 Woods Hole Oceanographic Institution High Frequency Radar

Summary
The Woods Hole Oceanographic Institution (WHOI) operates one HFR station, located on Race Point, MA for NERACOOS.

Data Flow
The data from this station is transmitted to WHOI where it undergoes QA/QC processing. The output of the WHOI data processing method is a ascii text file that follows the existing and required data file formats and metadata requirements of NOAA-IOOS and is transferred to the HFRNet at UCSD for ingestion into the total velocity calculation that is managed and distributed by HFRNet.

Quality Control
The WHOI-provided HF radar based observations of surface currents are processed from the raw spectral estimates of radar backscatter into surface currents following similar methods to those used and documented by Codar Ocean Sensors for the SeaSonde HF radar with a few key differences. Following Kirincich et al. (2012)[1] and de Poalo et al. (2015)[2], the raw results of the direction finding step are subjected to advanced quality control thresholds to eliminate suspect data and processed to spatially averaged surface currents using power-weighted radial averages. These two steps, in addition to the long time basis for the spectral estimates themselves have been shown to lead to higher quality radial velocity results. The WHOI QA/QC process will also being implemented for the UMaine HFR processing system described above.

Backup and Archival
HFR data generated by WHOI are archived at the HFRNet at UCSD.

---

[1] Kirincich, A. R., T. de Paolo, and E. Terrill (2012), Improving HF radar estimates of surface currents using signal quality metrics, with application to the MVCO high-resolution radar system, J. Atmos. Ocean. Technol., 29(9), 1377–1390, doi: 10.1175/JTECH-D-11-00160.1.

[2] de Paolo, T. D., E. Terrill, and A. Kirincich, (2015): Improving SeaSonde Radial Velocity Accuracy and Variance using Radial Metrics. Trans. Ocean. MTS/IEEE Genova, 6.

6.9 CDIP Wave Buoys

Summary
There are three CDIP buoys deployed in the NERACOOS region: Jeffrey's Ledge in the Gulf of Maine (maintained by UNH), Cape Cod Bay (maintained by USGS/Woods Hole Group) and Block Island in Rhode Island (maintained by USACE). The buoys record wave height, wave period, wave direction and water temperature. NERACOOS partially supports the UNH maintained buoy and will fully support the Cape Cod Bay buoy in 2017. NERACOOS ingests data from the CDIP program for all three buoys.

Data Flow
The data for all three CDIP buoys are managed and disseminated by the central data processing and archiving facility at University of California San Diego. The buoy telemeters data via satellite to the CDIP data center. There are failover processes as well that use HF radio to shore station. Data are retrieved from the data logger after recover and transmitted to the CDIP Data Center. NERACOOS uses an RSS feed from UCSD to acquire the real-time data hourly and make it available through data products on the website.

Quality Control
QA/QC for the CDIP buoys are conducted by the CDIP program.

Backup and Archival
The CDIP program provides data backup and archival

7 Data Sharing and Access

Data served through the NERACOOS website (http://www.neracoos.org) complies with the NOAA Data Sharing Procedural Directive by providing open access to real-time and historic data through a variety of services including data products and machine-readable services. The data and metadata formats used by NERACOOS adhere to IOOS guidance based on open standards (CF 1.6, ncSOS, ISO 19115). The website provides access to a several data services that support machine-to-machine data access as well as a variety of data products that allow end users to view and/or download near real-time data, historic data, and model predictions in a variety of formats. The NERACOOS website also provides access to information about NERACOOS including but not limited to strategic framework, board of directors, members, by-laws, grants, news, projects and partners.

The NERACOOS website utilizes a Responsive Web Design approach that enables data products and information to be effectively viewed on mobile devices.

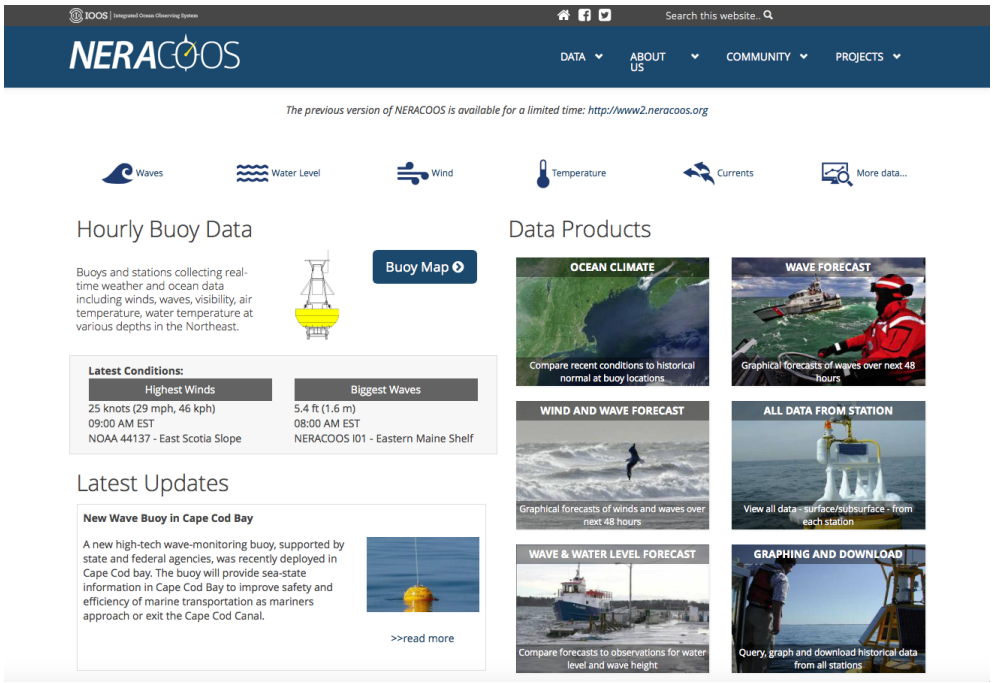


Figure 4: Screen capture of the NERACOOS home page.

7.1 Hourly Buoy Data

The NERACOOS Hourly Buoy Data product (http://www.neracoos.org/realtime_map) provides access to real-time ocean observations from regional observing assets including NERACOOS buoys and sensors, NDBC buoys, C-MAN stations, CDIP buoys, Environment Canada buoys, NWLON stations, USGS water level stations, NERRS stations and various short term research buoys in the region. This product provides a map of the region with clickable symbols that identify the location and type of ocean observing station. Selecting a station provides the user a view of the latest weather and sea surface observations from that station. The product also provides links to a series of related products described below.

