

Data Report: Comprehensive Inventory of Habitat Assessment and Evaluation Datasets to Support *Deepwater Horizon* Mesophotic and Deep Benthic Communities

Habitat Assessment and Evaluation Project

Authors

Rachel Bassett¹, Jennifer Herting², Janessy Frometa¹, Stephanie Sharuga³, Jacob Howell¹, Laughlin Siceloff¹, Jill Bourque⁴, Megan Cromwell⁵, Kirstie Francis⁵, Randy Clark⁶, Amanda Demopoulos⁴, Andy David⁷, Kristopher Benson⁸, and Stacey L. Harter⁷

¹ CSS, Inc., under contract to NOAA, National Ocean Service, National Centers for Coastal Ocean Science

² Tech Global, Inc., under contract to NOAA, National Marine Fisheries Service, Southeast Fisheries Science Center

³ Bureau of Ocean Energy Management, Office of Environmental Programs

⁴ U.S. Geological Survey, Wetland and Aquatic Research Center

⁵ NOAA, National Environmental Satellite, Data, and Information Service, National Centers for Environmental Information, National Centers for Environmental Information

⁶ NOAA, National Ocean Service, National Centers for Coastal Ocean Science

⁷ NOAA, National Marine Fisheries Service, Southeast Fisheries Science Center

⁸ NOAA, National Marine Fisheries Service, Office of Habitat Conservation

February 2023

DWH MDBC Data Report 2023-01



Suggested Citation

Bassett, R., Herting, J., Frometa, J., Sharuga, S., Howell, J., Siceloff, L., Bourque, J., Cromwell, M., Francis, K., Clark, R., Demopoulos, A., David, A., Benson, K., and Harter, S. L. (2023). Comprehensive inventory of Habitat Assessment and Evaluation datasets to support Deepwater Horizon Mesophotic and Deep Benthic Communities. National Oceanic and Atmospheric Administration. DWH MDBC Data Report 2023-01. <https://doi.org/10.25923/kz7t-4674>

Acknowledgments

Multiple individuals helped facilitate inventory compilation. Stephen Formel, Megan Cromwell, Errol Ronje and Michael Coyne contributed to data management. Lisa Clarke at NOAA Central Library was invaluable in her assistance with the Literature Search. Thanks to Ken Buja for his help with our GIS map images. Arliss Winship, Avery Paxton and Lauren Jackson served as technical advisors. Multiple other individuals across NOAA and other institutions provided data, technical expertise, and participated in the HAE and HAE/Habitat for Fish Workshops. We thank the MDBC Technical Team, Steering Committee, and Leadership Team for their guidance throughout the inventory process. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the National Oceanic and Atmospheric Administration or the Bureau of Ocean Energy Management. The report was reviewed by Mesophotic and Deep Benthic Communities Technical Team, Steering Committee and Open Ocean Trustee Implementation Group. CSS, Inc. employees were supported under NOAA Contract No. EA133C17BA0062.

For more information on MDBC Restoration, please visit:

<https://coastalscience.noaa.gov/project/scientific-support-for-mesophotic-and-deep-benthic-community-restoration-in-the-gulf-of-mexico/>

and

<https://www.fisheries.noaa.gov/southeast/habitat-conservation/mesophotic-and-deep-benthic-communities-restoration>

Disclaimer

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Deepwater Horizon Mesophotic and Deep Benthic Communities Restoration

This report is part of the NOAA Mesophotic and Deep Benthic Communities (MDBC) Series of publications that share the results of work conducted by the *Deepwater Horizon* MDBC restoration projects.

The 2010 *Deepwater Horizon* oil spill was an unprecedented event. Approximately 3.2 million barrels of oil were released into the deep ocean over nearly three months. The plume of oil moved throughout the water column, formed surface slicks that cumulatively covered an area the size of Virginia, and washed oil onto at least 1,300 miles of shoreline habitats. More than 770 square miles (2,000 square kilometers) of deep benthic habitat were injured by the oil spill, including areas surrounding the *Deepwater Horizon* wellhead and parts of the Pinnacles Trend mesophotic reef complex, located at the edge of the continental shelf.

