ERDC/CHL TR-18-11

Coastal and Hydraulics Laboratory

US Army Corps of Engineers®
Engineer Research and Development Center

ERDC
ENGINEER RESEARCH & DEVELOPMENT CENTER

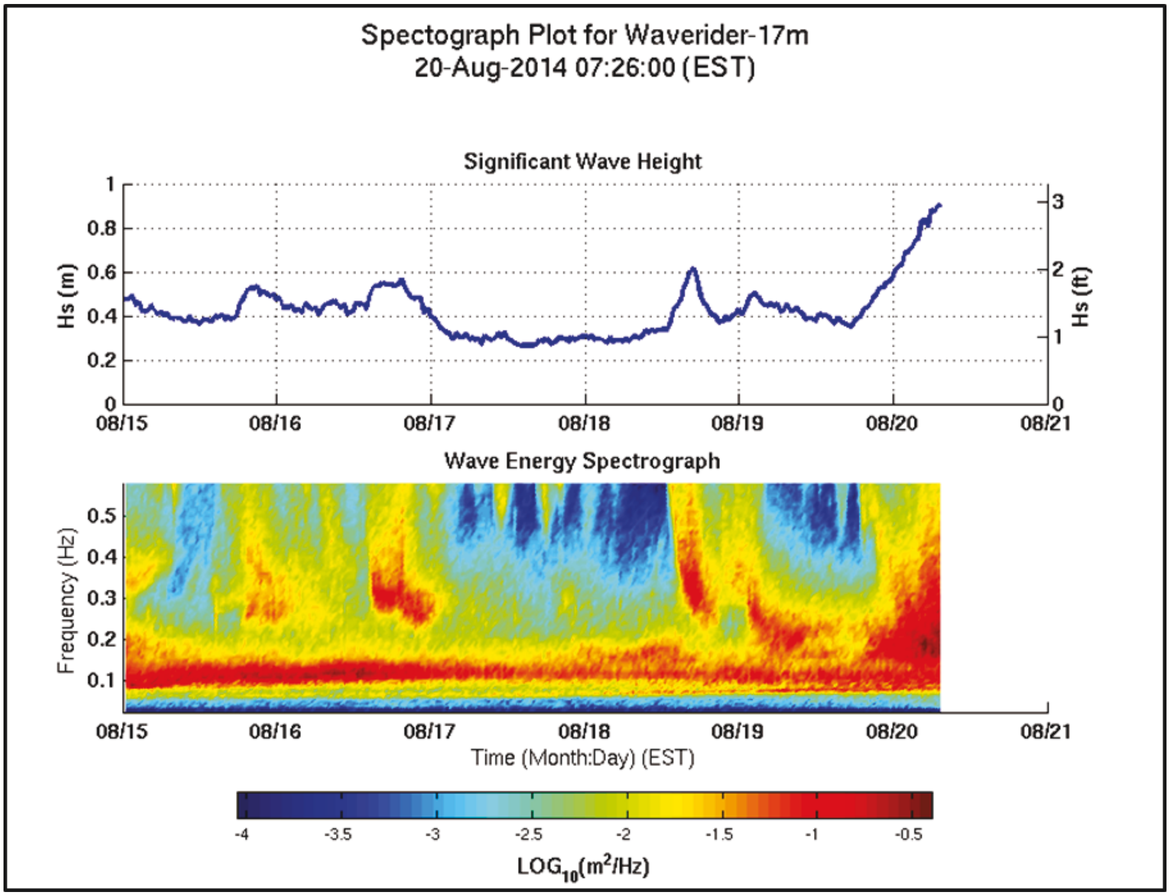
Coastal Ocean Data Systems Program

Oceanographic Observations Dataset

Data Management Plan

Kent K. Hathaway, Jeffrey L. Hanson, Kelly Knee, Andrew Bird, and Robert Fratantonio

August 2018



Coastal Ocean Data Systems Program
Navigation Business Line

ERDC/CHL TR-18-11
August 2018

Oceanographic Observations Dataset

Data Management Plan

Kent K. Hathaway and Jeffrey L. Hanson

Field Research Facility
U.S. Army Engineer Research and Development Center
1261 Duck Road
Kitty Hawk, NC 27949

Kelly Knee, Andrew Bird, and Robert Fratantonio

RPS Applied Science Associates
55 Village Square Drive
Wakefield, RI 02879

Final report

Prepared for U.S. Army Corps of Engineers
Washington, DC 20314-1000

Under Work Unit H70H6B, "Oceanographic Observations Dataset: Data Management Plan"

Abstract

This Data Management Plan details the Oceanographic Observations dataset, which is maintained at the U.S. Army Corps of Engineers (USACE) Engineer Research and Development Center Coastal and Hydraulics Laboratory (ERDC-CHL) Field Research Facility (FRF), Duck, NC. The plan was developed to support the FRF Data Integration Framework (FDIF) project. Information is organized in the following categories: general description, points of contact, data stewardship, data documentation, data sharing, initial data storage and protection, long-term archiving and preservation, hardware and software requirements, products/programs, tools, references, data catalog, and abbreviations and acronyms.

Contents

Figures and Tables

Figures

Tables

Preface

This study was conducted for Headquarters, U.S. Army Corps of Engineers (HQUSACE) under the USACE Coastal Ocean Data Systems (CODS) Program, work unit H70H6B, “Oceanographic Observations Dataset: Data Management Plan.” The technical monitor was Dr. J. P. Waters. Mr. Jeffrey A. McKee was the HQUSACE Navigation Business Line Manager overseeing the (CODS) Program.

The work was performed by the Coastal Observations and Analysis Branch (COAB) (CEERD-HFA) of the Flood and Storm Protection Division (CEERD-HF), ERDC-CHL. At the time of publication, Dr. J. P. Waters was Chief, CEERD-HFA; Dr. Cary A. Talbot was Chief, CEERD-HF; and Mr. W. Jeff Lillycrop (CEERD-HT) was the Technical Director for Civil Works and Navigation Research, Development, and Technology Transfer (RD&T) portfolio. The Acting Deputy Director was John P. Tucker III, and the Acting Director of ERDC-CHL was Mr. Jeffrey R. Eckstein.

COL Ivan P. Beckman was the Commander of ERDC, and Dr. David W. Pittman was the Director.

