

Request to Archive
With The National Centers for Environmental Information
For NERACOOS and CariCOOS non-Federal Station Assets through the University of
Maine
Provided by University of Maine SMS

2017-02-17

This information will be used by NCEI to conduct an appraisal and make a decision on the request.

1. Who is the primary point of contact for this request?

Bob Fleming
University of Maine SMS
Scientist, Systems Administrator
207-581-4379
bfleming@umeoce.maine.edu
email or phone is fine.

2. Name the organization or group responsible for creating the dataset.

IOOS > Integrated Ocean Observing System, NOAA, U.S. Department of Commerce, Caribbean Coastal Ocean Observing System (CariCOOS), Northeastern Regional Association of Coastal Ocean Observing Systems (NERACOOS)

3. Provide an overview summarizing the scope of data you want to archive. Describe the outputs, data variables, including their measurement resolution and coverage.

CariCOOS and NERACOOS are two of 11 Regional Associations established nationwide through the NOAA Integrated Ocean Observing System (IOOS). IOOS coordinates the multi-agency, cooperative effort to routinely collect realtime data and manage historical information based on a continuously operating network of buoys, ships, satellites, underwater vehicles, and other platforms. These data are needed for many purposes which include rapid detection and prediction of changes in our nation's ocean and coastal waters.

CariCOOS and NERACOOS will be starting the automation process with a group of non-federal in-situ data sets. These data sets are from two continuously operating networks of buoys in the Caribbean Sea and in the Gulf of Maine. Both buoy networks are operated by the Physical Oceanography Group at the University of Maine.

Continuously collected historical data since the creation of the buoys networks will be submitted by University of Maine for archival initially, with periodic ongoing submissions of realtime data to continue in the future from operating buoys.

These data are needed and used for many purposes which include:

- a) rapid detection and prediction of short- and long-term changes in our nation's ocean and coastal waters.
- b) providing data for assimilation into realtime weather forecast models.
- c) providing offshore meteorological and oceanographic data for storm hazard assessment
- d) assessment of weather and sea state for marine search and rescue, commercial shipping, fishing activities, and recreational boating.
- e) validation of atmospheric and oceanographic model data.

f) validation of satellite data

providing historical data for planning and engineering activities for transport, coastal infrastructure, and offshore mineral and wind energy production

g) assessment of marine ecosystem health

in-situ observational data includes (but is not limited to) the following data types:

wInd speed and direction, barometric pressure, air temperature, visibility, significant wave height, dominant wave period, mean wave direction, subsurface water temperature, subsurface salinity, subsurface water density, subsurface conductivity, surface currents, subsurface currents (ADCP profiles), dissolved oxygen

Physical, Meteorological, Surface Measurements, Current Measurements, Moored Buoy

4. What is the time period covered by the dataset? (YYYY-MM-DD, YYYY-MM or YYYY)

From 2001-07

Ongoing as continuous updates to the data record

5. Edition or version number(s) of the dataset:

N/A

6. Approximate date when the dataset was or will be released to the public:

2017-06

7. Who are the expected users of the archived data? How will the archived data be used?

Oceanographers, Integrated Ocean Observing System affiliates, climate researchers, offshore wind power engineers & researchers, commercial fishermen, commercial & military mariners, search and rescue personnel, civil & marine engineers, coastal planners, emergency response personnel, marine biologists, harbor pilots.

8. Has the dataset undergone user evaluation and/or an independent review process? Did NCEI participate in design reviews?

No

9. Describe the dataset's relationship to other archived datasets, such as earlier versions or related source data. If this is a new version, how does it improve upon the previous version(s)?

N/A

10. List the input datasets and ancillary information used to produce the data.

Buoy operator, University of Maine, maintains instrumentation and telemetry and generates and applies realtime and postprocessing QA/QC to the primary data. Type of instrumentation and sampling parameters are provided within the netcdf files, in accordance with Climate Forecasting (CF) convention guidelines.

11. List web pages and other links that provide information on the data.

<http://www.neracoos.org>

<http://cara.uprm.edu/>

<http://dm1.caricoos.org/thredds/catalog/buoys/catalog.html>

http://www.caricoos.org/drupal/data_download

<http://gyre.umeoce.maine.edu/caricoos/>

<http://gyre.umeoce.maine.edu/buoyhome.php>

12. List the kinds of documents, metadata and code that are available for archiving. For example, data format specifications, user guides, algorithm documentation, metadata compliant with a standard such as ISO 19115, source code, platform/instrument metadata, data/process flow diagrams, etc.

1. CF 1.6 compliant, ACDD 1.3 compliant, netcdf-3 files containing data, metadata, and information on platform and instrumentation.

13. Indicate the data file format(s).

1. netCDF-3
2. NCEI netCDF Templates v2.0

14. Are the data files compressed?

No

15. Provide details on how the files are named and how they are organized (e.g., file_name_pattern_YYYYMM.tar in monthly aggregations).

files are organized by mooring site/station, mooring deployment, and instrument package / depth designator if applicable.

