GULF OF MEXICO RESEARCH INITIATIVE
INFORMATION & DATA COOPERATIVE

Cruise Data Submission Guidance

Purpose

As a part of GRIIDC's mission to ensure a data and information legacy, GRIIDC submits data generated from research cruises to the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI) for long-term, 75-year archiving. NCEI requires that we submit data with specific documentation and file structure. This guidance provides information about the types of cruise data to submit to GRIIDC, supporting documentation that is required, and the file structure required for data submission. By submitting data to GRIIDC with the required documentation and file structure you will help facilitate the copying of your data to NCEI. GRIIDC will not accept cruise data without the required documentation or without the required file structure.

Types of Cruise Data to Submit to GRIIDC

The following data should be submitted as separate datasets to GRIIDC for each cruise:

System Type	Description/Examples
ADCP	Acoustic Doppler Current Profiler measuring ocean currents at stations or using transects
CTD	Conductivity, temperature, pressure and other sensors in the water column
Echosounder	Single, dual, or split-beam sonar; may be fixed frequency or chirped; measures depth to seafloor or midwater reflectors, shallow sediment profile, and/or water column image
Expendable Probe	XBT, XCTD, XSV, and/or XCP launched from hand or deck unit
Fluorometer	Fluorescence (usually for phytoplankton)
Gravimeter	Gravity field
Magnetometer	Magnetic field
Underway	• Meteorological data (wind speed/direction, temperature, humidity, turbulence, precipitation, radiation, etc.) • Navigation data (time, position, and motion e.g. Global Navigation Satellite System [D/GPS, GLONASS], Vertical/altitude Reference Unit [IMU/MRU], Speed log, and Gyrocompass • ADCP underway data • Flow-through data (CTD, fluorometers and other sensors, thermosalinograph [TSG])
Multibeam	Hull-mounted multiple beamforming sonar measures swath bathymetry, amplitude, and/or backscatter as well as water column image
pCO2	Partial pressure of dissolved carbon dioxide
SSV	Sea surface sound velocimeter

Winch	Wire tension, speed, payout, etc. recorded for vessel safety

CTD Data

CTD data should be submitted in ASCII or NetCDF format. Common ASCII file extensions include .txt, .cnv, and .csv. In addition to these files, providers should submit all available raw files (e.g., .hex, .bin, .hdr, .bl) together with configuration files (.con or .xmlcon). For more information, please see the GRIIDC CTD Data Submission guidance located on the GRIIDC website under the Help tab – Training and User Guides.

Underway Data

Underway data can be submitted as multiple datasets; however, the navigation files should be included in each dataset. If underway data is submitted as one dataset, please separate data types using folders.

Samples Collected During Cruise

If samples are collected during a cruise for analysis in the lab, for example genetic or chemical analysis, these data should be submitted as a separate dataset.

If samples are collected during a cruise and exposed to manipulated conditions in the lab, these data should be submitted as a separate dataset.

Cruise Data Submission Example:

R/V Pelican departs from St. Petersburg, Florida on May 16, 2018 and returns June 1, 2018. This cruise collected CTD, ADCP, and Meteorological (Met) data while underway. Sediment cores were obtained from 10 sites for analysis of chemical concentration of polycyclic aromatic hydrocarbons in the lab. Plankton tows were completed at five sites and species will be enumerated in the lab under light microscopy. Fish were caught via trawls and were measured in the field, and genetic and contaminant analysis will be performed in the lab. Water samples were taken to measure water quality parameters through lab analysis.

In this example, at least seven datasets would be submitted to GRIIDC:

> Dataset 1: CTD, ADCP, met data, navigation data
> Dataset 2: Sediment core PAH analysis
> Dataset 3: Species identified from plankton tows using light microscopy images
> Dataset 4: Species of fish caught and trawl information
> Dataset 5: Genetic analysis of fish caught
> Dataset 6: Contaminant analysis of fish caught
> Dataset 7: Water quality parameters

Dataset requirements

- Excel files should not contain any color formatting, formulas, graphs, signatures, or designs.
- Please use the GRIIDC Cruise Data Documentation Excel template for data documentation. Additional documentation can be submitted as a readme file (text format). Please do not submit a Word document.
- All headers and acronyms should be defined in the metadata or readme file. Please include definitions of all parameters and units.

- Please use a CF compliance checker for NetCDF files.
- Avoid leaving blank cells in spreadsheets. If blank cells are necessary please define them (e.g., data not collected, below detection limits, etc…) in the descriptive information.
- Datasets for different cruise ID's should be submitted as separate datasets.

Required Documentation

Specific documentation is required for cruise dataset submissions. A data documentation file must be submitted with each cruise dataset. This file should contain descriptive information about the contents of the dataset and include explanations of the folder structure and contents of folders and files. Please use the data documentation template to provide the required information:

Cruise Information Section:

- Start Date of Cruise: Formatted YYYYMMDD
- End Date of Cruise: Formatted YYYYMMDD
- Cruise purpose
- Sea name: Name of the body of water where samples were collected.
- Chief Scientist First Name Last Name

Keyword Section: The next section of the data documentation focuses on keywords that NCEI uses to facilitate search.