Unit Conversion Factors

Multiply	By	To Obtain
degrees (angle)	0.01745329	radians
feet	0.3048	meters
miles (nautical)	1,852	meters
miles (U.S. statute)	1,609.347	meters
yards	0.9144	meters

Abbreviations and Acronyms

1D—One-dimensional
2D—Two-dimensional
ADOP—Nortek Aquadopp
AMD—Advanced Micro Devices
ASA—(See RPS ASA)
ASCII—American Standard Code for Information Interchange
AST—Acoustic Surface Tracking
AWAC—Nortek Acoustic Wave and Current
CDIP—Coastal Data Information Program
CEERD-HF— U.S. Army Corps of Engineers, Engineer Research and Development Center, Flood and Storm Protection Division
CEERD-HF-A— U.S. Army Corps of Engineers, Engineer Research and Development Center, Flood and Storm Protection Division, Coastal Observations and Analysis Branch
CERC—Coastal Engineering Research Center
CESAM-OP-J— U.S. Army Corps of Engineers, Mobile District, Operations Division, Spatial Data Branch
CHL—Coastal and Hydraulics Laboratory
COAB—Coastal Observation and Analysis Branch
CPC—Central Processing Center
CSV—Comma-Separated Values (file format)
CS-W—Catalog Service for the Web
CTD—Conductivity, Temperature, and Depth
DAP—Data Access Protocol
DAQ—Data Acquisition
DMP—Data Management Plan
DOD—U.S. Department of Defense
ERDC—U. S. Army Engineer Research and Development Center
ERDC-CHL—U. S. Army Engineer Research and Development Center, Coastal and Hydraulics Laboratory
ERDDAP—National Oceanic and Atmospheric Administration (NOAA) Environmental Research Division's Data Access Program
FDIF—Field Research Facility Data Integration Framework
Fortran—Formula Translating System
FRF—Field Research Facility
GB—Gigabyte
GHz—Gigahertz
HF—High frequency

Hz—Hertz
ID—Identification
IEEE—Institute of Electrical and Electronic Engineers
IMLE—Interactive Maximum Likelihood Indicator
ISO—International Organization for Standardization
ISO/TC—International Organization for Standardization Technical Committee
J. Geophys. Res.—*Journal of Geophysical Research*
J. Mar. Res.—*Journal of Marine Research*
JPEG—Joint Photographic Experts Group (file format)
JSON—JavaScript Object Notation
km—Kilometer
m—meters
MAT—MATLAB (file format)
MATLAB—Matrix Laboratory
MB—Megabyte
MET—Meteorological
min—Minute
N/A—Not Available
NC—NetCDF (file format)
NDBC—National Data Buoy Center
NetCDF—Network Common Data Format
NOAA—National Oceanic and Atmospheric Administration
NODC—National Oceanic Data Center
OGC—Open Geospatial Consortium
OS—Operating system
PAR—Photosynthetically Active Radiation
PII—Personally Identifiable Information
PNG—Portable Network Graphics (file format)
POC—Point of Contact
PUV—Pressure and UV-current (wave analysis method)
QA/QC—Quality Assurance/Quality Control
RAM—Random Access Memory
RD&T— Research, Development, and Technology Transfer
RPS ASA—RPS Applied Science Associates
SOS—Sensor Observation Service
SUV—Surface-track and UV-current (wave analysis method)
THREDDS—Thematic Real-time Environmental Distributed Data Services
TIFF—Tagged Image File Format (file format)
TR—Technical Report

UI—User Interface
UNESCO—United Nations Educational, Scientific and Cultural Organization
USACE—U.S. Army Corps of Engineers
UV—Horizontal Orthogonal U and V Currents
WIS—Wave Information Studies
Xshore—Cross-shore

1 General Description

1.1 Background

The U.S. Army Engineer Research and Development Center's Coastal and Hydraulics Laboratory (ERDC-CHL) operates the Field Research Facility (FRF), an internationally recognized coastal observatory established in 1977 and located on the Atlantic Ocean near Duck, NC (Figure 1). The ERDC-CHL Coastal Observations and Analysis Branch (COAB) maintains a number of oceanographic datasets at the FRF, including winds, waves, water levels, currents, and beach morphology. Together, these datasets provide a 34-year record of coastal conditions and have become an important national resource for the coastal engineering and scientific communities.

Figure 1. The Field Research Facility, Duck, NC.

To protect and preserve this important national archive, COAB was charged with developing a Field Research Facility Data Integration Framework (FDIF) that would provide a modern, discoverable archive with easily accessible information for strategic analysis and reporting. To accomplish this task, COAB partnered with the U.S. Army Corps of Engineers (USACE) Mobile District (SAM), Spatial Data Branch (CESAM-OP-J), and Applied Science Associates (RPSASA). Along with the FRF data, ERDC-CHL Wave Information Studies (WIS) data are also being incorporated into the FDIF. It is intended that the results of this project will provide a framework that can be used by other data management groups within both ERDC-CHL and its District partners.

1.2 Objectives

The objective of this plan is to address the management of oceanographic observations data within the FDIF. Additional plans will be prepared for the management of survey and remote sensing data (e.g., radar, lidar, and camera imagery).

1.3 Approach

This project is being conducted in two phases. Phase I addresses continuous datasets, which are being collected continuously to form a historical record. Phase II addresses special project datasets, which were collected during the historic series for coastal experiments conducted at the FRF by federal agencies, academia, and private-sector companies.

1.4 Scope

This version of the plan addresses the management of only the FRF continuous oceanographic observations data. Once Phase II is in operation, the plan will be updated to include management of the special project datasets.

2 The dataset

2.1 Dataset name

The name of the dataset is the Field Research Facility Oceanographic Observations.

2.2 Data keywords

Critical keywords describing this dataset are listed below. These keywords were taken from the Global Change Master Directory Data Center, Instruments, and Science and Services Dictionaries (http://gcmd.nasa.gov/learn/keyword_release.html).

- Atmosphere > Atmospheric Pressure > Atmospheric Pressure Measurements
- Atmosphere > Atmospheric Radiation > Solar Radiation
- Atmosphere > Atmospheric Temperature > Air Temperature
- Atmosphere > Atmospheric Water Vapor > Dew Point Temperature
- Atmosphere > Atmospheric Water Vapor > Humidity
- Atmosphere > Atmospheric Winds > Surface Winds > Wind Speed/Wind Direction
- Atmosphere > Atmospheric Winds > Wind Chill
- Atmosphere > Precipitation > Rain
- DOD > DOD/USARMY/USACE/CHL/FRF > Field Research Facility, Coastal And Hydraulics Laboratory, U. S. Army Corps of Engineers, U.S. Army, U. S. Department of Defense
- Oceans > Bathymetry/Seafloor > Topography > Bathymetry
- Oceans > Bathymetry/Seafloor > Topography > Water Depth
- Oceans > Ocean Acoustics > Acoustic Velocity
- Oceans > Ocean Circulation > Ocean Currents
- Oceans > Ocean Optics > Photosynthetically Active Radiation
- Oceans > Ocean Optics > Secchi Depth
- Oceans > Ocean Pressure > Water Pressure
- Oceans > Ocean Temperature > Water Temperature
- Oceans > Ocean Temperature > Sea Surface Temperature
- Oceans > Ocean Waves > Wave Frequency
- Oceans > Ocean Waves > Wave Height
- Oceans > Ocean Waves > Wave Length
- Oceans > Ocean Waves > Wave Period

- Oceans > Ocean Waves > Wave Spectra
- Oceans > Ocean Waves > Wave Speed/Direction
- Oceans > Ocean Winds > Surface Winds
- Oceans > Salinity/Density > Conductivity
- Oceans > Salinity/Density > Density
- Oceans > Salinity/Density > Salinity
- Oceans > Sea Surface Topography > Sea Surface Height
- Oceans > Tides > Tidal Height
- Profilers/Sounders > CTD > Conductivity, Temperature, Depth

2.3 Data description

2.3.1 Waves

Hourly wave data from the numerous FRF wave gauges (e.g., the Nortek Acoustic Wave and Current (AWAC) and Aquadopp (ADOP) gauges, Datawell Waverider buoys, and FRF 8 meter (m) pressure sensor array) are stored as monthly MATLAB binary files. These files are converted into NetCDF files containing both the wave spectra and the bulk wave statistics (height, period, and direction). For non-directional wave gauges, only the 1D spectra are available in the wave spectra NetCDF file.