Under the Oil Pollution Act, state and federal natural resource trustees conducted a Natural Resource Damage Assessment (NRDA). The Trustees assessed damages, quantifying the unprecedented injuries to natural resources and lost services. They also developed a programmatic restoration plan to restore injured resources and compensate the public for lost services.

In April 2016, a settlement was finalized that included up to \$8.8 billion in funding for the *Deepwater Horizon* Trustees to restore the natural resource injuries caused by the oil spill as described in their programmatic restoration plan, Final Programmatic Damage Assessment and Restoration Plan and Final Programmatic Environmental Impact Statement. The *Deepwater Horizon* Open Ocean Trustee Implementation Group is responsible for restoring natural resources and their services within the Open Ocean Restoration Area that were injured by the oil spill. The Open Ocean Trustees include NOAA, U.S. Department of the Interior, U.S. Environmental Protection Agency, and U.S. Department of Agriculture.

In 2019, the Open Ocean Trustee Implementation Group committed more than \$126 million to implement four restoration projects to address the injury to MDBC. The MDBC projects are: Mapping, Ground-Truthing, and Predictive Habitat Modeling; Habitat Assessment and Evaluation; Coral Propagation Technique Development; and Active Management and Protection. NOAA and the Department of the Interior are implementing the projects, in cooperation with a range of partners, over eight years.

Together, the projects take a phased approach to meet the challenges involved in restoring deep-sea habitats. Challenges to restoration include a limited scientific understanding of these communities, limited experience with restoration at the depths at which these communities occur, and remote locations that limit accessibility.

More information about *Deepwater Horizon* restoration and the MDBC restoration projects is available at: www.gulfspillrestoration.noaa.gov.

Table of Contents

1. Introduction.....	1
1.1. <i>Deepwater Horizon</i> Oil Spill	1
1.2. <i>Deepwater Horizon</i> Restoration Plan.....	1
1.2.1. Mesophotic and Deep Benthic Communities (MDBC) Restoration.....	1
2. Goals and Objectives.....	3
3. Inventory Structure	4
3.1. Broad Inclusion Criteria	4
3.1.1. Taxa	4
3.1.2. Geographic Scope	4
3.2. Fine Inclusion Criteria.....	6
3.2.1. Data Characterization Attributes	6
3.2.2. Data Categories.....	7
3.2.3. Data Types.....	8
4. Methods	9
4.1. Data Sources (Inventory of Inventories).....	9
4.1.1. Literature Search	9
4.1.2. Portals and Databases.....	10
4.1.3. Organizations	10
4.1.4. Partner Engagement.....	11
4.1.5. External Engagement for Purpose of Discovering Datasets	11
4.2. Data Screening.....	11
4.2.1. Initial Screening	11
4.2.2. Subsequent Screening.....	12
4.2.3. Tracking the Screening Process.....	12
4.3. Downloading from Portals	12
4.3.1. Material from Data Protocols (Workflow)	12
4.3.2. Data Catalogs per working group	13
4.4. Inventory Products.....	14
4.4.1. Literature Database (EndNote™ Library).....	14
4.4.2. Topical Summaries.....	15
4.4.3. MDBC Geoplatform Inventory	15
4.5. Updating the Inventory.....	16
4.5.1. Data Accretion Plan	16
4.5.2. Periodicity of Inventory Assembly	17
5. Archiving.....	18

6. Planned Uses of Data Inventory	19
6.1. Workshop	19
6.2. Gap Analysis	19
6.3. Prioritization of Topical and Spatial Data Gaps.....	19
6.4. Operations Planning.....	20
7. Results.....	21
7.1. Fish Ecology.....	22
7.1.1. Data Types.....	22
7.1.2. Data Sampling by Instrument.....	22
7.1.3. Spatial Coverage of Data	23
7.1.4. Data Sources	26
7.1.5. Data Collection per Decade.....	27
7.1.6. Data Collection by Depth	28
7.2. Benthic Invertebrates.....	29
7.2.1. Data Types.....	29
7.2.2. Data Sampling by Instrument.....	29
7.2.3. Spatial Coverage of Data	30
7.2.4. Data Sources	32
7.2.5. Data Collection per Decade.....	33
7.2.6. Data Collection by Depth	34
7.3. Corals.....	35
7.3.1. Data Types.....	35
7.3.2. Data Sampling by Instrument.....	35
7.3.3. Spatial Coverage of Data	36
7.3.4. Data Sources	37
7.3.5. Data Collection per Decade.....	38
7.3.6. Data Collection by Depth	39
7.4. Environmental and Oceanographic	40
7.4.1. Data Types.....	40
7.4.2. Data Sampling by Instrument.....	40
7.4.3. Spatial Coverage of Data	41
7.4.4. Data Sources	41
7.4.5. Data Collection per Decade.....	42
7.5. Genetics	43
8. Conclusions	47
Literature Cited.....	48