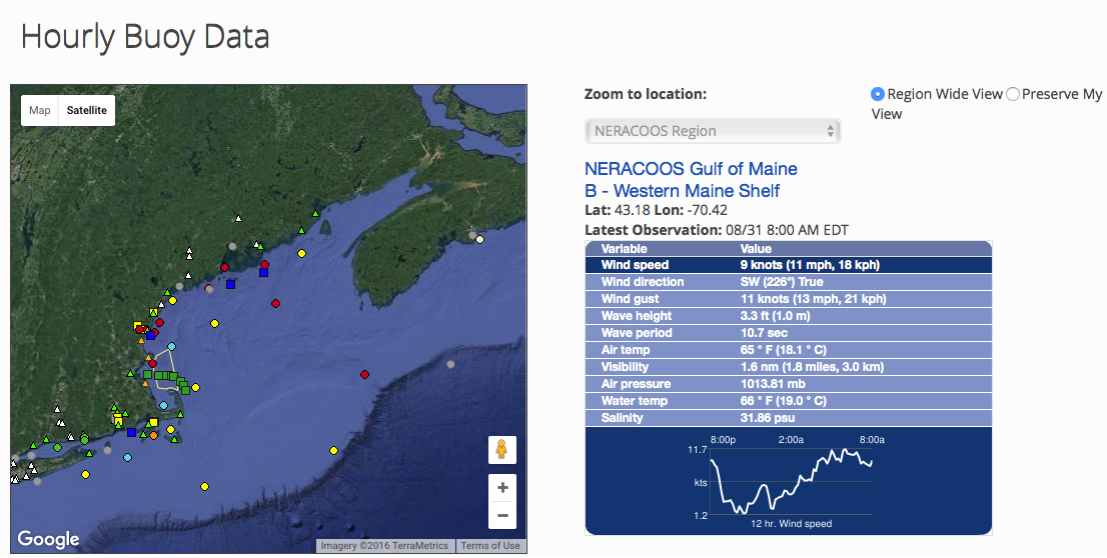
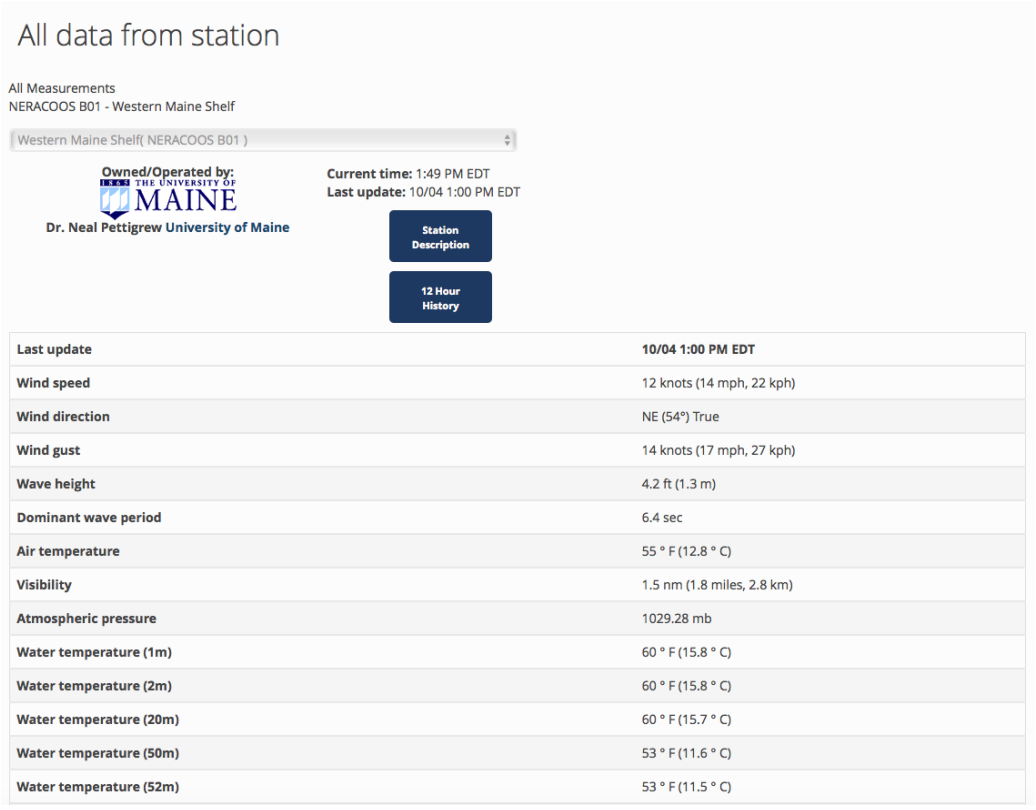


Figure 5: Screen capture of the NERACOOS hourly buoy data product.

7.2 All Data From Station

The All data from station display allows a user to see all the data collected by a buoy in the past hour. Specifically, a user can view hydrographic conditions throughout the water column that are not available from the initial view of the data.

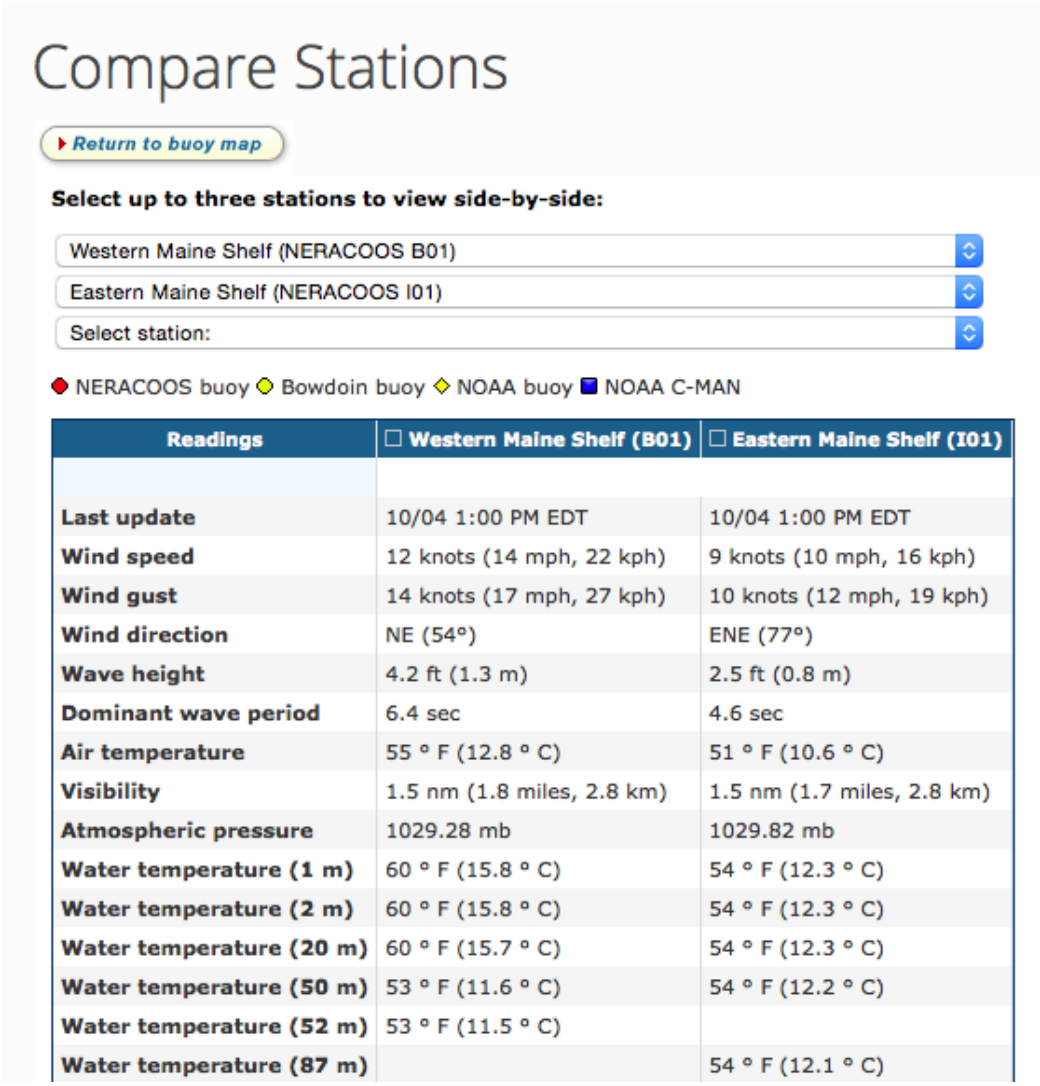


Last update	**10/04 1:00 PM EDT**
Wind speed	12 knots (14 mph, 22 kph)
Wind direction	NE (54°) True
Wind gust	14 knots (17 mph, 27 kph)
Wave height	4.2 ft (1.3 m)
Dominant wave period	6.4 sec
Air temperature	55 ° F (12.8 ° C)
Visibility	1.5 nm (1.8 miles, 2.8 km)
Atmospheric pressure	1029.28 mb
Water temperature (1m)	60 ° F (15.8 ° C)
Water temperature (2m)	60 ° F (15.8 ° C)
Water temperature (20m)	60 ° F (15.7 ° C)
Water temperature (50m)	53 ° F (11.6 ° C)
Water temperature (52m)	53 ° F (11.5 ° C)

Figure 6: Screen capture of the NERACOOS all data from stations product

7.3 Compare stations

The Compare stations product allows a user to select up to three buoys or stations and compare the latest observations side-by-side. This tool helps mariners see how conditions may be changing in the region or what conditions are like at a location they may be heading to.