- Wave data from 1980 to 1986 were non-directional, using Datawell Waverider buoys approximately 3–5 kilometers (km) offshore at nominal depths of 15–18 meters (m) and Baylor Wave Staffs at several locations under the pier. The non-directional Waverider buoy was upgraded to a Datawell Directional buoy (Mk III) in 1997. While the original buoy data were collected with radio transmission of the vertical displacements in an analog format, the newer directional buoys transmit horizontal and vertical displacements in digital form. COAB created custom software to acquire and process this data.
- In 2008, a second directional Waverider was deployed approximately 16 km offshore of the FRF at a nominal depth of 26 m. Data from this buoy was telemetered via an Iridium satellite link through the Coastal Data Information Program (CDIP) (http://www.cdip.ucsd.edu). The FRF nearshore buoy was upgraded to an Iridium telemetered Datawell buoy (via CDIP) in 2013. Since 2013, buoy wave spectra and statistics have been computed at 30-minute (min) intervals from 30-min time series, using directional Fourier coefficients computed onboard the buoy. COAB processes these coefficients through an Iterative Maximum Likelihood Estimator (IMLE) (Pawka 1983; Oltman-Shay

and Guza 1984) to obtain 2D spectra that are archived in monthly MATLAB files and then converted to NetCDF files. Occasionally, the IMLE produces a poor directional estimation, and the direct Fourier method is substituted (Longuet-Higgins et al. 1963).

- Five Baylor wave staffs, surface-piercing gauges that make direct measurements of the sea surface, were deployed under the pier in the early 1980s to provide cross-shore measurements of wave height. This first-principle measurement was expected to provide highly accurate, non-directional wave data. However, uncertainties in water depth under the pier from the scour trough, which changed throughout a storm, and wave interference with the pilings made the data difficult to interpret. Nevertheless, the data showed cross-shore variation of wave height from many storms. It was expected that the Baylor gauges would measure storm surge and wave setup. However, it was discovered that the gauge means were dependent on the power supply voltage, which varies with fluctuations on the supply and line loss, which, in turn, was a function of the wire temperature. These mean water levels are flagged as suspect; however, they are still provided for surge estimates. A user could make pre- and post-storm comparison to compute offsets and improve surge estimates.
- In 1986, a bottom-mounted array of 10 pressure sensors was deployed at an 8 m depth northeast of the pier with an alongshore dimension of 255 m. This formed a linear phased array that provided high-directional resolution wave spectra. However, a linear array has only a 180° range to detect wave direction, based on the assumption that all wave energy comes from offshore. In 1990, the array was modified with five additional cross-shore pressure sensors, using a cross-shore dimension of 120 m. This made the array two-dimensional, with a 360° directional range, and thus able to detect reflected energy from the shore. The array design was based on the work of Davis and Regier (1977). The 8 m Array operated between 1986 and 2012 and was rebuilt in July 2015. IMLE analysis is used to compute the two-directional wave spectra. A detailed explanation of the array and analysis method can be found in Long and Oltman-Shay (1991, 1993).
- In 2008, a cross-shore array of wave gauges was created with bottom-mounted Nortek AWAC and Aquadopp current profilers deployed between 2 m and 11 m depths (Figure 2). Seaward of the bottom-mounted gauges are the two Datawell Waverider bouys, approximately 3.2 km offshore at 17 m depth and 16 km offshore at 26 m depth. Near the shelf break is a National Data Buoy Center (NDBC) directional

wave buoy with meteorological instruments. The NDBC buoy is approximately 100 km from the FRF and in 46 m water depth. Funding for the NDBC buoy is provided by ERDC-CHL.
- The COAB created customized directional analysis programs for the processing of AWAC time series, using a combination of methods in spectra computation. It was shown that analysing the data as a point source and as a spatial array and then merging the two spectra provided an improvement over the individual methods (Hathaway and Hanson 2011).
- Spectral wave analysis use the Welch (1967) method of overlapping and windowing data segments to compute multiple spectra that are then averaged to produce an estimate of the spectral energy density.

AWACs have four acoustic beams, one vertical and three slanted at 25°. The vertical beam, which is called the Acoustic Surface Tracking (AST) beam, tracks the surface displacements, sampling at 4 hertz (Hz), which is twice that of the other beams. The lower end of the spectrum uses the orbital currents from the three oblique velocity beams, converted to horizontal and vertical orthogonal currents, and the AST signals to compute wave frequency-direction spectra. An IMLE is used to estimate the directional distributions. This is termed Surface and horizontal (UV) currents (SUV) method, and it is treated as a point measurement. The upper end of the spectrum uses IMLE processing of the AST and individual radial-velocity beams, treating it as a spatial array (Array method). If the AST signal fails quality checks, the pressure (P) signal is substituted (PUV method). Frequency cutoffs are computed based on the horizontal spatial separation of the velocity beams (H) and wavelength (L). The SUV (or PUV) method is exclusively used at frequencies below L = 10 H, and the Array method is used exclusively at frequencies above L = 6 H (Figure 3). Between these frequencies, a linear combination of the two spectra is used. A maximum frequency cutoff (f_c) is set at the lesser frequency value when the wavelength is two lag distances (L = 2 H) or when linear wave theory wave pressure attenuation is 0.1. IMLE is used for computing 2D directional spectra.

Figure 2. FRF Cross-shore array.

