Submitting Marine Geophysical Data
to NOAA National Centers for Environmental Information
& the co-located IHO Data Centre for Digital Bathymetry

Introduction

This document describes procedures to prepare multibeam bathymetric data, subbottom profiler data, water column sonar data, and supplemental data sets for submission to NOAA National Centers for Environmental Information (NCEI) and the co-located International Hydrographic Organization (IHO) Data Centre for Digital Bathymetry (DCDB).

File Formats

1. Multibeam Bathymetry

General Information:

The multibeam bathymetry database at NCEI/IHO DCDB primarily maintains raw (as collected) data files in the instrument's vendor specific format (e.g., .all, .s7k, .xse). However, any other supplemental data (sound speed profiles, tides, vessel offsets, cruise reports, etc.) and/or processed versions of the multibeam data are also accepted. In all submissions, the data files and cruise/survey should be well documented using metadata.

MB Data File Formats:

NCEI can accept bathymetric data from most of the commercial multibeam sonars and acquisition systems. The multibeam bathymetry data management pipeline at NCEI relies heavily on the open source software suite, MB-System. Data formats supported by the software are listed on their website. Data submitted in unsupported formats will still be accepted but will not be discoverable through the web services provided at NCEI (e.g., Bathymetric Data Viewer). These data can only be accessed from the archive upon request to mb.info@noaa.gov.

Processed data (if submitted) need to be delivered in an MB-System processed format or other non-proprietary format. The majority of processed data in the multibeam bathymetry database are processed MB-System, XYZ, or GSF format.

If your data are not in one of the supported formats, email mb.info@noaa.gov to discuss the options available at NCEI for your data.

2. Subbottom Profiler Data

General Information:

The trackline geophysical database at NCEI/IHO DCDB primarily maintains raw (as collected) or processed data files in SEGY format. In all submissions, the data files and cruise/survey should be well documented using metadata.

Subottom Data File Formats:

NCEI encourages data providers to provide data in SEGY format. Most commercial subbottom profilers can either collect in this format or convert from the proprietary format to community-friendly SEGY. The subbottom data management pipeline at NCEI relies heavily on SEGY as a means of extracting navigation necessary to generate tracklines that display the location of the data in the Trackline Geophysical Data Viewer. Data submitted in unsupported formats will still be accepted but will not be discoverable through the web services provided at NCEI. These data can only be accessed from the archive upon request to trackline.info@noaa.gov.

3. Water Column Sonar Data (WCSD)

General Information:

The water column sonar database NCEI/IHO DCDB primarily maintains raw (as collected) data files in the instrument's vendor specific format (e.g., .raw, .wcd). However, any other supplemental data (sound speed profiles, tides, vessel offsets, biological data, cruise reports, etc.) are also accepted. NCEI has developed a data packaging toolpackager_10_01_16 that facilitates submitting water column sonar data and creation of accompanying metadata. To enable discovery and access of the data through the Water Column Sonar Data Viewer, please ensure navigation datagrams are included in the water column data files (e.g., *.wcd).

WCSD Data File Formats:

NCEI can accept files from most instruments capable of collecting water column sonar data. Currently supported instruments include Kongsberg EM series multibeam sonars (EM122, EM2040, EM302, EM3002, EM710), Reson SeaBat 7125, Simrad multibeam sonars (ME70, MS70), and Simrad single beam systems (EK500, EK60, ES60, EK80). It is preferable to have position data within the water column sonar data files. Files lacking position information may not be publicly accessible through NCEI data discovery portals. If your data are not in one of the supported formats, email wcd.info@noaa.gov to discuss the options available at NCEI for your data.

Metadata:

Proper metadata are very important for documenting the history of the data and providing insight into the means of long-term preservation. Please include any metadata that have already been created for each cruise/dataset. NCEI uses and prefers ISO standard metadata, but accepts all other standards. If you are not familiar with metadata, NCEI has developed a metadata primer available here: http://www.ncddc.noaa.gov/metadata-standards/. Example ISO standard metadata records for dataset level (multibeam, singlebeam/subbottom, water column), collection/cruise level, and multibeam file level are provided with this document.

If cruise level metadata have not been created, the minimum requested metadata fields and examples are listed in the following table. An Excel spreadsheet to easily populate this information for the data submission can be requested from mb.info@noaa.gov.