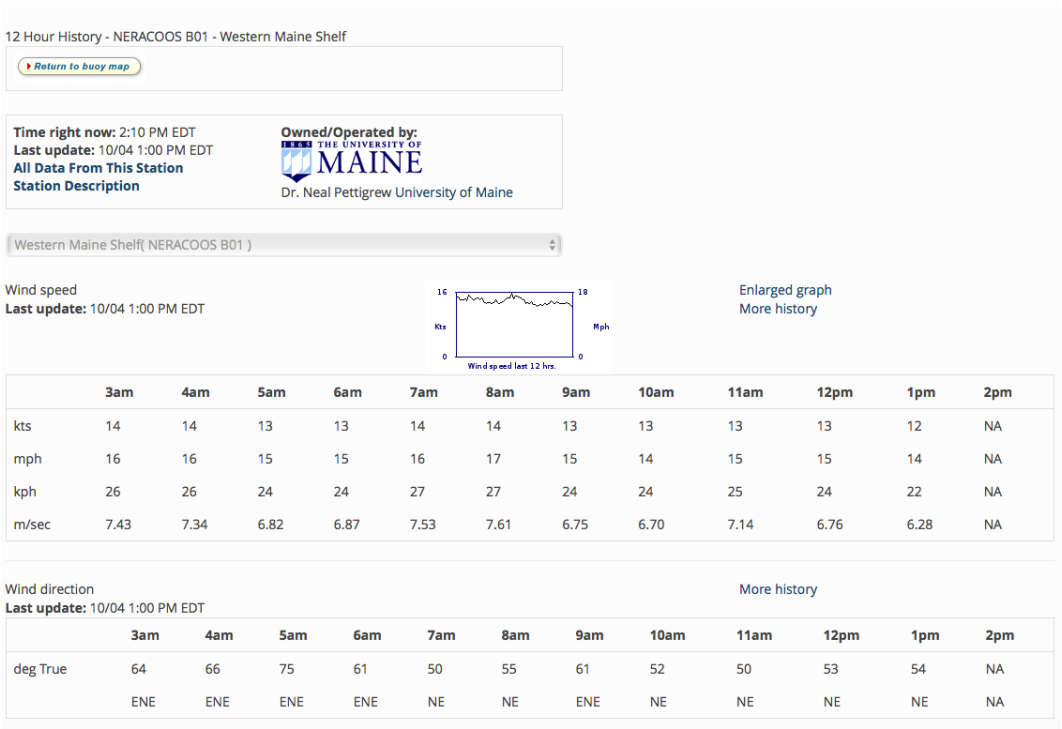


Readings	Western Maine Shelf (B01)	Eastern Maine Shelf (I01)
Last update	10/04 1:00 PM EDT	10/04 1:00 PM EDT
Wind speed	12 knots (14 mph, 22 kph)	9 knots (10 mph, 16 kph)
Wind gust	14 knots (17 mph, 27 kph)	10 knots (12 mph, 19 kph)
Wind direction	NE (54°)	ENE (77°)
Wave height	4.2 ft (1.3 m)	2.5 ft (0.8 m)
Dominant wave period	6.4 sec	4.6 sec
Air temperature	55 ° F (12.8 ° C)	51 ° F (10.6 ° C)
Visibility	1.5 nm (1.8 miles, 2.8 km)	1.5 nm (1.7 miles, 2.8 km)
Atmospheric pressure	1029.28 mb	1029.82 mb
Water temperature (1 m)	60 ° F (15.8 ° C)	54 ° F (12.3 ° C)
Water temperature (2 m)	60 ° F (15.8 ° C)	54 ° F (12.3 ° C)
Water temperature (20 m)	60 ° F (15.7 ° C)	54 ° F (12.3 ° C)
Water temperature (50 m)	53 ° F (11.6 ° C)	54 ° F (12.2 ° C)
Water temperature (52 m)	53 ° F (11.5 ° C)	
Water temperature (87 m)		54 ° F (12.1 ° C)

Figure 7: Screen capture of the NERACOOS compare stations product.

7.4 12 Hour History

The 12 Hour history product allows a user to view the past 12 hours of observations in both a table and graph format. Mariners specifically requested this display so that they could see how the ocean conditions were changing over the past 12 hours, which allows them to see how conditions are trending.



12 Hour History - NERACOOS B01 - Western Maine Shelf

Return to buoy map

Time right now: 2:10 PM EDT
Last update: 10/04 1:00 PM EDT
All Data From This Station
Station Description

Owned/Operated by:
THE UNIVERSITY OF MAINE
Dr. Neal Pettigrew University of Maine

Western Maine Shelf(NERACOOS B01)

Wind speed
Last update: 10/04 1:00 PM EDT

Enlarged graph
More history

	3am	4am	5am	6am	7am	8am	9am	10am	11am	12pm	1pm	2pm
kts	14	14	13	13	14	14	13	13	13	13	12	NA
mph	16	16	15	15	16	17	15	14	15	15	14	NA
kph	26	26	24	24	27	27	24	24	25	24	22	NA
m/sec	7.43	7.34	6.82	6.87	7.53	7.61	6.75	6.70	7.14	6.76	6.28	NA

Wind direction
Last update: 10/04 1:00 PM EDT

More history

	3am	4am	5am	6am	7am	8am	9am	10am	11am	12pm	1pm	2pm
deg True	64	66	75	61	50	55	61	52	50	53	54	NA
	ENE	ENE	ENE	ENE	NE	NE	ENE	NE	NE	NE	NE	NA

Figure 8: Screen capture of the NERACOOS 12 hour history product.

7.5 Ocean Climate

The Ocean Climate product allows a user to view the latest buoy observations in comparison to a 10 + year climatology for the same meteorological or oceanographic parameter. This allows the user to see how current conditions compare to the average conditions for that location. The tool allows a user to look from year to year and view how any given years data compares to the average conditions.

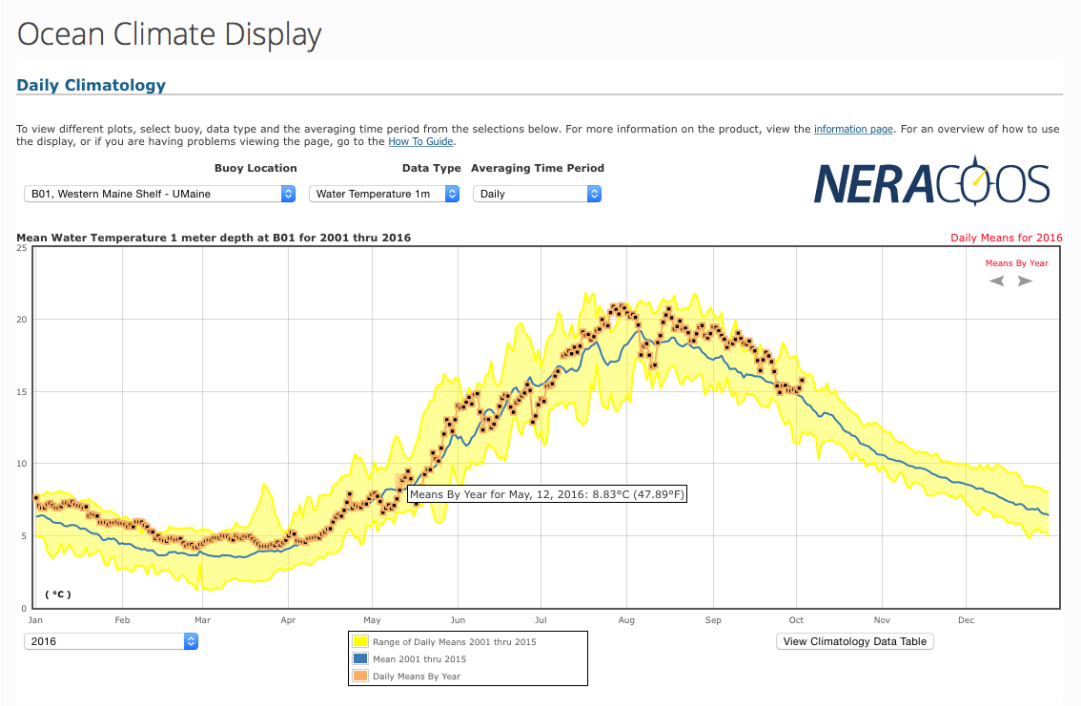


Figure 9: Screen capture of the NERACOOS ocean climate product.

7.6 Region Wide Buoy Conditions

The Region Wide Buoy Conditions product provides a regional display of the latest meteorological and oceanographic observations. A user can select up to two parameters from the buoys and the product will display all the latest observations on a map.

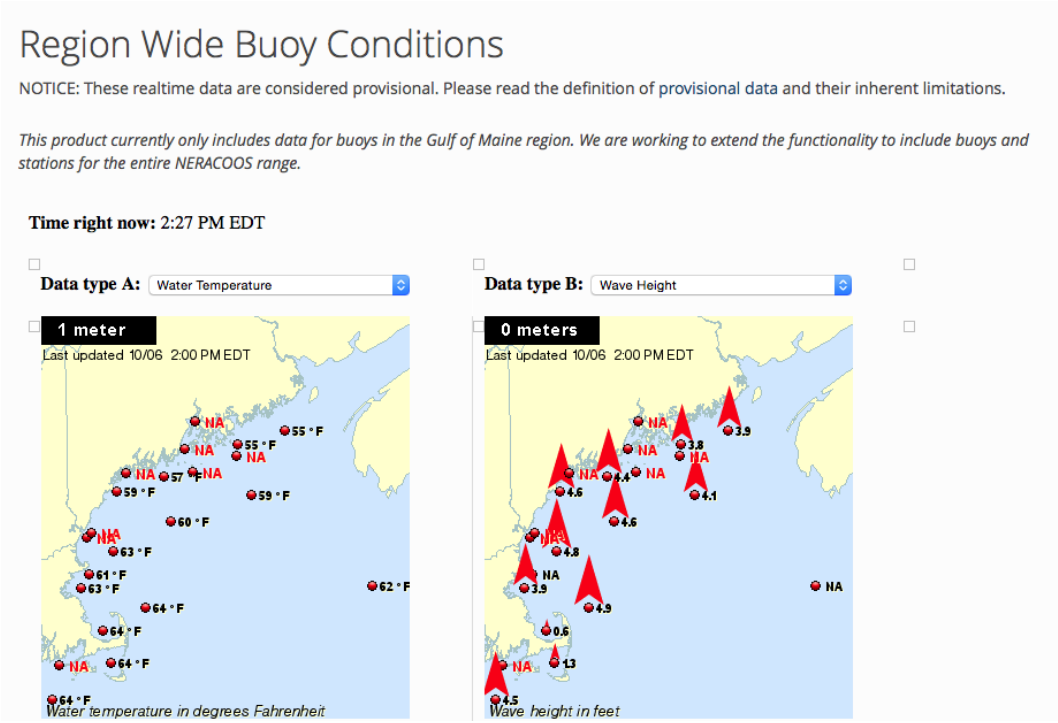


Figure 10: Screen capture of the NERACOOS region wide buoy conditions product.