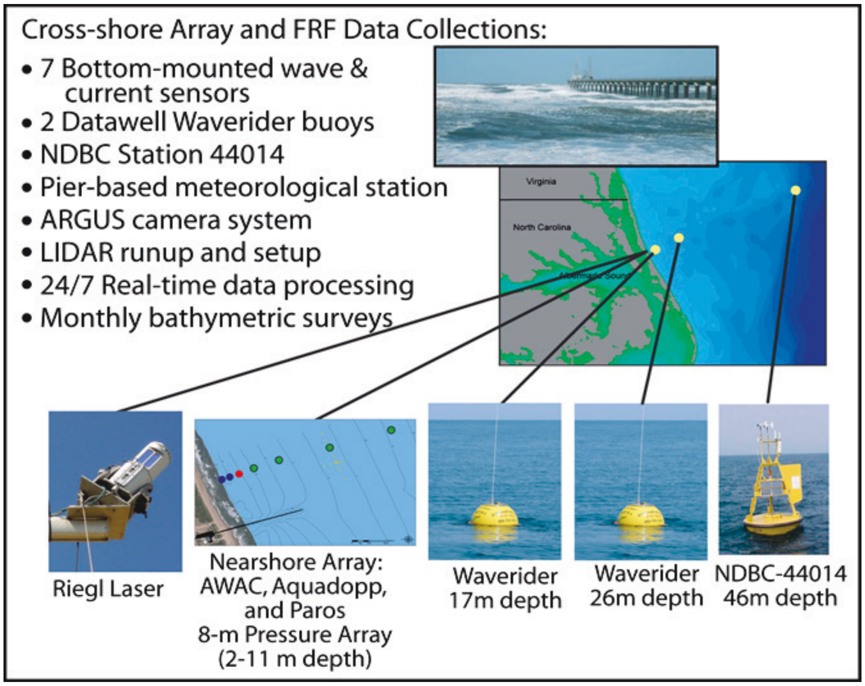
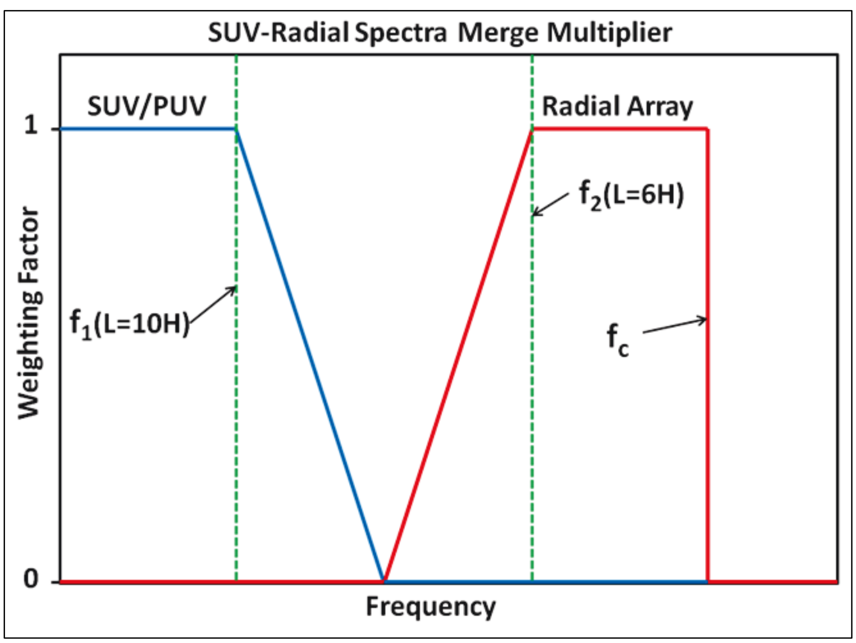


Figure 3. AWAC merged spectra weighting factor.



2.3.2 Meteorology

2.3.2.1 Winds

Wind data is recorded from multiple wind sensors, typically four, each of which produce datasets of raw time series sampled at 1 or 2 Hz. These data are saved in American Standard Code for Information Interchange (ASCII) and MATLAB binary files. In addition, a derived 10-minute statistics dataset compiles the best data from the four sensors. Each of these datasets is converted to NetCDF.

2.3.2.2 Other meteorological (MET) data

All other meteorological (MET) data are available as a 10-min statistics-derived dataset stored in monthly ASCII text files and are converted to NetCDF. The following MET data types are available:

- Average temperature
- Average humidity
- Dew point
- Wind chill or heat index
- Barometer
- Photosynthetically Active Radiation (PAR)
- Solar radiation
- Wind gust (highest 5 second (sec) average), sustained (highest 1 min average), and 10 min average; vector-averaged wind direction
- Rain

Current profiles are collected either hourly or every two hours (hr) from the AWAC and ADOP gauges. The data is available in MATLAB binary file format and converted to NetCDF. Averaging duration is typically 10 minutes.

2.3.3 Oceanographic

2.3.3.1 Conductivity, temperature, and depth (CTD) casts

Conductivity, temperature, and depth (CTD) profiles are collected daily at the FRF pier from an Ocean Sensor's CTD gauge. Sound speed, water density, and salinity are computed from the CDT data using the UNESCO 1978 standard (UNESCO 1981). The data is stored in two datasets: a monthly statistics file in MATLAB binary format and a monthly raw profile in ASCII text format. Both datasets are converted to NetCDF format.

2.3.3.2 Water visibility

Water visibility is recorded daily (weather permitting) at the end of the FRF pier (approximately 450 m offshore) using a Secchi disk. The entire

30+ year record is recorded in a single MATLAB binary file and is converted into a single NetCDF file.

2.3.3.3 End of pier water temperature

Water temperature is recorded hourly at the end of the FRF pier at six different depths and stored as monthly files in raw ASCII text format and MATLAB binary statistics format. Both datasets are available in NetCDF format.

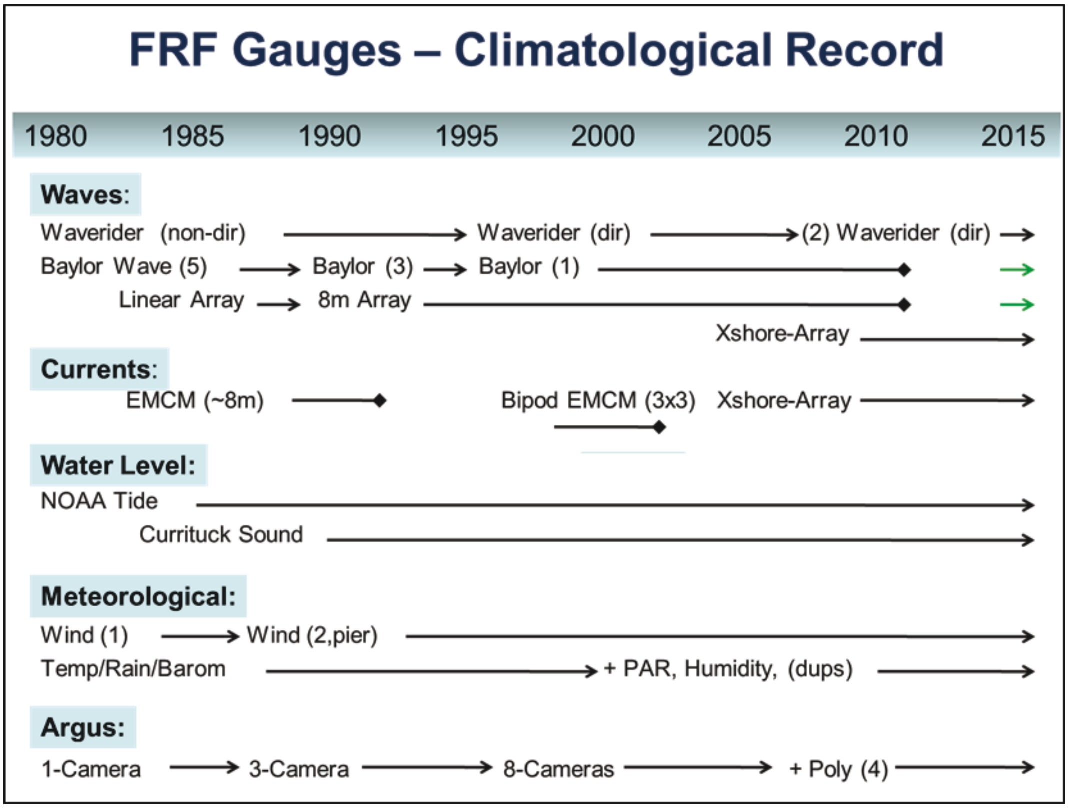
2.3.4 Water level

Water level measurements, predictions, and differences are recorded in 6-min intervals for two locations: the FRF pier and Currituck Sound. File formats include ASCII text and NetCDF. Water level data from the pier are obtained from the National Oceanic and Atmospheric Administration (NOAA) station (ID 8651371) located at the pier's seaward end. Predicted tides are computed from NOAA tidal coefficients (Fortran program supplied by Chuck Brown of NOAA, 1991).