Appendices.....	50
Appendix A: Data Characterization Attributes.....	50
Appendix B: Data Sources.....	55
Appendix C: Literature Review Framework.....	57

List of Figures

Figure 1. Tier 1 of the geographic scope for the benthic invertebrate and fish ecology working groups inventory	5
Figure 2. Tier 2 of the geographic scope for the benthic invertebrate and fish ecology working group inventory.....	5
Figure 3. Workflow diagram for the identification, acquisition and prioritization of datasets.....	13
Figure 4. A pie chart displaying the percentages of data types in the datasets collected for the fish ecology (FIS) data inventory	22
Figure 5. A bar graph displaying the number of fish ecology (FIS) datasets that used the listed sampling instruments, fish collection techniques and other associated techniques such as water quality and current measurements	23
Figure 6. Submersible vehicle and towed camera system tracklines used during various surveys collected for the fish ecology data inventory.....	24
Figure 7. Deployment sites of various underwater camera gears, such as submersible vehicles and submerged baited underwater video arrays used during various surveys collected for the fish ecology data inventory.....	24
Figure 8. Deployment sites of traps by the Panama City NOAA NMFS Southeast Fisheries Science Center to identify and collect fish(es), supplementary to their bottom video survey.....	25
Figure 9. Deployment sites of various hook and line gears, such as bottom longline and vertical longlines, used during various surveys collected for the fish ecology data inventory.....	25
Figure 10. Trawl survey tracklines and deployment sites in the GOM used during various surveys collected for the fish ecology data inventory.....	26
Figure 11. Survey sites per deployment of various gears, such as trawls, drop cameras, and hook and line gears used during various surveys collected for the fish ecology data inventory.....	26
Figure 12. The total number of datasets provided by various sources for the fish ecology data inventory.....	27
Figure 13. A plot of the total number of datasets surveyed during various decadal periods of time	28
Figure 14. A plot of the total number of datasets compared to the depths in which they were surveyed.....	28
Figure 15. A pie chart displaying the percentages of data types in the datasets collected for the Benthic Invertebrates (BIN) group.....	29
Figure 16. A bar graph with the number of benthic invertebrate (BIN) datasets that used the listed sampling instruments, sample collection techniques, and other associated techniques such as water quality and current measurements	30
Figure 17. ROV and HOV dive sites within the datasets collected for the benthic invertebrate data inventory.....	31
Figure 18. Core sample sites within the datasets collected for the benthic invertebrate data inventory.....	31
Figure 19. Trawl and MOCNESS sampling sites within the datasets collected for the benthic invertebrate data inventory.....	32