7.7 Graphing and Download

The graphing and download tool provides a user-friendly interface to graph and/or download hourly meteorological and oceanographic data from buoys and stations in the NERACOOS region. Graphing and Download allows a user to select the parameter(s) they are interested in, select a time frame they want to see/access the data for, select the buoy location(s) they want data from and then they can view a graph of the data or access the data in several common file formats.

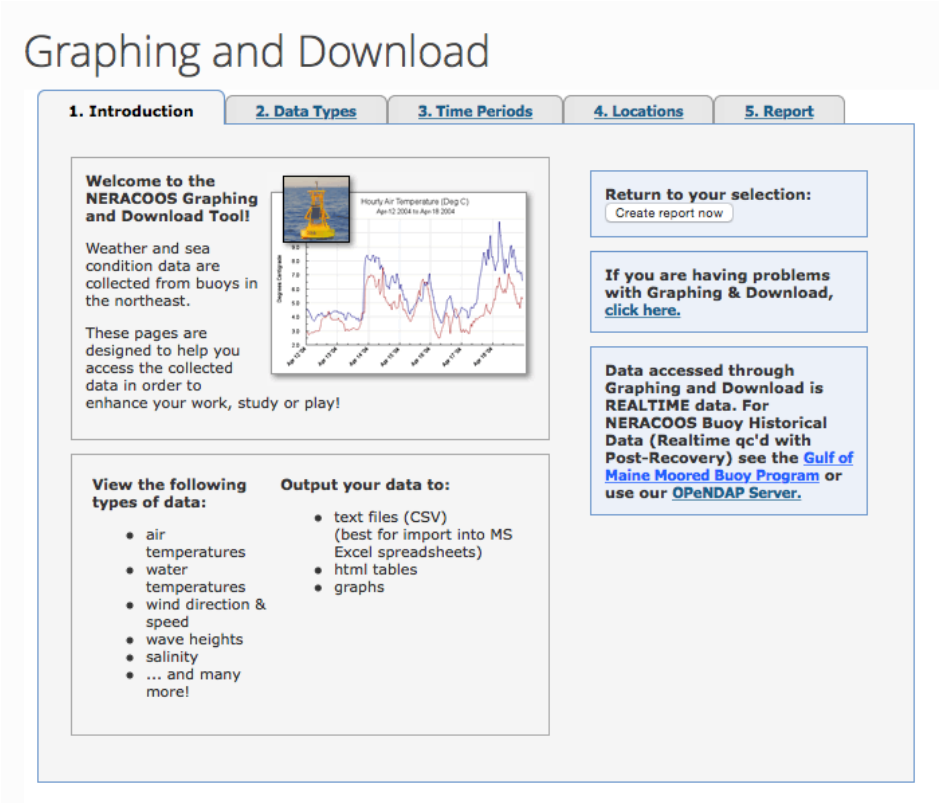


Figure 11: Screen capture of the NERACOOS graphing and download product.

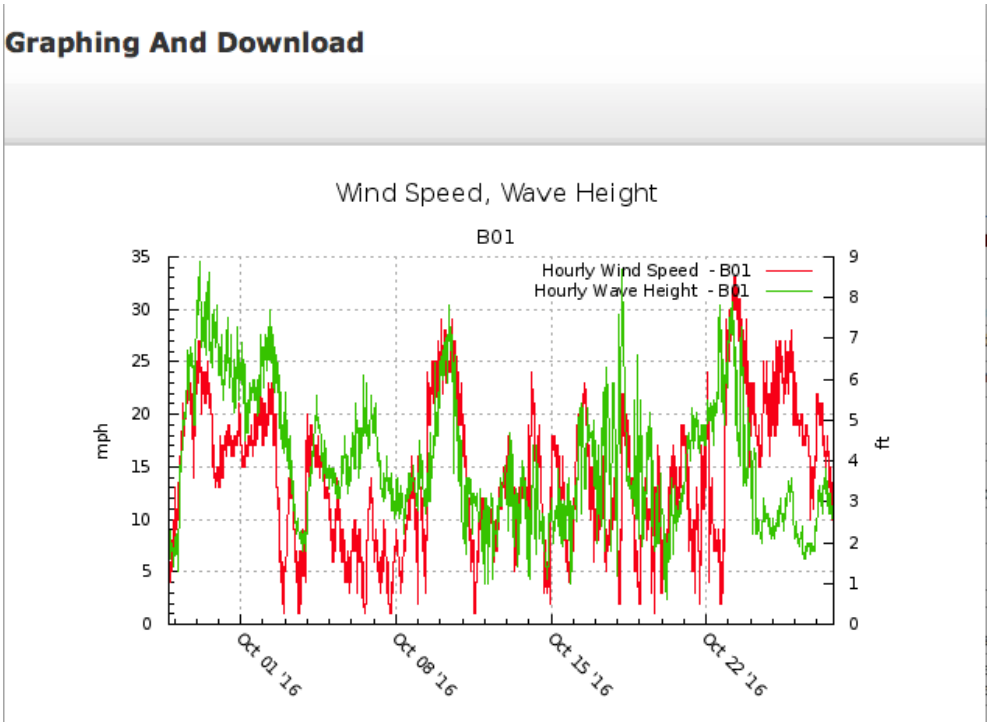


Figure 12: Screen capture of the NERACOOS graphing and download output.

7.8 Wave Forecast

The Wave Forecast product is a graphical display of predicted wave conditions for the next 48 hours in the region. The color on the map depicts the predicted significant wave height and the arrows depict the predicted direction and period of waves. The forecast maps are updated twice daily and are generated from the WaveWatch III model.

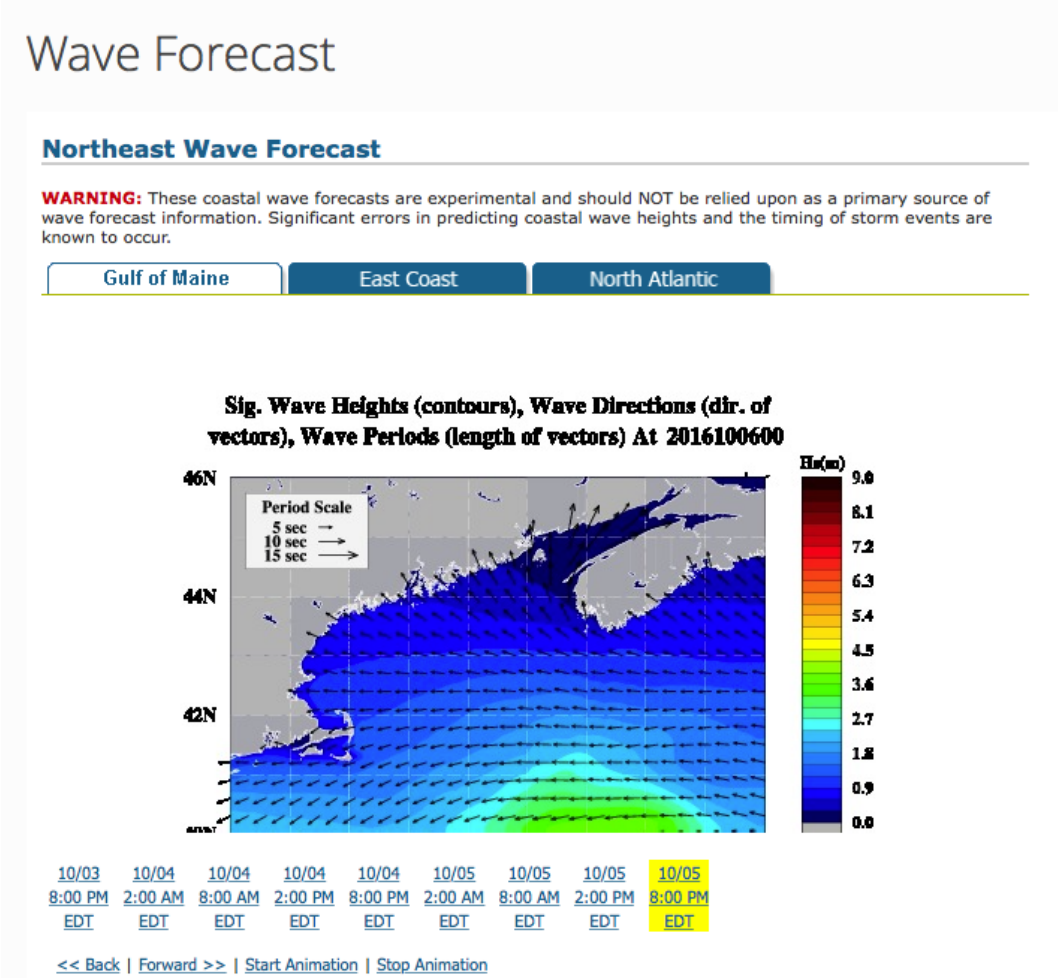


Figure 13: Screen capture of the NERACOOS wave forecast product.