2.4 Data temporal extent

The temporal extent of the data varies according to location and sensor type. The longest FRF dataset covers the period from 1979 to the present. Figure 4 illustrates the temporal coverage of the various data types included in the FRF data record. The data inventory provides exact time spans for each dataset integrated into the repository to date.

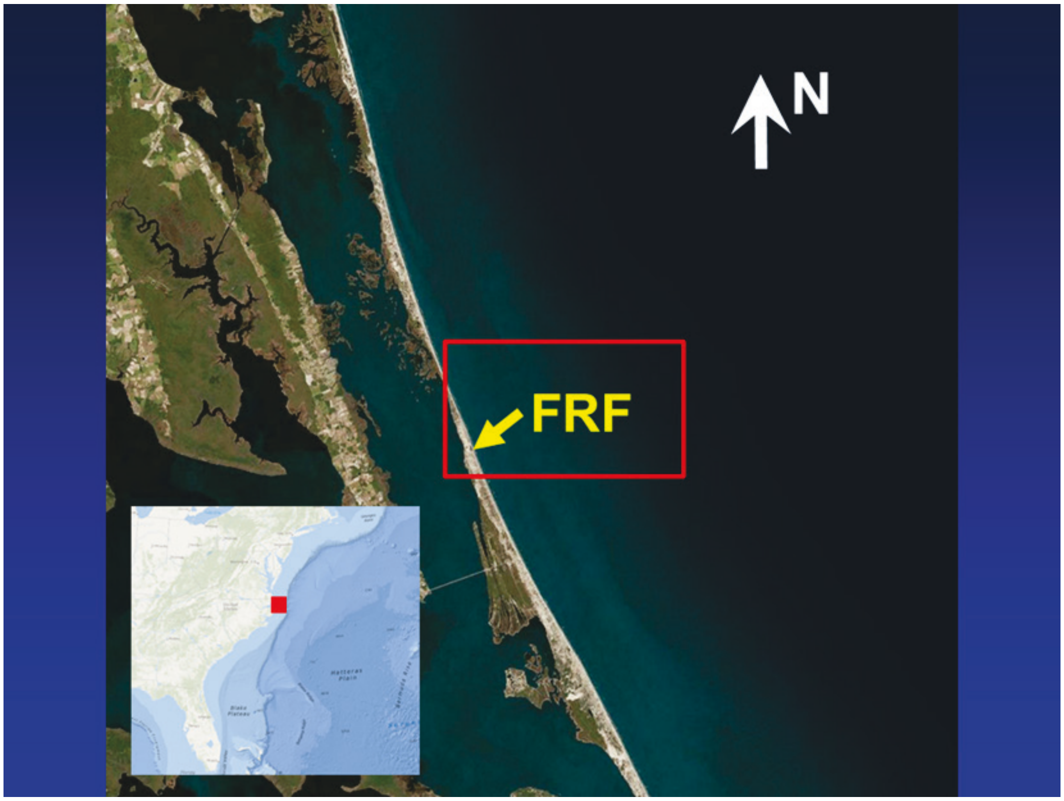
Figure 4. Oceanographic observation record.



2.5 Data geographic extent

The final data archive is expected to cover multiple locations in U.S. coastal waters. However, the datasets that have been integrated into the oceanographic data repository to date are limited to the region around the FRF in Duck, NC. The encompassing bounding box is 75.742 ° W Longitude, 36.186 ° N Latitude, 75.560 ° W Longitude, and 36.260 ° N Latitude, as shown in Figure 5.

Figure 5. Geographic extent of data.



2.6 Data types

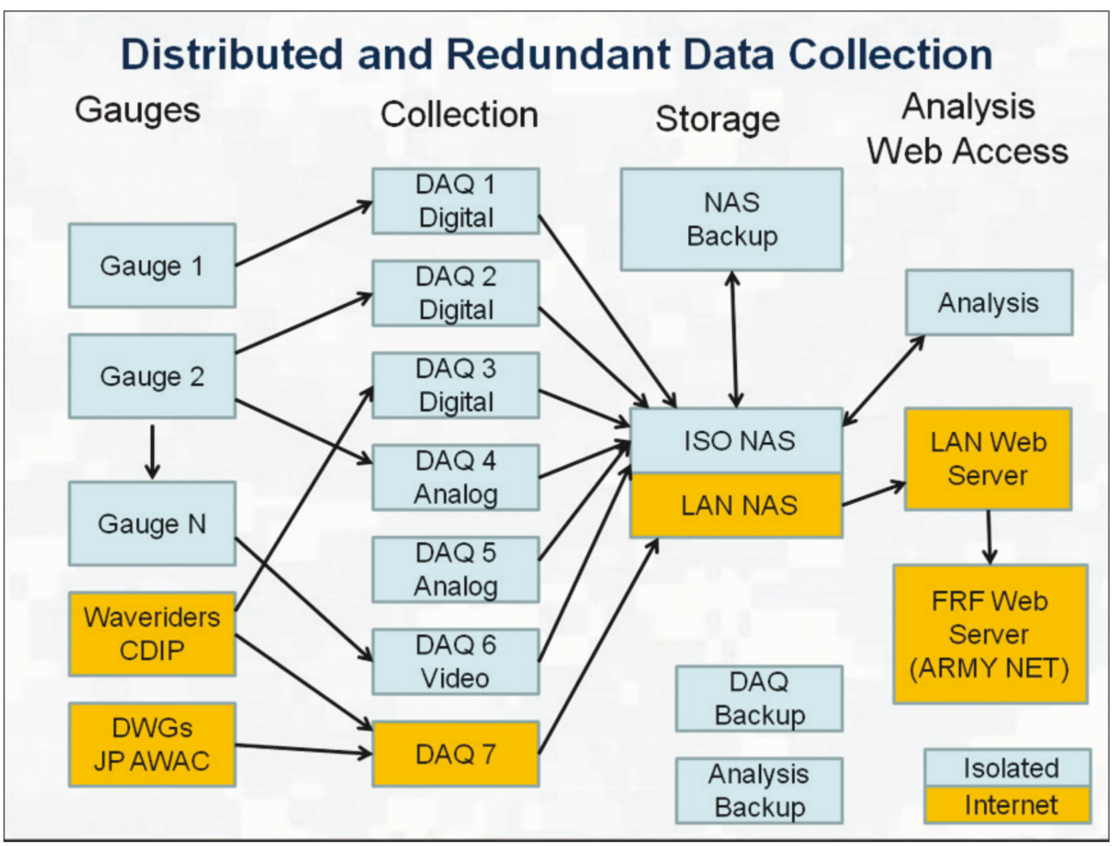
The following data types have been integrated into the oceanographic observations data repository to date:

- Wave bulk statistics: Wave height, period, and direction
- Wave spectra data: Frequency (1D) and frequency-direction spectra (2D)
- Current data: Velocity profiles and vertically integrated velocities
- Wind data: 10 min averaged speed, direction, sustained, and gust
- Meteorological parameters: Air temperature, humidity, pressure, precipitation, PAR, and solar radiation (10 min averaged)
- CTD profiles (daily, weather dependent)
- Water temperature profiles
- Water level (NOAA tide and others)
- Water visibility (daily Secchi depth measurements)

2.7 Data capture/creation method

Data is captured by a variety of meteorological and oceanographic instrumentation, and derived, spectral, and statistics datasets are created using a variety of technologies. An instrumentation list is provided in Appendix A. Data acquisition (DAQ) is accomplished in a five-tiered approach, as depicted in Figure 6. The instruments with only analog, only digital, and both analog and digital output are in the first tier. The second tier consists of the data acquisition computers with either multi-port serial interfaces or analog-to-digital converters. Prior to 1992, all data were analog; since 1992, an increasing number of gauges have provided digital data streams, which are now the most common signal type at the FRF. Data collection is performed almost entirely with custom in-house programs. The third tier is the storage of the *raw* data, the fourth tier is the data analysis, and the fifth tier is the analysis archive and web access. Automatic backups of the data acquisition computers and data archives are executed daily. When possible, data collection and processing are isolated from the Internet.