SSS/DDDDD.IIIIII.III.merged.nc , where:

SSS: "Mooring site" (or "station") refers to a geographic location where a series of buoys are deployed sequentially. For this data set, all mooring sites are designated by a 3-character string, named and described in the "station" variable, and associated with a specific wmo_id attribute). Examples: "A01", "PR2".

DDDDD: "Mooring deployment", a designator for a specific deployment of a buoy at a mooring site. All mooring deployments are designated by a 5-character string (named in the "platform_1" variable in each netcdf file), which comprises the 3-character station followed by a 2 digit chronological serial number. Individual deployments typically span a few months to slightly more than a year. Specific instruments, sensors, and sampling parameters do not change or vary during a deployment. Examples: A0101, A0102, A0103, PR201, PR202

IIIII.III Instrument package/ depth designator (variable length). Instrument groups in the initial data submission include:

met, waves.summit, waves.axys, currents.002m, currents.adcp, ocean.001m, ocean.020m, position.gps

"merged": constant value, does not change.

Example filenames: A0101.currents.adcp.merged.nc, PR103.met.merged.nc, PR103.ocean.001m.merged.nc

Files are organized into folders based on mooring site (station).

Each folder contains all data files for that station.

Each data file comprises all the data for an instrument or instrument group for an entire deployment.

16. Explain how to access sample data files and/or a file listing for previewing. If it is not available now, when will it be available?

A partial set of samples is available now with more to follow to complete a representative set. The data will be provided through a Web Accessible Folder (WAF) to NCEI for preview.

WAF: <http://gyre.umeoce.maine.edu/ncei>

files are located in two top-level folders

<http://gyre.umeoce.maine.edu/ncei/caricoos>

<http://gyre.umeoce.maine.edu/ncei/neracoos>

and in subfolders designated by 3 character mooring site (station_name) indicating a particular mooring site.

17. What is the total data volume to be submitted?

Historic Data: all historic data or data submitted as a completed collection.

Total Data Volume: 7GB

Number of Data Files: 2100

Continuous Data: data volume rate for a continuous data production.

Total Data Volume Rate: 200MB per Year

Data File Frequency: 130 per Year

Data Production Start: 2017-05

18. Are later updates, revisions or replacement files anticipated? If so, explain the conditions for submitting these additional data to the archive.

Realtime data are processed and distributed in a short time frame (minutes to hours) by buoy operator. When mooring platforms and instrumentation is retrieved from the ocean, typically months after realtime data was available, higher resolution or formerly missing data may be available from recovered instruments or dataloggers. This data may be resubmitted to the archive with a similar file name (as described in #15 above, regarding file naming)

The packages will be established as follows:

New, never-before seen data files will be archived based on which mooring site they are associated with: each mooring site will be assigned an accession number.

New, data from a previously submitted buoy: The AIP for that buoy will be updated (NODC's major-revision) with the new data file.

Revised, data that was previously submitted that needs to be updated: If the naming conventions match and the checksums do not match, then the most recent submission of that file will be assumed to be the latest and greatest submission and will replace the previous file.

19. Describe the server that will connect to the ingest server at NCEI for submitting the data.

Physical Location: University of Maine, Orono, Maine

System Name: gyre.umeoce.maine.edu

System Owner: UMAINE SMS PhoG> University of Maine Physical Oceanography Group

Additional Information: open http access to Web Accessible Folder

20. What are the possible methods for submitting the data to NCEI? Select all that apply.

1. FTP PULL

NCEI will pull from UMaine web accessible folder <http://gyre.umeoce.maine.edu/ncei/>.

21. Identify how you would like NCEI to distribute the data. Web access support depends on the resources available for the dataset.

1. Direct download links

2. Advanced web services (e.g., THREDDS Catalog Service)

22. Will there be any distribution, usage, or other restrictions that apply to the data in the archive?

No known constraints apply to the data.

23. Discuss the rationale for archiving the dataset and the anticipated benefits. Mention any risks associated with not archiving the dataset at NCEI.

Data Set fits within NOAA's mission. OSTP: Increasing Access to the Results of Federally Funded Scientific Research. Collocation of this data within a national, searchable repository for other meteorological and environmental data sets has obvious benefits for accessibility.

24. Are the data archived at another facility or are there plans to do so? Please explain.

Primary data, raw telemetry and operating information remains with UMaine. Processed near-realtime data is distributed to NERACOOS and CariCOOS in realtime and may be archived by those organizations. Subsets of the data are transferred to the National Data Buoy Group (NDBC) in realtime and may be available there.

25. Is there an existing agreement or requirement driving this request to archive? Have you already contacted someone at NCEI?

There is an SLA between the IOOS Program Office and NCEI to archive non-federal station data. We've been in contact with Mathew Biddle at NCEI.

26. Do you have a data management plan for your data?

No

27. Have funds been allocated to archive the data at NCEI?

Through the Integrated Ocean Observing System Data Discover, Access, and Archiving Project.

28. Identify the affiliated research project, its sponsor, and any project/grant ID as applicable.

Physical Oceanography Group, School of Marine Sciences, University of Maine

29. Is there a desired deadline for NCEI to archive and provide access to the data?

No deadlines for archive or access.

30. Add any other pertinent information for this request.

Please include Mathew Biddle in discussions concerning this data set.