7.9 Wind and Wave Forecast

The wind and wave forecast product provides an easy to use view of predicted sea state conditions. The product combines a graphical and numeric display of WaveWatch III forecasted wave height, period and direction along with NOAA NCEP forecasted wind speed and direction. This product, requested by mariners combines critical sea state predictions into a single display.

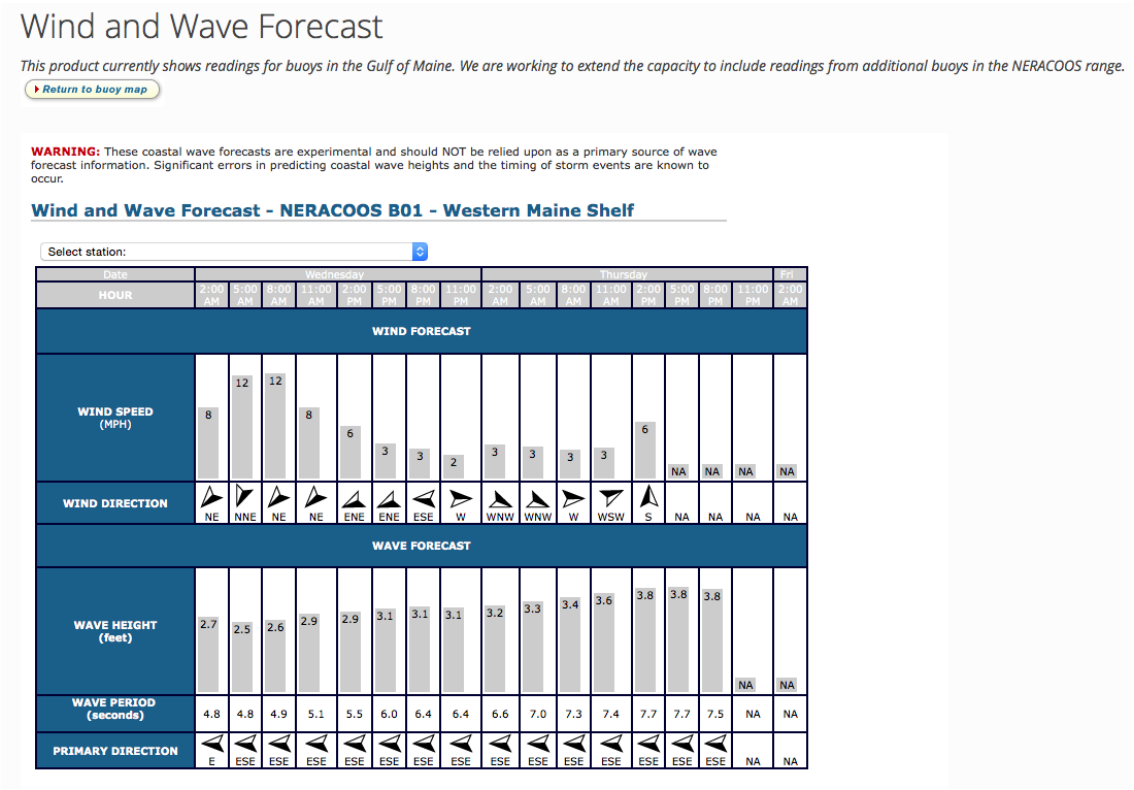


Figure 14: Screen capture of the NERACOOS wind and wave forecast product.

7.10 Water and Wave Level Observation Viewer

This product allows user to select a buoy or water level station and view a plot of observed data (wave height or water level) from that station along with the model forecast (WWIII wave height or FVCOM water level) for that location.

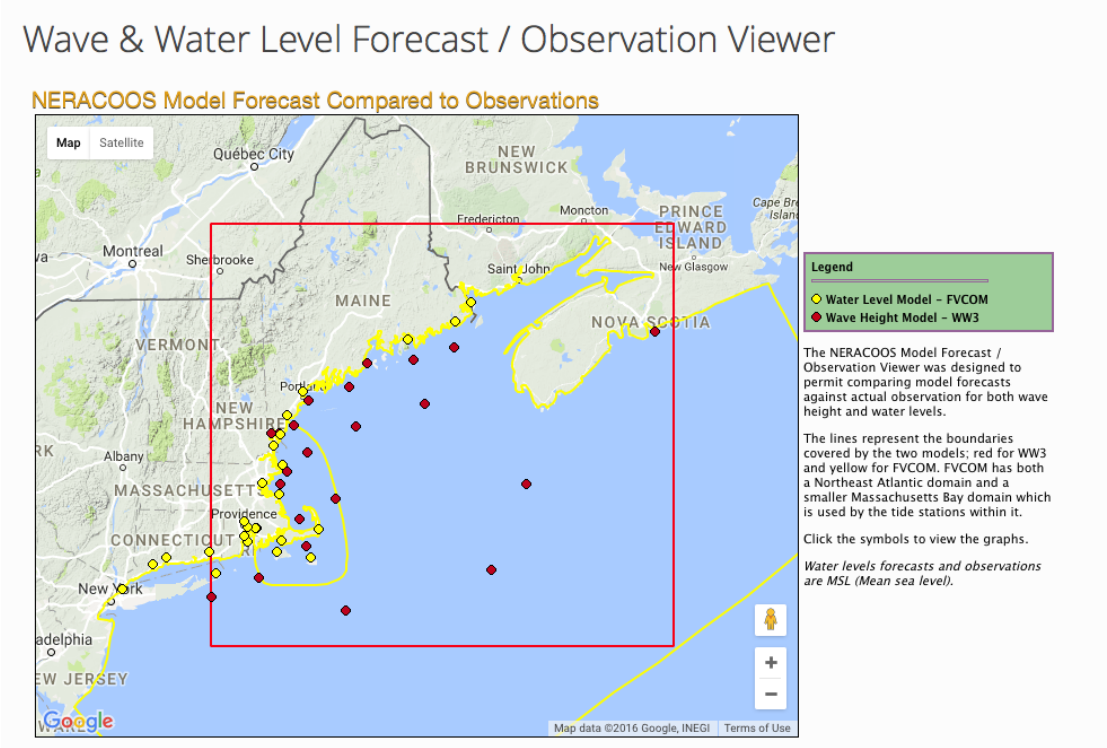


Figure 15: Screen capture of the NERACOOS wave & water level viewer entry page.

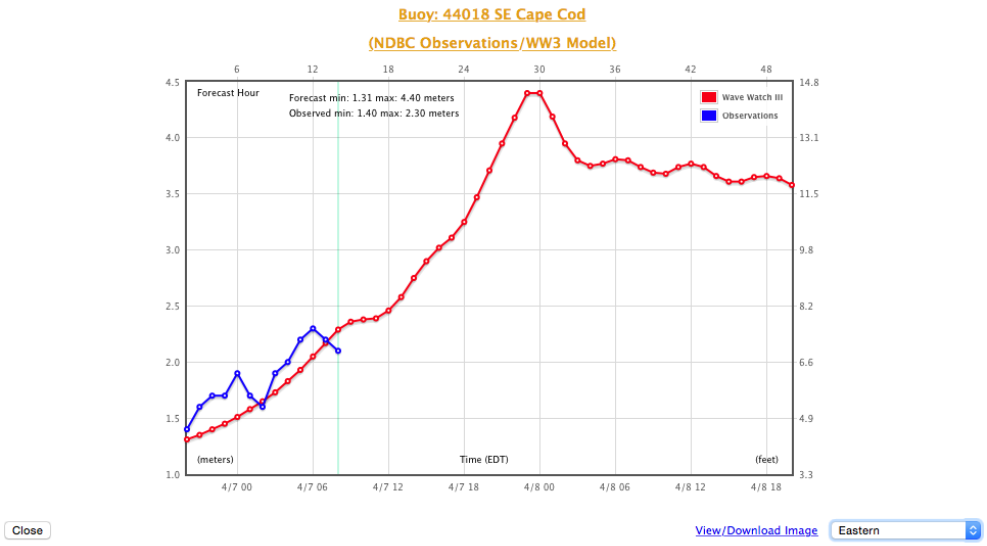


Figure 16: Screen capture of the NERACOOS wave & water level viewer output.