Figure 6. FRF tiered data acquisition with internet connected (blue) and internet isolated (yellow) computers.



Analog data collections from 1980 to mid-1987 were generally taken at 6-hr intervals. When storm conditions (defined as wave H_{mo} greater than 2 m at the pier end) existed, the interval was shortened to 3 hr. Most data were collected at 2 Hz, occasionally at 4 Hz, with corresponding record lengths of 34 or 17 min. In February 1987, four contiguous records (136 min) were collected at 6-hr intervals (3-hr intervals during storms). Beginning in January 1991, five contiguous records (171 min) were recorded at 3-hr intervals, leaving a data gap of approximately 9 min every three hours. In March 2013, all data collection became hourly with an approximate three-second gap between one-hour records.

Digital data collections for non-wave systems (e.g., meteorological data and water temperature) are continuously measured at rates ranging from 0.1 to 1 Hz and archived in monthly files (see the individual gauge metadata for sampling details). Wave data are recorded at 2 Hz for bottom-mounted gauges (4 Hz for the Nortek AWAC Acoustic Surface Track signal) with hourly 34-min records. Datawell Directional Waverider spectra are from contiguous 30-min records. The new directional buoys internally record raw time series at 1.28 Hz. These time series were downloaded during servicing and archived after processing. Directional wave data from the FRF Waverider buoy 630 were collected via high-frequency (HF) radio between December 1996 and March 2013 in hourly 53.3-min records (4,096 points at 1.28 Hz). Hourly directional wave spectra were computed from the 53-min records. After that time, the buoy's internally processed 30-minute records were used.

2.8 Data volume

An estimated 11 Gibabytes (GB) of data have been collected annually since 2008 for a present total of approximately 130 GB. Almost 90 GB of data are currently integrated into the oceanographic observations data repository with the remaining added by the end of FY16.

2.9 Data restrictions

Personally Identifiable Information (PII) or other information whose distribution may be restricted by law or national security is not included in this dataset.

2.10 Data storage plan

These data will be stored at the ERDC, Vicksburg, MS CHL Thematic Real-time Environmental Distributed Data Services (THREDDS) server. COAB will maintain local storage of the original and processed data at the FRF, Duck, NC.

3 Points of Contact (POC)

The physical address of all POCs is USACE Field Research Facility, 1261 Duck Road, Kitty Hawk, NC 27949.

Project POC

Michael F. Forte, a Research Physical Scientist, is the project POC. (252) 261-6840 x228 or at Michael.F.Forte@usace.army.mil.

Overall POC

Jeffrey P. Waters, Chief, Coastal Observations and Analysis Branch, is the overall POC. (252) 261-6840 x233 or at Jeffrey.P.Waters@usace.army.mil.

Data Quality POC

Kent K. Hathaway, Research Oceanographer, is the data quality POC. His phone number is (252) 261-6840 x224, email address is Kent.K.Hathaway@usace.army.mil.

Data Collection POC

Kent K. Hathaway, Research Oceanographer, is the data collection POC. See above for contact information.

Documentation/Metadata POC

Kent K. Hathaway, Research Oceanographer, is the documentation/metadata POC. Contact information can be found under Data Quality POC.

Data Storage/Disaster Recovery POC

Kent K. Hathaway, Research Oceanographer, is the data storage/disaster recovery POC. See above for contact information.

DMP Adherence/Implementation POC

Michael F. Forte, Research Physical Scientist, is the Data Management Plan (DMP) adherence, implementation POC. Contact information is listed under Project POC.

4 Data Stewardship

4.1 Quality control

A variety of Quality Assurance/Quality Control (QA/QC) tests are applied to the datasets, depending on the gauge type. These tests can be for a single data signal or an inter-comparison with other signals or gauges. A short summary of some QA/QC checks is presented below; specific tests are found with each dataset. Generally, the QA/QC follows the guidelines of the NOAA Integrated Ocean Observing System (IOOS) Quality Control of Real-Time Ocean Data (QARTOD) Standards (http://www.ioos.noaa.gov/qartod/welcome.html). Some additional tests are applied for non-standard gauges (e.g., the 8m Array).

Non-wave data:

- Thresholds (min/max for means, standard deviation, or variance)
- Variance of variance—stationarity check
- Spikes and jumps in the raw time series

Wave data:

- Thresholds (means, wave height, period, and direction)
- Wave height comparison with adjacent gauges (8m Array)
- Wave variance between sensors on the same gauge, such as vertical to horizontal displacement of the Datawell Waverider buoy or orbital currents to pressure in a current profiler
- Coherence between data channels (e.g., current and pressure)
- Signal-to-noise ratio (ratio of spectral peak to tail)
- Pressure spectra noise floor

As datasets are converted from MATLAB to NetCDF format, validation and verification scripts are run to ensure that no data is lost during the conversion process.

4.2 Data lifecycle

Real-time data are processed and stored internally on FRF servers within 10 min of completing collection. The data is generally publicly available within 1 hr, though latency can be up to 2.5 hr; however, steps are underway to reduce this delay. As real-time data feeds are integrated to the FDIF, this delay will be reduced to minutes. At that time, options for synchronizing the FRF server to match the shorter delay times of the FDIF server will be investigated and implemented.

5 Data Documentation

5.1 Metadata repository

Esri's Geoportal Server, which was set up by CESAM-OP-J, is used as the FDIF metadata repository. The Geoportal Server is a free, open source product for managing and publishing metadata. It facilitates discovery of resources, including datasets and web services.

The Geoportal Server can be used to create a Geoportal website, where data managers can register data resources for discovery by users. It also stores and catalogs metadata and retrieves information for each registered data resource. All metadata is contained in an Open Geospatial Consortium (OGC) compliant geoportal CS-W 2.0.2 catalog service.

In addition to Geoportal, metadata for the oceanographic data is also available as an ncISO service on a THREDDS server.

5.2 Additional metadata

Additional metadata may be included in the *Abstract* field.

5.3 Data/metadata standards

The data and metadata elements in this data collection are represented according to International Organization for Standardization (ISO) 19115-2, *Geographic Information—Metadata*, from International Organization for Standardization Technical Committee (ISO/TC) 211 Geographic Information/Geomatics. This standard is based on a tiered approach where metadata is divided into required and important categories.

6 Data Sharing

6.1 Data availability

Data in this dataset have been available to the public since 1979. Some of the data have been publicly available through the FDIF since February 2016. Data collection is ongoing, and there is an approximate 1 to 2.5-hour delay between data collection and publication.