Figure 20. Non-ROV underwater camera sampling sites within the datasets collected for the benthic invertebrate data inventory.....	32
Figure 21. A plot of the total number of datasets provided by various sources for the benthic invertebrate data inventory.....	33
Figure 22. A plot of the total number of benthic invertebrate datasets surveyed during various decadal periods of time.....	34
Figure 23. A plot of the total number of datasets compared to the depths in which they were surveyed.....	34
Figure 24. A pie chart displaying the percentages of data types in the datasets collected for the Coral (COR) group.....	35
Figure 25. A bar graph displaying the number of coral (COR) datasets that used the listed sampling instruments and coral collection techniques, and other associated techniques such as water quality and current measurements.....	36
Figure 26. ROV and HOV dive sites within the datasets collected for the coral data inventory.....	37
Figure 27. Non-ROV underwater camera sampling sites within the datasets collected for the coral data inventory.....	37
Figure 28. A plot of the total number of datasets provided by various sources for the coral data inventory.....	38
Figure 29. A plot of the total number of datasets surveyed during various decadal periods of time. Some surveys have been conducting work over several decades and are represented under the Multi-Decade section.	39
Figure 30. A plot of the total number of datasets compared to the depths in which they were surveyed.....	39
Figure 31. A pie chart displaying the percentages of data types in the datasets collected for the Environmental/Oceanographic (ENV) group.....	40
Figure 32. A bar graph displaying the number of environmental/oceanographic datasets that used the listed sampling instruments, sediment collection techniques, and other associated techniques such as water quality and current measurements.....	41
Figure 33. A plot of the total number of datasets provided by various sources for the environmental/oceanographic data inventory.	42
Figure 34. A plot of the total number of datasets surveyed during various decadal periods of time. Some surveys have been conducting work over several decades and are represented under the Multi-Decade category.	43
Figure 35. Total number of deep-sea coral species with genetic data available (solid fill) and no data available (pattern fill), organized by scientific taxonomic group.	46

List of Tables

Table 1. The working groups selected the broad inclusion criteria that were pertinent to their inventory searches.....	4
--	---

Table 2. The working groups selected the minimum metadata requirements to be met with an “X” or not required with a null or “NA”	7
Table 3. Working groups focused their data search and discovery on those data categories which have an “X”	7
Table 4. The working groups identified data types to be considered for their inventory searches	8
Table 5. The working groups summarized the knowledge gathered in the literature review by writing topical summaries based on the literature review framework	15
Table 6. Overview of the datasets acquired by each HAE working group	21
Table 7. The number of layers and datasets that were represented spatially in the fish ecology data inventory by data type.	23
Table 8. The total count as well as percentages were calculated per each source of surveys inventoried for fish ecology datasets	27
Table 9. The number of layers and datasets that were represented spatially in the benthic invertebrate data inventory by data type.	30
Table 10. The total count as well as percentages were calculated for each source of datasets inventoried for benthic invertebrate datasets.	33
Table 11. The number of layers and datasets that were represented spatially in the coral data inventory by data type.	36
Table 12. The total count as well as percentages were calculated for each source of datasets inventoried for coral datasets	38
Table 13. The total count as well as percentages were calculated for each source of datasets inventoried for environmental/oceanographic dataset	42
Table 14. Mesophotic and deep-sea coral taxa with no existing publicly available genetic data. Taxa are organized by taxonomic group: octocorals, stony corals, black corals, and sea pens.....	45
Table A1. Cross-project Attribute Table	50

Acronyms

ADCP - Acoustic Doppler Current Profiler
AUV - Autonomous Underwater Vehicle
BOEM – Bureau of Ocean Energy Management
BP – BP Exploration and Production Inc.
CPT - Coral Propagation Techniques
CSS, Inc. – Consolidated Safety Services, Incorporated
CTD - Conductivity, Temperature, Depth (instrument)
DOI - U.S. Department of the Interior
DWH - Deepwater Horizon
EPA - U.S. Environmental Protection Agency
FWCI - Fishes and Water Column Invertebrates
GIS – Geographic Information System
GOM - Gulf of Mexico
HAE - Habitat Assessment and Evaluation
HOV - Human Operated Vehicle
MDBC - Mesophotic and Deep Benthic Communities
MOCNESS - Multiple Opening/Closing Net and Environmental Sensing System
NCBI - National Center for Biotechnology Information
NCEI – National Centers for Environmental Information
NCCOS – National Centers for Coastal Ocean Science
NMFS – National Marine Fisheries Service
NOAA - National Oceanic and Atmospheric Administration
OOTIG - Open Ocean Trustee Implementation Group
PDARP/PEIS - Final Deepwater Horizon Programmatic Damage Assessment and Restoration
Plan/Programmatic Environmental Impact Statement
RAD-seq – Restriction Site-Associated DNA Sequencing
ROV - Remotely Operated Vehicle
SNP - Single Nucleotide Polymorphism
USDA - U.S. Department of Agriculture
USGS – United States Geological Survey

Definitions

Several terms are referenced repeatedly throughout the Inventory Report and are defined as follows:

Data extraction:

The process of retrieving data and associated metadata from datasets that passed the data screening process based on their adherence to defined inclusion criteria.