7.11 Satellite Imagery

This product allows user to view satellite imagery of sea surface temperature and chlorophyll concentration provided by the Satellite Oceanography Data Laboratory at UMaine.

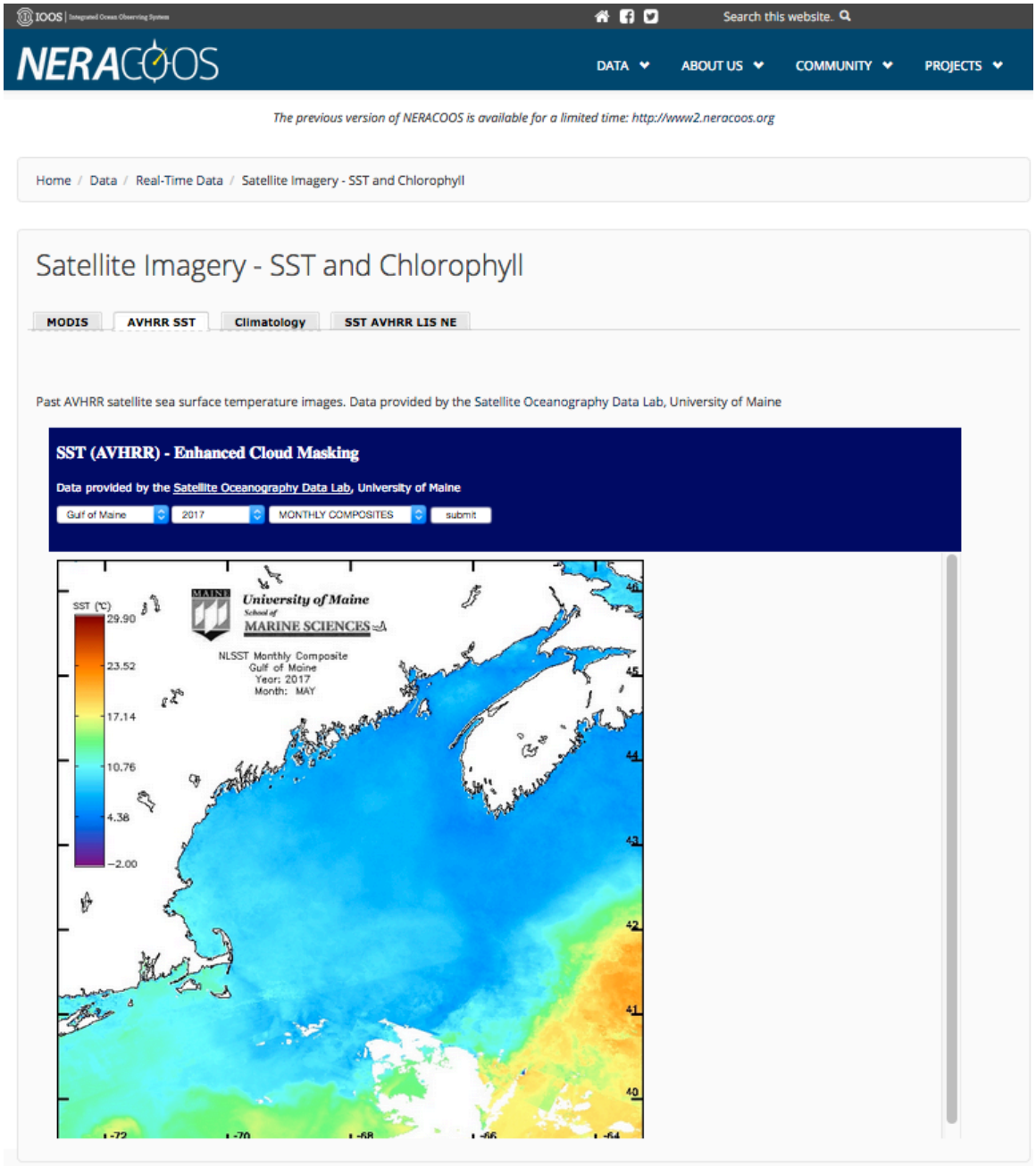


Figure 17. Screen capture of the NERACOOS satellite imagery product.

7.12 Surface Current

This product, developed by UCSD, provides a regional visualization of surface current data from HF Radar stations.

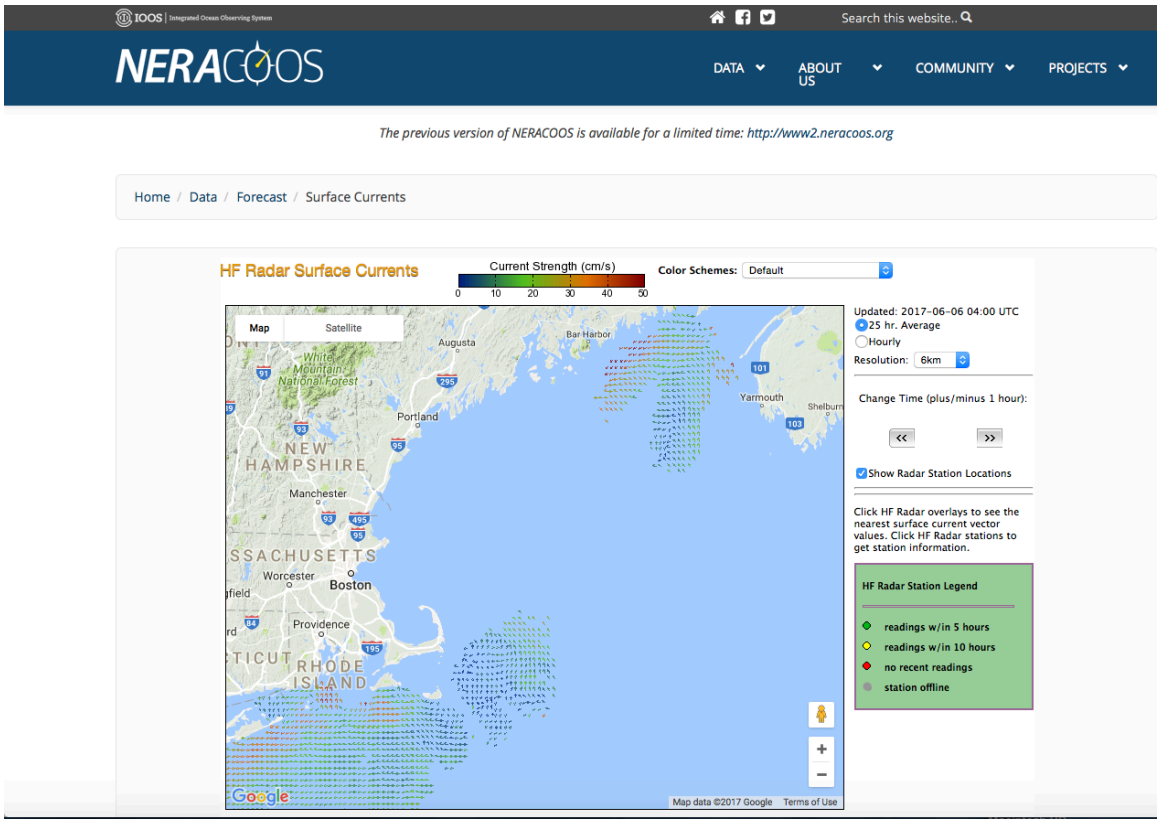
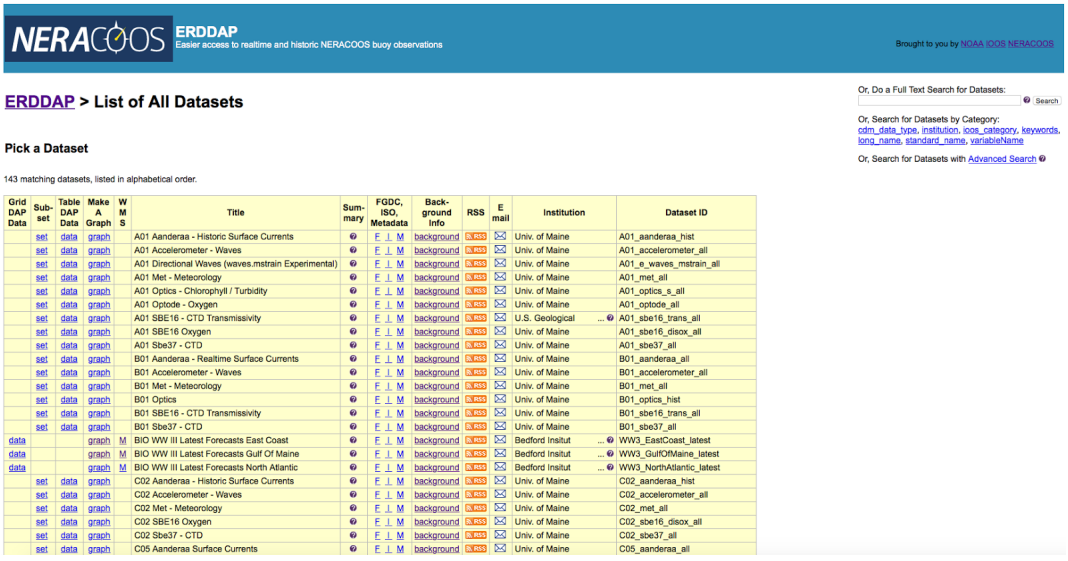