Information Field	Example	Comments
SURVEY_NAME	NF1309	Typically "ship ID, year, cruise number"
SHIP_NAME	Nancy Foster	
CHIEF_SCIENTIST	John Smith	"None" for transits
CHIEF_SCI_ORGANIZATION	USGS	"None" for transits
DEPARTURE_PORT	San Juan, Puerto Rico	City, State for US ports. City, Country for international ports
ARRIVAL_PORT	Charleston, SC	City, State for US ports. City, Country for international ports
START_TIME	30-SEP-13	date only. "DD-MMM-YY"
END_TIME	01-OCT-13	date only. "DD-MMM-YY"
NAV1	DGPS (or GPS)	Equipment used in determining data positioning
INSTRUMENT	Reson 7125	Sonar instrument used in data collection
HORIZONTAL_DATUM	WGS84	If projected data, which projection zone
VERTICAL_DATUM	MLLW	not required for transit
SHIP_OWNER	NOAA	
PROJECT_NAME	Corals in the Florida Keys	Specified project name or "Transit"
SOURCE	NOAA	Source organization of data being provided
ABSTRACT	Text summary	Brief narrative summary of the resource contents. Abstract narrative should include information on general content and features; dataset application: GIS, CAD, image, database; geographic coverage: county/city name; time period of content: begin and end date or single date; and special data characteristics or limitations.
PURPOSE	Text summary	Summary of the intentions for which the dataset was developed. Purpose includes objectives for creating the dataset and what the dataset is to support.
PROPRIETARY	yes/no	Pertains to temporary data access restrictions.
COMMENTS	Proprietary hold until Oct 1, 2014	General comments regarding the cruise or dataset, if any
PROCESSING_STEPS	Text summary	Paragraph describing processing performed on data, if any
DOI*	doi:10.7289/V56T0JNC	If a DOI is not provided, NCEI will create one upon request.
OUTSIDE_LINK	http://www.....	Web link to additional information regarding cruise, project, or funding

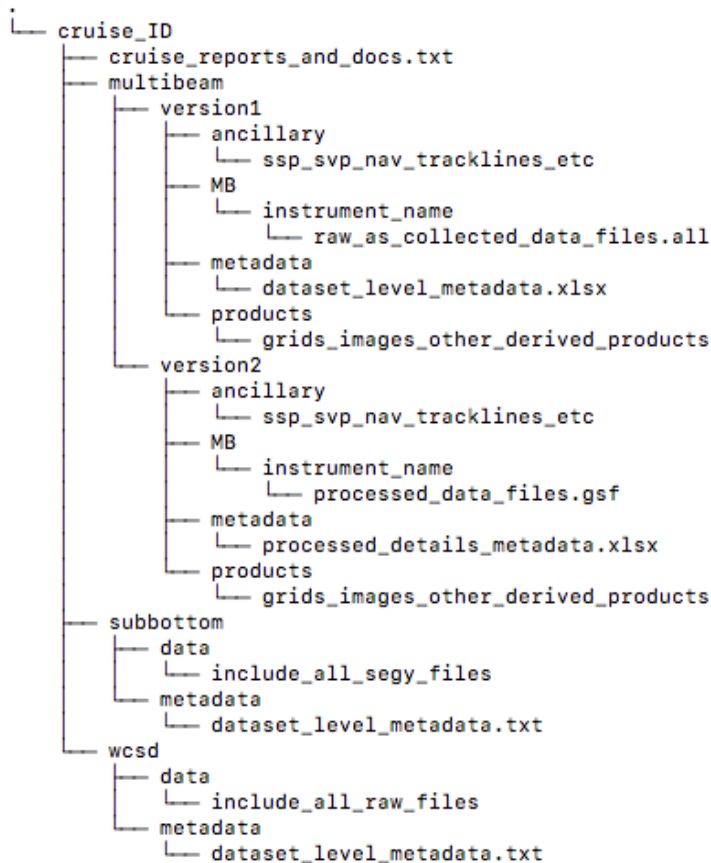File level metadata are also accepted if available, but not necessary.

For documenting processed data, a separate file or xml document may be submitted in addition to the above table. Please include any corrections to soundings that were applied, the nature and means by which the data were cleaned/edited, and any other corrections or conversions applied to the data.

*A DOI (digital object identifier) is a type of persistent identifier used to uniquely identify objects (in this case a particular survey or dataset). Metadata about the object is stored in association with the DOI name. For more information see: https://en.wikipedia.org/wiki/Digital_object_identifier

Data File Structure:

The data may be delivered in one archived file (e.g., tar or zip) or in a well-defined directory structure. Please include an MD5 checksum with the delivery so NCEI can verify the integrity of the files and the completeness of the data transfer. For questions regarding MD5 checksums, contact mb.info@noaa.gov.