6.2 Data availability restrictions

There are no data availability restrictions.

6.3 Data access restrictions

Users are subject to the standard *hold harmless* clause and *include credit* caveat. Standard language for providing credit is available to users on all websites developed to support the FDIF project.

6.4 Data access protocols

THREDDS serves the data in Data Access Protocol (DAP) and Sensor Observation Service (SOS) formats and allows it to be downloaded in NetCDF (.nc) format. The Environmental Research Division's Data Access Program (ERDDAP), which allows data to be downloaded in CSV, JSON, MAT, and dozens of other formats, may also be implemented.

6.5 Catalog registration

These data are expected to be registered in NOAA's National Geophysical Data Center catalog. Other catalogs, such as data.gov, will also be considered.

7 Initial Data Storage and Protection

7.1 Initial data storage location

Data are currently stored at the FRF in Duck, NC. Multiple copies of the raw data are saved. Data collection computers store the original data files for approximately one year, and hourly updates are made on the analysis computers and redundant storage computers. This copy is also archived on the main backup system. The FDIF data is also stored on a THREDDS server at the USACE Central Processing Center (CPC) in Vicksburg, MS; see Chapter 7, *Long-Term Archiving and Preservation*, for more details on THREDDS.

7.2 Data storage protection

Data protection is provided through availability on a mirrored system of the main archive that is backed up nightly. The mirrored system is housed in a separate building. Portable disk drives are made for backup annually and stored off site in the event of a catastrophic loss of the facility.

7.3 Data access protection

Access to restricted data is managed via password. Additional details will be provided if restricted data is designated to be included in the FDIF system. Otherwise, there are no limitations to data access. In the event of unauthorized access, copies are maintained on a non-Internet-accessible backup system as well as on-the-shelf backup disks.

8 Long-Term Archiving and Preservation

8.1 Data archiving

This data will be stored on a THREDDS server at the USACE CPC in Vicksburg, MS. The project team will also consider long-term archiving at the National Oceanic Data Center (NODC). However, discussions with the NODC have not yet been initiated. This data is publicly available at http://wisdata.erdc.dren.mil/thredds/catalog.html. To access data stored on the THREDDS server, users must have direct access to the server. Data cannot be edited or deleted through the THREDDS catalog. As such, only those data managers and administrators with access to the server will be able to modify or delete data. Any accidental modifications or deletions can be addressed by restoring data from the mirrored system or archive.

A backup copy of the data archive will be maintained at the COAB FRF.

Copies of Coastal Data Information Program (CDIP)-collected Waverider data are also maintained at CDIP.

8.2 Data archiving costs

Costs of long-term data archiving will be provided and maintained by the Coastal Field Data Collection Program.

8.3 Data archiving procedures

Data will be transferred automatically from the FRF to the THREDDS server in Vicksburg using a number of scripts, each designed to handle a specific data type provided by the FRF. Real-time NetCDF data files will be transferred to the Vicksburg THREDDS data server as they are generated. Historical data processed as part of the Phase I continuous database will be transferred to the Vicksburg server via external hard disk. Additional historical data processed following setup of the continuous data archive may also be transferred via external hard disk, depending on the file size.

8.4 Data retention period

Due to the importance of collecting and maintaining a long-term climatological record, the data retention period is indefinite.

9 Hardware and Software Requirements

9.1 Hardware requirements

Based on the expected oceanographic data volume, an initial minimum of 300 GB of disk storage is required. This breaks down to 250 GB for data storage and 50 GB for the OS and utilities. A minimum of 4 GB (RAM) and a 2 GHz or faster multiple-core Intel or Advanced Micro Devise (AMD) processor are also required.

9.2 Software requirements

9.2.1 Server software requirements

The server software requirements are as follows:

- THREDDS Data Server 4.3
- Java version 6+
- Tomcat version 6+
- ERDDAP
- Python 2.7
- MATLAB

9.2.2 Client-side software requirements

The client-side software requirements are the following:

- Web browser (i.e., Chrome, Firefox, or Internet Explorer)
- Data analysis tools of choice (i.e., MATLAB or Python)

10 Products/Programs

A number of plots and data products are available for the different data types. The highest priority plots include time series and wave spectra plots for those continuous oceanographic observations data parameters already received. Table 1 summarizes the plots included initially as well as those to be added in subsequent phases. Data export options include commonly used oceanographic observations data types, such as NetCDF (.nc), MATLAB (.mat), and CSV (.csv) files. Plots can be exported into .png, .jpeg, and .tiff formats.

Table 1. Summary of expected plots for FDIF Oceanographic Observations data.

Implementation Schedule	Plot Description
Phase 1	Bulk statistics time series for wind speed, wind direction, meteorology, wave heights, wave periods, water level, current magnitudes and directions, water quality parameter, and visibility
Phase 1	Wave 1D spectra plots (linear, log)
Phase 1	Wave 2D spectra plots (polar, Cartesian)
Phase 1	Current magnitude and direction stat plots
Phase 2	Current profiles
Phase 2	Current (u, v) time-stacks (time vs depth axes)
Phase 2	Physical water quality (temperature, salinity, sigma-t, and sound speed) time-stack plots
Phase 2 (Climatology Displays)	Wind, wave and current roses—All year, annual, seasonal, and monthly
Phase 2 (Climatology Displays)	Wind, wave and current histograms—All year, annual, seasonal, and monthly
Phase 2 (Climatology Displays)	Wave height and period scatter plots—All year, annual, seasonal, and monthly
Phase 3	Multiple gauge temporal stat plots (wind speed, wind direction, meteorology, wave heights, wave periods, water level, current magnitude and direction, water quality parameter, and visibility)
Phase 3	Multiple gauge spatial stat plots (wind speed, wind direction, meteorology, wave heights, wave periods, water level, current magnitude and direction, water quality parameter, and visibility)
Phase 3	Multiple gauge transect plot (Xshore array waves and currents)
Phase 3	Wave vector plot
Phase 3 (Climatology Displays)	Parameter extremal plot (monthly, annually)

11 Tools

11.1 Visualization tools

The FDIF User Interface (UI) tools will serve as the primary visualization tools for the FRF oceanographic observations data. The present vision for the UI is a map-based view that shows the location of all FRF assets and provides tabular and graphical summaries of the various data types.

Before the prototype UI was released in December 2014, IPython Notebooks were used for data visualization. An example, showing wave spectra and time series, is available for viewing at http://nbviewer.ipython.org/github/Bobfrat/notebooks/blob/master/plot_wave_spectrum_and_timeseries.ipynb.

11.2 Analysis tools

The analysis tools for oceanographic observations data were developed in Python and were implemented as IPython Notebooks until they were integrated with the UI.

11.3 Upload tools

All data are being pushed to the THREDDS data server using Python tools.

11.4 Download tools

Any interfaces offered by THREDDS or ERDDAP, including DAP tools, can be used to download data. THREDDS and ERDDAP servers can be accessed directly from MATLAB and Python tools as well as from ArcGIS and R (an open-source programming language) via middleware tools such as NOAA's Environmental Data Connector (http://www.pfeg.noaa.gov/products/EDC/).

11.5 Other tools

Validation and verification tools have been developed in Python to ensure that data is not lost during the conversion from MATLAB to NetCDF files.