Data Inventory and Analysis Implementation Plan:

A systematic and reproducible plan for compiling the comprehensive inventory. Also referred to as “Inventory Plan.”

Data screening:

The process of reviewing datasets for inclusion in the inventory based on whether the data meet defined inclusion criteria. Data screening occurred in multiple phases where data were reviewed for adherence to broad-level inclusion criteria and then reviewed again at a narrower level for adherence to finer-scale inclusion criteria. Datasets that passed the screening steps are those from which data was extracted and/or compiled.

Inclusion criteria:

Objective standards or criteria used to evaluate datasets for inclusion in the inventory. Inclusion criteria can be broad or fine-level and set the coarse and fine filters for data compilation, respectively.

Inventory:

The collection of raw data or derived products, tabular, or spatial data. The inventory includes data from different sources, such as direct observations and photo libraries of MDBC taxa from deep submergence vehicles (ROVs, HOVs, AUVs), scuba or rebreather diving, fisheries independent monitoring (e.g., trawls, sleds, traps, or Multiple Opening/Closing Net and Environmental Sensing System (MOCNESS)), fisheries dependent data, isotopic or other physiologically related studies, drop cameras or lander cameras, passive and active hydroacoustics, organismal tagging observations, discrete collections of specimens, genetic sequences, sediment community data, and environmental data. The inventory may also include literature citations and other documentation of existing MDBC habitat characterization data sets (e.g., sample collections).

Mesophotic benthic habitats:

Areas on the seafloor that exist in the depth zone where light levels are low and are inhabited by an interacting group of benthic species. Mesophotic habitats are typically found at depths ranging from approximately 100 feet (30 m) and extending to over 500 feet (approximately 150 m) in tropical and subtropical regions (Puglise et al. 2009, NOAA 2011, Hourigan et al. 2015, Sulak and Dixon 2015, DWH NRDA Trustees 2016). For the purpose of these MDBC projects, work in mesophotic habitats will generally focus on biological communities existing between approximately 165 feet (50 meters) and 980 feet (300 m) deep. The use of 980 feet (300 m) as the lower limit of this zone is consistent with BOEM and BSEE’s current management practices for benthic resources (BOEM 2021). This depth range has also been referred to as the “twilight zone” (Pyle 1996, Kahng et al. 2010). The dominant

communities providing structural habitat among mesophotic benthic habitats comprise coral, sponge and algal species (NOAA 2011a).

Deep or deepwater benthic habitats:

Areas on the seafloor that exist in the depth zone where light is not present and are inhabited by an interacting group of benthic species. For the purpose of this MDBC project, work in deep benthic habitats will generally focus on depths deeper than 980 feet (300 m) deep.

Mesophotic and Deep Benthic Communities (MDBC):

MDBC are an interacting group of benthic species, including fish, mobile, and sessile invertebrates and benthic infauna, that grow within mesophotic and deepwater benthic habitats. Some taxa, such as mesophotic corals, create biogenic habitats and are both a taxonomic unit of the community and define the habitat. In the context of the DWH Programmatic Damage Assessment and Restoration Plan, MDBC is defined as a restoration type.

MDBC habitats:

Areas on the seafloor inhabited and/or associated with mesophotic and deep benthic communities. MDBC habitats include mesophotic reefs, deep-sea corals, soft-bottom habitats, and non-natural substrates (e.g., oil and gas structures, artificial reefs) located on the seafloor. MDBC habitats may be characterized by geomorphological forms, substrates, or biological cover depending on data availability.

1. Introduction

1.1. *Deepwater Horizon* Oil Spill

On April 20, 2010, the *Deepwater Horizon* (DWH) mobile drilling unit exploded, resulting in a massive release of oil from the BP Exploration and Production Inc. (BP) Macondo well, causing loss of life and extensive natural resource injuries in the Gulf of Mexico (GOM). Oil spread from the deep ocean to the surface and nearshore environment from Texas to Florida. Extensive response actions were undertaken to try to reduce harm to people and the environment. However, many of these response actions had collateral impacts on the environment and on natural resource services. The DWH Trustees documented a footprint of over 770 square miles (2,000 square km) of injury to deep-sea benthic habitat surrounding the wellhead and extending up the continental slope (depths greater than 1000 ft [greater than 300 m]), within zones of varying impact (DWH NRDA Trustees, 2016).