- **Observation Type:** Provide at least one observation type from the table below. Please include "profile" for data packages containing CTD profile data.

aerial photography	ice hole sampling	sediment analysis
atlas	imagery	site samples
benthic	in situ	surface measurements
Benthic FOCE-type study	integrated profiles	surface underway
biological	laboratory analysis	survey
chemical	laboratory experiments	survey – biological
composition & location	manual sample collection	survey - coral reef
continuous	marine mammal observation	survey – swimmer/diver
current measurements	meteorological	time series
current measurements – surface currents	model output	time series profile
daily maximum	moored buoy	tows
derived products	navigational	tows – oblique tows
discrete sample	optical	tows – plankton tows
documentation only – no observation type	physical	tows – undulating tows
drifting buoy	phytoplankton abundance	tows – vertical tows
fish examination	plankton	trawl
geological	pore water chemistry	underway
geophysical	profile	undulating profile
GIS product	pump cast	visual estimate
height of breaking surf	satellite data	visual observation

ice core	satellite photography	water chemistry

- **Instrument Type:** Provide at least one instrument type that was used to collect the data.
  Examples: CTD, Thermosalinographs, Multibeam Swath Bathymetry System

- **Datatype:** Provide at least one word or short phrase that describes type of data.
  Examples: Water temperature, conductivity, turbidity, northward sea water velocity, weather observations

- **Platform:** Please include at least one platform name. If you are submitting data collected from a ship, this should be the ship name.
  Examples: Weatherbird II, Pelican

People and Institution Section: Include a list of people that are important to the dataset. This will facilitate search by people and institutions via NCEI.

- **People:** Contributors to the dataset. Please add at least one and include the following information:
  Last name, First name
  Organization/Institution
  Role (Primary point of contact, point of contact, principal investigator, collaborator)
  Email address

- **Collecting Institution:** The institution(s) that collected the data. This is usually the same as the PI's institution.
  Example: Louisiana Universities Marine Consortium, Texas A&M University

- **Author List:** Provide an ordered list of authors who should be included in the citation for the dataset. This may or may not include the people listed above. The format should be: Last name, First name (or first initial); Last name, First name (or first initial). Please separate authors using a semicolon.
  Example: Carson, Rachel; Jones, E.

Folder and File Information Section: A description of each folder and the files in each folder should be provided. Specifically:
- A definition of any folder or file naming conventions used.
- If multiple versions of the same data are included (example: .csv and .xlsx) information about the differences between the files.
- If processed and raw data are included, information about which folders and files contain raw data and which folders and files contain processed data.

Dataset File Structure
The GRIIDC data management system can only accept a single dataset file for a dataset. Dataset files should be packaged into a single compressed archive file (zip). The structure for cruise data files is provided below for a few of the different cruise data types. The documentation folder can contain any

available supporting documentation for the dataset, including but not limited to, cruise reports, station location maps/images, latitude and longitude files, data dictionaries, and navigation logs. The same documentation files may be provided for all datasets from the same cruise. The only required file in this folder is the data documentation file described in the previous section of this document.

Dataset 1: CTD Data

PELICAN_CTD_20170810-20170812.zip
    Folder: Documentation
        Cruise_Data_Documentation_20170811.xlsx*
        DataDocumentationReadme.txt
        CruiseReport.docx
        StationCoordinates.xlsx
    Folder: Data
        Folder: CTD
            CTD20170811.cnv
            CTD20170811.hdr
            CTD20170811.hex
            CTD20170811.ros
            CTD20170811.txt
            CTD20170811.xml
            CTD20170811.xmlcon
            CTD20170811_RawData.csv
            CTD20170811_Downcast.csv
            CTD20170811_Binned_Averages.csv

Dataset 2: Meteorological and Navigation Data

PELICAN_Met_20170810-20170812.zip
    Folder: Documentation
        Cruise_Data_Documentation_20170812.xlsx*
        DataDocumentationReadme.txt
    Folder: Data
        Folder: Met
            WS-1-Barometer-Device-DRV_20170810.Raw
            WS-1-Derived-True-Wind-Average-Speed-10-Seconds-DRV_20170810.Raw
            WS-1-Dervived-True-Wind-DRV_20170810.Raw
        Folder: Navigation
            ASHTECH-$GPGGA-RAW_20180518-123000.Raw
            ASHTECH-$GPHDT-RAW_20180518-123000.Raw
            ASHTECH-$GPVTG-RAW_20180518-123000.Raw

Dataset 3: Longline Fisheries Data

PELICAN_Longline_20170810-20170812.zip
    Folder: Documentation
        Cruise_Data_Documentation_20170811.xlsx*
        DataDocumentationReadme.txt

        Folder: Data
                FishCatch20170811.csv

    *these files are generated from the Cruise_Data_Documentation_Template_2018105.xlsx distributed by GRIIDC.

Additional Information

If you would like to provide more specific keywords for each section of the data documentation file please reference the vocabularies for each keyword listed below. In your data documentation file please add the name of the vocabulary above the keyword list.

Seanames GCMD: http://gcmdservices.gsfc.nasa.gov/static/kms/locations/locations.csv
Seanames NODC: https://www.nodc.noaa.gov/General/NODC-Archive/seanamelist.txt
Instrument Type GCMD: http://gcmdservices.gsfc.nasa.gov/static/kms/instruments/instruments.csv
Science Keywords GCMD:
http://gcmdservices.gsfc.nasa.gov/static/kms/sciencekeywords/sciencekeywords.csv
Climate and Forecast Standard Names Table: http://cfconventions.org/standard-names.html
Data Type GCMD: http://gcmdservices.gsfc.nasa.gov/static/kms/sciencekeywords/sciencekeywords.csv
Platforms - ICES: http://seadatanet.maris2.nl/v_bodc_vocab_v2/search.asp?lib=C17
Platforms - NODC: https://www.nodc.noaa.gov/cgi-bin/OAS/prd/platform
Project Name - NODC: https://www.nodc.noaa.gov/cgi-bin/OAS/prd/project
World Meteorological Organization (WMO) ISO Profile Categories - https://geo-ide.noaa.gov/wiki/index.php?title=World_Meteorological_Organization_ISO_Metadata_Profile