Please provide the following information, and submit to the NOAA DM Plan Repository.

Reference to Master DM Plan (if applicable)

As stated in Section IV, Requirement 1.3, DM Plans may be hierarchical. If this DM Plan inherits provisions from a higher-level DM Plan already submitted to the Repository, then this more-specific Plan only needs to provide information that differs from what was provided in the Master DM Plan.

URL of higher-level DM Plan (if any) as submitted to DM Plan Repository:

1. General Description of Data to be Managed

1.1. Name of the Data, data collection Project, or data-producing Program:
Bathymetry (Alaska and surrounding waters)

1.2. Summary description of the data:
A Raster having 20 m resolution with decimal values was assembled from 18.6 billion bathymetric soundings that were obtained from the National Center for Environmental Information (NCEI) https://www.ncei.noaa.gov. Bathymetric soundings extends from Kuril-Kamchatka Trench in the Bearing Sea along the Aleutian Trench to the Gulf of Alaska, and in the Arctic Ocean from Prince Patrick Island to the International Date Line. Bathymetric soundings were scrutinized for accuracy using statistical analysis and visual inspection with some imputation. Editing processes included: deleting erroneous and superseded values, digitizing missing values, and referencing all data sets to a common, modern datum.

1.3. Is this a one-time data collection, or an ongoing series of measurements?
Ongoing series of measurements

1.4. Actual or planned temporal coverage of the data:
2013 to Present

1.5. Actual or planned geographic coverage of the data:
W: -133, E: 170, N: 88, S: 40
Alaska and surrounding waters

1.6. Type(s) of data:
(e.g., digital numeric data, imagery, photographs, video, audio, database, tabular data, etc.)
Map (digital)

1.7. Data collection method(s):
(e.g., satellite, airplane, unmanned aerial system, radar, weather station, moored buoy, research vessel, autonomous underwater vehicle, animal tagging, manual surveys, enforcement activities, numerical model, etc.)

1.8. If data are from a NOAA Observing System of Record, indicate name of system:

1.8.1. If data are from another observing system, please specify:

2. Point of Contact for this Data Management Plan (author or maintainer)

2.1. Name:
Steve Lewis

2.2. Title:
Metadata Contact

2.3. Affiliation or facility:

2.4. E-mail address:
steve.lewis@noaa.gov

2.5. Phone number:
907-586-7858

3. Responsible Party for Data Management
Program Managers, or their designee, shall be responsible for assuring the proper management of the data produced by their Program. Please indicate the responsible party below.

3.1. Name:
Steve Lewis

3.2. Title:
Data Steward

4. Resources
Programs must identify resources within their own budget for managing the data they produce.

4.1. Have resources for management of these data been identified?
Yes

4.2. Approximate percentage of the budget for these data devoted to data management (specify percentage or "unknown"):
Unknown

5. Data Lineage and Quality
NOAA has issued Information Quality Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information which it disseminates.

5.1. Processing workflow of the data from collection or acquisition to making it publicly accessible
(describe or provide URL of description):
Lineage Statement:

-Currently, our process keeps a record of the survey from which that point originated. Some older data does not have this level of metadata. Servers are crawled for relevan t data, and a list of download URLS to data files is returned. Data files are then retrieve d using custom python scripts. Raw data is downloaded from online NCEI web servers. Data is converted to CSV or XYZ files Points missing one or more of their XYZ points val ues are removed and archived. Data is evaluated as a component of the bathymet ry map to identify outliers (instances where data point(s) are not consistent with the expected variability of the surrounding environments) using a variety of statistical a nd manual methods.   K-Natural Neighbors (KNN) - Python and SciPy Percentil es with Standard Deviations - Python and SciPy Slope and Neighbors - ArcGIS Models Manual/visual selection Human   Imputation of points to make them consistent with sur rounding terrain tracklines and satellite altimetry.   Upon integration into the dataset, be tween 0-25% of the deepest and shallowest points are immediately removed from each tr ackline based on the StdDev of each induvial trackline. All data is archived; a dataset with the outliers could be built within 7 weeks. The percentage of points removed (R) is d etermined by a non-linear function of the datasets standard deviation (s), and can be seen below: $R = -4.746813 + 30.059/(1+(s/9.584625)^{0.9983})$) This function wa s derived using a best-fit curve tool, which was instructed to return a naturally-logar ithmic function which was equal to ~25 when s=0, and which decreased asymptotically t o 0 as s grew larger. The function was then tailored to have what the developer felt wa s a reasonable slope   The logic here is that datasets with low standard deviations would be relatively flat and featureless. Since they have a lower level of topographic com plexity, they can undergo a higher rate of removal while still retaining the essential topographic character of the surface they represent.   Data is then organized in to a Kd-Tree structure in which data points are organized based on their values with th e data sorted between levels in the tree (i.e., the first level is split along the x axis, the next level is split along the y axis, the next along the z axis, and then the fourth along th e x axis again. The result was a tree which can be searched in O (log(n)) time, and which was optimized for quick spatial searches, critical for the next step.   K-Nearest-Nei ghbors (KNN) statistical model is used on the data. This uses the value of each data point's K spatially nearest neighbors (k-value) to produce an 'expected value.' The expected value is then subtracted from the point's observed value, and the absolute value of this difference is the point's 'residual.' After calculating the residuals for all of the poin ts in the data set, we remove the 5% of points with the highest residuals.   After these s teps are completed, the remaining data are converted to Feature Classes. This data str ucture is composed of not only the raw data, but also a host of metadata calculated from t he raw data, such as vessel name and tracklines number. Spatial indexes are added to the data to optimize operations. All internal data is point data and is stored in Alaska Alb ers project.   ArcMap and Arc Pro Slope and Neighbor Outlier Tools. The ArcMap and ArcPro function analyzes each data point based on the slope of the rendered terrain pol ygon and the point's immediate adjacent neighbors. If a sufficient portion of these slopes exceeded a manually pre-defined threshold, the point is flagged as a potential outlier but not removed. After this function has identified all potential outliers, the set is visua lly reviewed and flagged. Flagged points are manually removed from the terrain but st