References

Davis, R. E., and L. Regier. 1977. Methods of estimating directional spectra from multi-element arrays. *Journal of Marine Research*. 35(3): 453–478.

Hathaway, K. K., and J. L. Hanson. 2011. Wave transformation and setup from a cross-shore array of acoustic Doppler profilers. *Proceedings, 12th International Workshop on Wave Hindcasting and Forecasting*, 30 Oct–4 Nov, Kohala Coast Hawaii. (http://www.waveworkshop.org/12thWaves/index.htm, Session P12).

Long, C. E., and J. M. Oltman-Shay. 1993. *Preliminary estimates of frequency-direction spectra derived from the SAMSON pressure gage array, November 1990 to May 1991*. MP-CERC-93-3. Vicksburg, MS: U.S. Army Engineer Waterways Experiment Station, Coastal Engineering Research Center.

Long, C. E., and J.M. Oltman-Shay. 1991. *Directional characteristics of waves in shallow water*. TR-CERC-91-1. Vicksburg, MS: U.S. Army Engineer Waterways Experiment Station, Coastal Engineering Research Center.

Longuet-Higgins, M. S., D. E. Cartwright, and N. D. Smith. 1963. Observations of the directional spectrum of sea waves using the motions of a floating buoy. In *Ocean Wave Spectra, 111–136*. Englewood Cliffs, NJ: Prentice Hall.

Oltman-Shay, J., and R. T. Guza. 1984. A data adaptive ocean wave directional spectrum estimator for pitch and roll type measurements. *Journal of Physical Oceanography* 14(11): 1800–1810.

Pawka, S. S. 1983. Island shadows in wave directional spectra. *Journal of Geophysical Research* 88(C4): 2579–2591.

United Nations Educational, Scientific and Cultural Organization (UNESCO). 1981. The practical salinity scale 1978 and the international equation of state of seawater 1980. Tenth Report of the Joint Panel of Oceanographic Tables and Standards (JPOTS). *Marine Science* 36: 25.

Welch, P. D. 1967. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics* AU-15: 70–73.

Appendix A: Data Catalog

Waves

Description: Multiple wave gauges are used to measure wave height and direction at the FRF.

Table A1. Wave data

Source	#/ Stations	Frequency	Duration	Format	Archive Size	Products/Notes
AWAC	5	Hourly	Continual	MATLAB	22 GB	Stat series plots, 1D/2D spectra plots
AWAC	5	2 Hz	Continual	MATLAB	N/A	Raw time series
ADOP	2	Hourly	Continual	MATLAB	4 GB	Stat series plots, 1D/2D spectra plots
ADOP	2	2 Hz	Continual	MATLAB	N/A	Raw time series
Waverider	3	Hourly	Continual	MATLAB	23 GB	Stat series plots, 1D/2D spectra plots
Waverider	3	2 Hz	Continual	MATLAB	N/A	Raw time series
8m Array	1	Hourly	Continual	MATLAB	N/A	Stat series plots, 1D/2D spectra plots
8m Array	1	2 Hz	Continual	MATLAB	N/A	Raw time series
Baylor Staff		Hourly	Continual	MATLAB	N/A	Stat series plots, 1D spectra plots
Baylor Staff		2 Hz	Continual	MATLAB	N/A	Raw time series

Appendix B: Meteorology

Description: Multiple wind stations collect data at the FRF.

Table B1. Meteorology data

Source	#/ Stations	Frequency	Duration	Format	Archive Size	Products/Notes
Wind sensor	4	10 min	Continual	ASCII/MATLAB	248 MB	Stat series plot, magnitude/ direction stats plot, wind rose Note: A fifth dataset is derived from the best data of the 4 available sensors
Wind sensor	4	2 Hz	Continual	MATLAB	38 GB	Raw time series
Other MET (such as PAR, air temperature, humidity, rain, and solar radiation)	1	10 min	Continual	ASCII	22 MB	Stat series plot

Appendix C: Currents

Description: Nortek AWAC and ADOP current profilers provide 2D current profiles and raw time series datasets.

Table C1. Currents data

Source	#/ Stations	Frequency	Duration	Format	Archive Size	Products/Notes
AWAC	5	1 – 2 hr	Continual	MATLAB	137 MB	Magnitude/direction stats plot, profile plots, time stack plots, current rose
ADOP	2	1 – 2 hr	Continual	MATLAB	25 MB	Magnitude/direction stats plot, profile plots, time stack plots, current rose

Appendix D: Oceanographic

Description: Additional oceanographic observations datasets collected at the FRF include CTD profiles and water visibility.

Table D1. Oceanographic data

Source	#/ Stations	Frequency	Duration	Format	Archive Size	Products/Notes
Ocean Sensors OS200 CTD	1	Daily	Continual	MATLAB	16 MB	Profile plots, time stack plots Note: Depth variable and interpolated to standard depths
Ocean Sensors OS200 CTD	1	Daily	Continual	ASCII	11 MB	Raw time series
Temperature Sensor Array	1	1 hr	Continual	MATLAB	690 KB	Stat series plots, profile plots
Temperature Sensor Array	1	1 hr	Continual	ASCII	5 MB	Raw profile
Visibility Secchi Disk	1	Daily	Continual	MATLAB	45 KB	Raw time series

Appendix E: Water Level

Description: Water level measurements, predictions, and differences are recorded in 6-minute intervals for two locations: The FRF pier (NOAA station) and the Currituck Sound.

Table E1. Water level data

Source	#/ Stations	Frequency	Duration	Format	Archive Size	Products/Notes
End of Pier (NOAA)	1	6 min	Continual	ASCII	N/A	Predicted and measured time series
Currituck Sound	1	6 min	Continual	ASCII	N/A	Predicted and measured time series

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* August 2018	**2. REPORT TYPE** Final report	**3. DATES COVERED** *(From - To)*
4. TITLE AND SUBTITLE Oceanographic Observations Dataset: Data Management Plan		**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Kent K. Hathaway, Jeffrey L. Hanson, Kelly Knee, Andrew Bird, and Robert Fratantonio		**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER** H70H6B
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Engineer Research and Development Center, Field Research Facility U.S. Army Engineer Research and Development Center 1261 Duck Road Kitty Hawk, NC 27949		**8. PERFORMING ORGANIZATION REPORT NUMBER** ERDC/CHL TR-18-11
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Headquarters, U.S. Army Corps of Engineers Washington, DC 20314-1000		**10. SPONSOR/MONITOR'S ACRONYM(S)** **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This Data Management Plan details the Oceanographic Observations dataset, which is maintained at the U.S. Army Corps of Engineers (USACE) Engineer Research and Development Center Coastal and Hydraulics Laboratory (ERDC-CHL) Field Research Facility (FRF), Duck, NC. The plan was developed to support the FRF Data Integration Framework (FDIF) project. Information is organized in the following categories: general description, points of contact, data stewardship, data documentation, data sharing, initial data storage and protection, long-term archiving and preservation, hardware and software requirements, products/programs, tools, references, data catalog, and abbreviations and acronyms.

15. SUBJECT TERMS

Coastal engineering
Data Libraries
Data Curation
Duck, NC
Oceanography
Oceanographic
Research Stations

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR	45	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. 239.18