Figure 18. Screen capture of the NERACOOS HF Radar visualization map.

7.13 NERACOOS ERDDAP

NERACOOS has implemented an ERDDAP (the Environmental Research Division's Data Access Program) data server that provides a simple, consistent way to download subsets of scientific datasets in common file formats and make graphs and maps.



Grid DAP Data	Sub-set	Table DAP Data	Make A Graph	WMS	Title	Sum-mary	FGDC, ISO, Metadata	Back-ground Info	RSS	E mail	Institution	Dataset ID
	set	data	graph		A01 Aanderaa - Historic Surface Currents		F I M	background	RSS		Univ. of Maine	A01_aanderaa_hist
	set	data	graph		A01 Accelerometer - Waves		F I M	background	RSS		Univ. of Maine	A01_accelerometer_all
	set	data	graph		A01 Directional Waves (waves.mstrain Experimental)		F I M	background	RSS		Univ. of Maine	A01_e_waves_mstrain_all
	set	data	graph		A01 Met - Meteorology		F I M	background	RSS		Univ. of Maine	A01_met_all
	set	data	graph		A01 Optics - Chlorophyll / Turbidity		F I M	background	RSS		Univ. of Maine	A01_optics_s_all
	set	data	graph		A01 Optode - Oxygen		F I M	background	RSS		Univ. of Maine	A01_optode_all
	set	data	graph		A01 SBE16 - CTD Transmissivity		F I M	background	RSS		U.S. Geological ...	A01_sbe16_trans_all
	set	data	graph		A01 SBE16 Oxygen		F I M	background	RSS		Univ. of Maine	A01_sbe16_disox_all
	set	data	graph		A01 Sbe37 - CTD		F I M	background	RSS		Univ. of Maine	A01_sbe37_all
	set	data	graph		B01 Aanderaa - Realtime Surface Currents		F I M	background	RSS		Univ. of Maine	B01_aanderaa_all
	set	data	graph		B01 Accelerometer - Waves		F I M	background	RSS		Univ. of Maine	B01_accelerometer_all
	set	data	graph		B01 Met - Meteorology		F I M	background	RSS		Univ. of Maine	B01_met_all
	set	data	graph		B01 Optics		F I M	background	RSS		Univ. of Maine	B01_optics_hist
	set	data	graph		B01 SBE16 - CTD Transmissivity		F I M	background	RSS		Univ. of Maine	B01_sbe16_trans_all
	set	data	graph		B01 Sbe37 - CTD		F I M	background	RSS		Univ. of Maine	B01_sbe37_all
data			graph	M	BIO WW III Latest Forecasts East Coast		F I M	background	RSS		Bedford Insitut ...	WW3_EastCoast_latest
data			graph	M	BIO WW III Latest Forecasts Gulf Of Maine		F I M	background	RSS		Bedford Insitut ...	WW3_GulfOfMaine_latest
data			graph	M	BIO WW III Latest Forecasts North Atlantic		F I M	background	RSS		Bedford Insitut ...	WW3_NorthAtlantic_latest
	set	data	graph		C02 Aanderaa - Historic Surface Currents		F I M	background	RSS		Univ. of Maine	C02_aanderaa_hist
	set	data	graph		C02 Accelerometer - Waves		F I M	background	RSS		Univ. of Maine	C02_accelerometer_all
	set	data	graph		C02 Met - Meteorology		F I M	background	RSS		Univ. of Maine	C02_met_all
	set	data	graph		C02 SBE16 Oxygen		F I M	background	RSS		Univ. of Maine	C02_sbe16_disox_all
	set	data	graph		C02 Sbe37 - CTD		F I M	background	RSS		Univ. of Maine	C02_sbe37_all
	set	data	graph		C05 Aanderaa Surface Currents		F I M	background	RSS		Univ. of Maine	C05_aanderaa_all

Figure 19: Screen capture of the NERACOOS ERDDAP list of data sets.

7.14 Website Analytics

NERACOOS uses Google analytics to monitor Activity on our website. Google analytics allows us to track overall usage and look at the usage of specific products as well as how users navigate within our website.

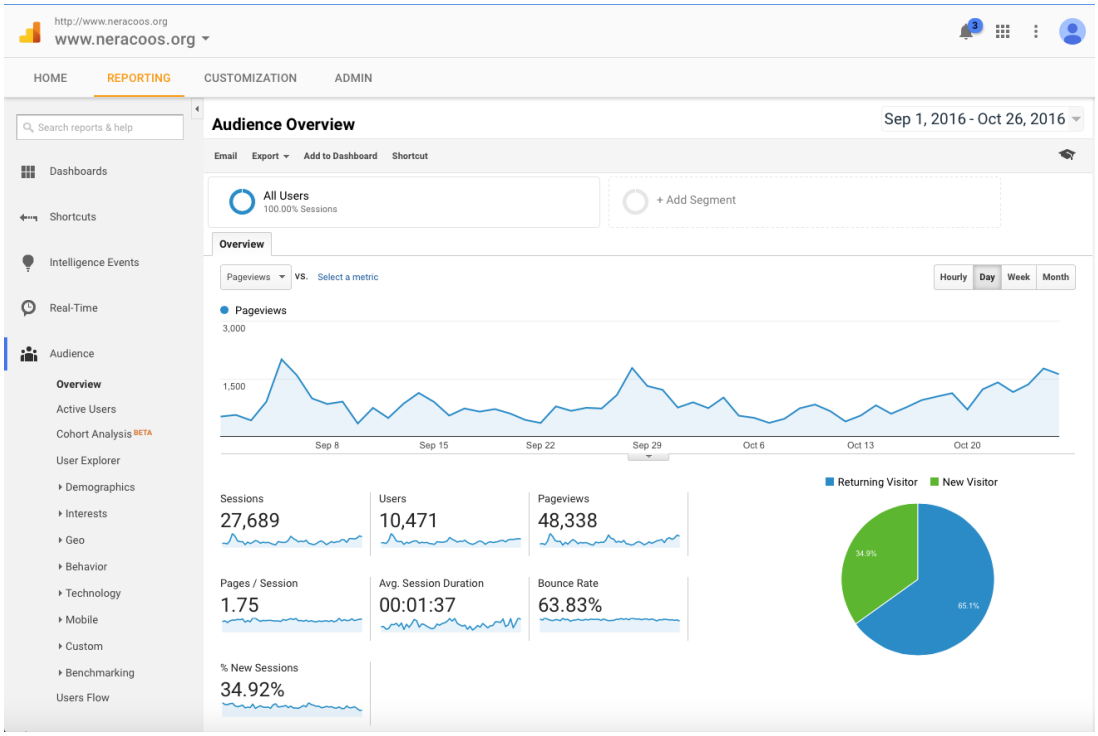


Figure 20: Screen capture of NERACOOS Google Analytics dashboard.

8 Data Backup and Archival

Security of NERACOOS data is ensured using the “here, near, there” paradigm outlined by the Long Term Ecological Research (LTER) guidance on digital research data and documents (http://intranet2.lternet.edu/content/protecting-your-digital-research-data-and-documents). NERACOOS funded data providers keep a local version of the raw data (here); submit data to the NERACOOS data management system (near), and will ultimately archive data with the NCEI (there). The GMRI Data Manager is responsible for working with funded partners to implement processes around data handling, curation and submission to the NERACOOS data management system and NCEI.

8.1 Data Provider Backup

NERACOOS funded data providers maintain local versions of NERACOOS data and backup data using their organization's backup systems, which are described in section 6. The table below provides an overview of the data providers local backup systems.