ored as an independent shapefile and thus no data removed from the active dataset were truly deleted.

5.1.1. If data at different stages of the workflow, or products derived from these data, are subject to a separate data management plan, provide reference to other plan:

5.2. Quality control procedures employed (describe or provide URL of description):
Used K-natural neighbors, Percentiles, and ArcGIS slope tools to location and remove outliers.

6. Data Documentation

The EDMC Data Documentation Procedural Directive requires that NOAA data be well documented, specifies the use of ISO 19115 and related standards for documentation of new data, and provides links to resources and tools for metadata creation and validation.

6.1. Does metadata comply with EDMC Data Documentation directive?
No

6.1.1. If metadata are non-existent or non-compliant, please explain:
Missing/invalid information:
- 1.7. Data collection method(s)
- 7.2. Name of organization of facility providing data access

6.2. Name of organization or facility providing metadata hosting:
NMFS Office of Science and Technology

6.2.1. If service is needed for metadata hosting, please indicate:

6.3. URL of metadata folder or data catalog, if known:
https://www.fisheries.noaa.gov/inport/item/27377

6.4. Process for producing and maintaining metadata
(describe or provide URL of description):
Metadata produced and maintained in accordance with the NOAA Data Documentation Procedural Directive:
https://nosc.noaa.gov/EDMC/DAARWG/docs/EDMC_PD-Data_Documentation_v1.pdf

7. Data Access

NAO 212-15 states that access to environmental data may only be restricted when distribution is explicitly limited by law, regulation, policy (such as those applicable to personally identifiable information or protected critical infrastructure information or proprietary trade information) or by security requirements. The EDMC Data Access Procedural Directive contains specific guidance, recommends the use of open-standard, interoperable, non-proprietary web services, provides information about resources and tools to enable data access, and includes a Waiver to be submitted to justify any approach other than full, unrestricted public access.

7.1. Do these data comply with the Data Access directive?
Yes

7.1.1. If the data are not to be made available to the public at all, or with limitations, has a Waiver (Appendix A of Data Access directive) been filed?

7.1.2. If there are limitations to public data access, describe how data are protected from unauthorized access or disclosure:

7.2. Name of organization of facility providing data access:

7.2.1. If data hosting service is needed, please indicate:
No

7.2.2. URL of data access service, if known:
https://alaskafisheries.noaa.gov/arcgis/rest/services
https://alaskafisheries.noaa.gov/arcgis/rest/services/bathy_40m/MapServer
https://alaskafisheries.noaa.gov/arcgis/rest/services/bathy_40m/MapServer

7.3. Data access methods or services offered:
By Raster format download

7.4. Approximate delay between data collection and dissemination:
1 year

7.4.1. If delay is longer than latency of automated processing, indicate under what authority data access is delayed:
NA

8. Data Preservation and Protection
The NOAA Procedure for Scientific Records Appraisal and Archive Approval describes how to identify, appraise and decide what scientific records are to be preserved in a NOAA archive.

8.1. Actual or planned long-term data archive location:
(Specify NCEI-MD, NCEI-CO, NCEI-NC, NCEI-MS, World Data Center (WDC) facility, Other, To Be Determined, Unable to Archive, or No Archiving Intended)
OTHER

8.1.1. If World Data Center or Other, specify:

8.1.2. If To Be Determined, Unable to Archive or No Archiving Intended, explain:

8.2. Data storage facility prior to being sent to an archive facility (if any):
Alaska Regional Office - Juneau, AK

call or email or visit web

8.3. Approximate delay between data collection and submission to an archive facility:
NA

8.4. How will the data be protected from accidental or malicious modification or deletion prior to receipt by the archive?
Discuss data back-up, disaster recovery/contingency planning, and off-site data storage relevant to the data collection
NA

9. Additional Line Office or Staff Office Questions
Line and Staff Offices may extend this template by inserting additional questions in this section.