A preferred data structure would be the following:

```
.
└── cruise_ID
    ├── cruise_reports_and_docs.txt
    ├── multibeam
    │   ├── version1
    │   │   ├── ancillary
    │   │   │   └── ssp_svp_nav_tracklines_etc
    │   │   ├── MB
    │   │   │   └── instrument_name
    │   │   │       └── raw_as_collected_data_files.all
    │   │   ├── metadata
    │   │   │   └── dataset_level_metadata.xlsx
    │   │   └── products
    │   │       └── grids_images_other_derived_products
    │   └── version2
    │       ├── ancillary
    │       │   └── ssp_svp_nav_tracklines_etc
    │       ├── MB
    │       │   └── instrument_name
    │       │       └── processed_data_files.gsf
    │       ├── metadata
    │       │   └── processed_details_metadata.xlsx
    │       └── products
    │           └── grids_images_other_derived_products
    ├── subbottom
    │   ├── data
    │   │   └── include_all_segy_files
    │   └── metadata
    │       └── dataset_level_metadata.txt
    └── wcsd
        ├── data
        │   └── include_all_raw_files
        └── metadata
            └── dataset_level_metadata.txt
```

CruisePack Data Packager Tool:

NCEI CruisePack is an optional data packager tool that makes packaging cruise-based data for submission to NCEI quick and easy. This free, stand-alone tool can be easily used on Windows computers. It provides a simple user interface to facilitate entering metadata about the cruise and datasets and a way to specify the location of the data so that CruisePack can automatically organize the data into an accepted directory structure. Finally, it packages the data to a specified location (an external hard drive or another system) for transfer to NCEI/DCDB. Using CruisePack can dramatically simplify the process of organizing data and ensure all required metadata are provided. If you are interested in using CruisePack or would like more information, please email mb.info@noaa.gov.

Data Submission:

Email mb.info@noaa.gov to alert the multibeam data manager of incoming data (multibeam, subbottom, wcsd), set up your data submission, or ask any questions.

Data can be delivered to NCEI via (1) shipping external hard drives, (2) uploading to NCEI's FTP server, or (3) data copy using rsync through a secure shell login (linux).

1. External hard drives containing a data submission can be shipped to the following address
   Jennifer Jencks
   NOAA NCEI
   325 Broadway E/NE42
   Boulder, CO 80305

2. NCEI maintains a number of public FTP servers. All the FTP servers allow anonymous ftp access (login: anonymous, password: *email address*). The servers share a single 1.0TB file system. If delivering large volumes of data (> 200GB), please notify mb.info@noaa.gov prior to uploading the data.

   The incoming FTP server (ftp://incoming.ngdc.noaa.gov/) allows external users to upload files. Files should be placed in the /pub/incoming/ directory. After 14 days, the files will be deleted. No external user can read files from this server. Directory listings have been disabled for the incoming FTP site.

   There are multiple ways to access the incoming FTP server.
   - Using a Windows FTP client
     - Two common clients are FileZilla and WinSCP. Both are free and easy to use. Use the following FTP settings to connect
       - File protocol: FTP
       - Host: incoming.ngdc.noaa.gov
       - Username: anonymous
       - Password: <leave blank>
     - Once connected, change the remote site directory to */pub/incoming*
       - You will see notifications stating that the "Directory listing failed" or "550 Permission denied". These are okay as long as you are in */pub/incoming*
     - Create a directory with a name that uniquely identifies the source of your submission (e.g., your name, your institution's name) in */pub/incoming*
       - Open this new directory
     - Copy and paste (or drag and drop) your files to this directory

   - Using a linux FTP client
     - Navigate to the directory containing your data submission
     - Type *ftp incoming.ngdc.noaa.gov*
     - Navigate to */pub/incoming*
     - Create a directory with a name that uniquely identifies the source of your submission (e.g., your name, your institution's name)
     - cd into the directory that you just created
     - Upload the data using the put or mput commands

When using FTP upload to submit your data, always email mb.info@noaa.gov and let the data manager know the number of files uploaded and the total volume. This will allow the data manager to verify the data were successfully loaded to the FTP server and downloaded at NCEI.

3. Providing data through rsync via a secure shell login is reserved for recurring submissions from trusted providers. Email mb.info@noaa.gov if you feel that this method would be best for your data. The data manager and the IT security admin at NCEI will assess your request and determine if this method is allowable.