Table 3: Data provider data backup procedures.

Data Provider	Backup Protocol
Charybdis Group	All data are maintained on a cloud based hosted service and archive on raid 6 and higher disk arrays and periodically archived in accordance to best industry standards.
University of Rhode Island (URI)	URI data files are backed up on the base station computer at Marine Ecosystems Research Laboratory building within URI-GSO.
University of Connecticut (UConn)	Data files are backed up to a local external hard drive at UConn Avery Point Campus and onto UConn's Enterprise File Services network drive, located offsite on the UConn's Storrs campus.
University of Maine	Once daily the entire directory tree on the application server (which contains historical and post-recovery as well as real-time netCDF data files is replicated to the web server for distribution. Both the application server and the web server are backed up daily by incremental backup and weekly by full backup to rotating hard drives on a backup server located in a different on-campus building.
University of New Hampshire (UNH)	Data files are archived on the local UNH computer, an external hard drive, and backed up to a cloud data storage service (Box@UNH).

8.2 NERACOOS Data System Backup

The data stored on the NERACOOS Amazon EC2 environment is backed up using Amazon Simple Storage Service (S3). This includes netCDF files from providers, data stored in the observation databases and website content. The scalable, secure solution provides nightly incremental backups of the entire Amazon EC2 instances. GMRI manages the account on Amazon EC2 and has full access to configure and deploy instances and manage the back up systems.

Data are retained over a 30-day revolving window. From this scalable, secure solution data can be recovered from a point in time during that window and restored quickly in the event of a system failure.

The collection of software code that includes data processing scripts, web services and web-delivered applications and products is managed internally through version control systems. GMRI uses Git, which is managed, in an online private repository on Bitbucket (https://bitbucket.org/). GMRI maintains a library of technical documentation on Google Docs. These documents include detailed descriptions of the Amazon Web Services configuration, protocols for disaster recovery, data framework configuration and product development.

8.3 Archival at NCEI

NERACOOS is currently working with NCEI to establish a process of submitting data to NCEI. NERACOOS data providers will create netCDF files compatible with IOOS CF 1.6 and NCEI netCDF Templates v2.0 data formats and metadata standards. Historical data will be produced in yearly or monthly files and made available on an http accessible web site by the data providers. NERACOOS has established a Web Accessible Folder (WAF) to allow NCEI to pull the data as needed. Each data provider's data will be submitted to NCEI via the single NERACOOS WAF. Newer, near real-time files will be created monthly and stored on the WAF site. NCEI will pull these files monthly. NERACOOS will provide guidance and tools to the data providers to ensure their netCDF files meet the standards required.

We are implementing this system in a phased approach and starting with UMaine. The NCEI Request to Archive has been initiated with the University of Maine, which oversees the buoy program for both NERACOOS and CariCOOS. Please see the NCEI Request to Archive document (http://www.neracoos.org/sites/neracoos.org/files/documents/NCEI_Request_to_Archive_NERACOOS.pdf) and the NERACOOS Submission Agreement Document (http://www.neracoos.org/sites/neracoos.org/files/documents/NCEI_Submission_Agreement_NERACOOS.pdf).

At this time, the University of Maine effort is entering the submission-testing phase. We will leverage the outcome of this initial effort with NCEI to provide a model for our other data providers. We anticipate the system to submit data from UMaine to NCEI will be fully implemented by December 2017. After UMaine is completed we will work with UNH to implement their archival process and then work with UConn to implement their process. We anticipate that the UNH archival implementation will be complete by May 2018 and the UConn implementation will be completed by August 2018.

9 Appendix I: NERACOOS Data Source Table

Provider Type Key: NF = NERACOOS Funded, F/S = Federal/State, P = Private

Observations

Provider	Type	Locations	Provider Type		
			NF	F/S	P
UMaine Buoys	Real Time Historic	A01 Mass Bay B01 Western Maine Shelf E01 Central Maine Shelf F01 Penobscot Bay I01 Eastern Maine Shelf M01 Jordan Basin N01 Northeast Channel *C01 Casco Bay *J01 Eastport *K01 Saint Johns, NB *L01 Yarmouth, NS (*buoys no longer deployed, historical records available)	X		
UMaine HF Radar	Real Time	Grand Manaan Island, New Brunswick Cape St. Mary, Nova Scotia Green Island, Maine	X		
UConn Buoys	Real Time Historic	44022 Execution Rocks Long Island Sound 44039 Central Long Island Sound 44040 Western Long Island Sound 44060 Eastern Long Island Sound LDLC3 New London Ledge Light	X		X
URI Buoys and Sensors	Real Time Historic	Narragansett Bay - Mount View Narragansett Bay - Quonset Point			X
UNH Buoys and Sensors	Real Time Historic	CML Coastal Marine Lab Field Station CO2 Appledore Island GREAT_BAY Great Bay, NH	X		
Charybdis Group Tide Gauges	Real Time Historic	Hampton, NH Tide Gauge Gloucester, MA Tide Gauge Scituate Tide, MA Gauge	X		
WHOI HF Radar	Real Time	Race Point, MA	X		

WHOI Sensors	Real Time Near Real Time	MVCO SeaNode Martha's Vineyard, MA ESP HAB - several locations, Gulf of Maine			X
NOAA NEFSC Sensors	Historic	eMOLT - various locations, Gulf of Maine		X	
NOAA NDBC Buoys and Stations	Real Time Historic	44005 Cashes Ledge 44007 Casco Bay 44008 Nantucket 44011 Georges Bank 44013 Boston Harbor 44017 23 Nautical Miles SW of Montauk Point, NY 44018 SE Cape Cod 44020 NANTUCKET SOUND 44027 Jonesport, ME CMAN Stations BUZM3 Buzzards Bay * IOSN3 Isle of Shoals * MDRM1 Mt Desert Rock * MISM1 Matinicus Rock *		X	
NERRS Sensors	Real Time Historic	BGXN3 Great Bay Reserve, NH NAXR1 Naragansett Bay Reserve, RI NAXR1 Narragansett Bay Reserve, RI WAXM3 Waquoit Bay Reserve, MA WEXM1 Wells Reserve, ME		X	
Environment Canada Buoys	Real Time	44137 East Scotia Slope * 44150 La Have Bank * 44258 Halifax Harbor *		X	
NOAA CMAN	Real Time	CMAN MDRM1 - Mt Desert Rock CMAN MISM1 - Matinicus Rock CMAN IOSN3 - Isle of Shoals CMAN BUZM3 - Buzzards Bay		X	
CDIP Buoys	Real Time	CDIP154 Block Island, RI 44098 Jeffrey's Ledge (UNH) CDIP Cape Cod Bay	X	X	
SmartBay Canada	Real Time	SB01 Mouth of Placentia Bay, NL, Canada SB04 Placentia Bay, NL, Canada		X	
NOAA NOS CO-OPS Water Level	Real Time	8410140 Eastport, ME 8411250 Cutler Naval Base, ME 8413320 Bar Harbor, ME 8418150 Portland, ME 8419317 Wells, ME 8423898 Fort Point, NH 8443970 Boston, MA 8447386 Fall River, MA		X	

		8447387 Borden Flats Light at Fall River, MA 8447930 Woods Hole, MA 8448725 Menemsha Harbor, MA 8449130 Nantucket Island, MA 8447435 Chatham, MA 8452660 Newport, RI 8454049 Quonset Point, RI 8452944 Conimicut Light, RI 8452944 Providence, RI 8461490 New London, CT 8465705 New Haven, CT 8467150 Bridgeport, CT 8510560 Montauk, NY 8516945 Kings Point, NY Scituate Harbor, MA (NWS)			
USGS Water Services Water Level	Real Time	Saco, ME Tide Gauge		X	

Models

Provider	Type	Description	NF	F/S	P
UMass	Forecast Historic	NECOFS/FVCOM oceanographic model system Northeast Atlantic and Mass. Bay domains. (via TDS SOS at NERACOOS)	X		
BIO	Forecast	WW3 wave models for Northeast Atlantic Domain. (via TDS SOS at NERACOOS)	X		
UConn	Forecast Nowcast	HFR STPS (nowcast/forecast)	X		
UMaine	Forecast	GOM Circulation model (currents, temp, salinity)			X
NOAA NOMADS/ NCEP	Forecast	Wind Forecasts for Gulf of Maine (daily via OPenDAP)		X	
NOAA/ NOMADS/ ESTOFS	Forecast	Extratropical Surge and Tide Operational Forecast System		X	

Satellite

Provider	Type	Description	NF	F/S	P
UMaine	Near Real Time	Sea surface temperature, chlorophyll imagery Historic, climatology	X		X
NOAA Coastwatch	Near Real Time			X	