Mesophotic and deep benthic communities (MDBC) are largely composed of corals, sponges, infauna, fish and mobile invertebrates, and are distributed across hard and soft bottom habitats deeper than 50 m. MDBC are vast and complex ecosystems on the ocean floor that are a foundation of GOM food webs, as well as provide food, refuge and reproductive opportunities for multiple species of fish and invertebrates, critical for successful fisheries. Additionally, a multitude of organisms from all levels of the food chain, from bacteria to large predatory fishes, inhabit the GOM water column for portions or the entirety of their life cycle.

As described in the Final *Deepwater Horizon* Programmatic Damage Assessment and Restoration Plan/Programmatic Environmental Impact Statement (DWH PDARP/PEIS), the full extent of injuries to MDBC as well as to fishes and water-column invertebrates (FWCI) from the DWH oil spill remains unknown. Attempts were made to quantify injuries to fishes and invertebrates inhabiting the water column, but these efforts were hampered by availability of information to identify pre-DWH baseline conditions and a limited understanding of compounded injuries associated with loss of mesophotic and deep-water habitats. Only small fractions of the mesophotic and deep-water habitats in this region of the GOM have been previously surveyed, and the current distributions of MDBC in the impacted area has not been determined. These knowledge gaps regarding MDBC distribution and connectedness to FWCI create a significant challenge for restoring both MDBC and FWCI communities and achieving restoration goals prescribed by the DWH PDARP/PEIS. The DWH PDARP/PEIS identified MDBC and FWCI as two Restoration Types to help guide restoration planning, project selection, and evaluation of the impacts of restoration activities at both project-specific and cross-project and cross-resource, ecosystem scales. (DWH NRDA Trustees, 2016).

1.2. *Deepwater Horizon* Restoration Plan

1.2.1. Mesophotic and Deep Benthic Communities (MDBC) Restoration

The Habitat Assessment and Evaluation (HAE) project is one of four projects selected by the Open Ocean Trustee Implementation Group (OOTIG) in the Open Ocean Trustee Implementation Group Final Restoration Plan 2/ Environmental Assessment: Fish, Sea Turtles, Marine Mammals, and Mesophotic and Deep Benthic Communities to restore natural resources injured by the 2010 DWH oil spill in the Gulf of Mexico in mesophotic and deep-water habitats (OOTIG, 2019).

The PDARP/PEIS identifies goals for each Restoration Type. These Restoration Type-specific goals help to guide restoration planning and project selection. In addition, the PDARP/PEIS identifies restoration approaches that describe options for implementation and in some cases, techniques and methods. In accordance with the ecosystem approach to DWH restoration, the PDARP/PEIS recognized that the portfolio of approaches to restore for FWCI would include a three-fold approach. Two of these approaches recognize increases in fish and invertebrate productivity that could be achieved via habitat restoration in 1) coastal and nearshore habitats, and 2) offshore habitats. These offshore habitats specifically refer to restoration activities undertaken by the MDBC Restoration Type. Additionally, fish and invertebrate restoration would be achieved by reductions in FWCI mortality by addressing known sources of mortality to fish and invertebrates, such as bycatch reduction and fisheries interactions.

Data inventory and analysis is an important foundational component of the HAE project. The development of this inventory has identified those data types and parameters needed to characterize MDBC in the GOM. The results of the inventory will help guide and direct the subsequent activities for the HAE project. This will allow the team to assess informational and taxonomic data gaps useful for prioritizing future MDBC HAE activities.

This project is interdependent with several other MDBC restoration projects and cross-project teams. As such, many of the datasets, criteria, evaluations, findings, and products will be relevant and available for use by other MDBC projects.

The Inventory and Analysis Team consisted of Rachel Bassett, Janessy Frometa, Jennifer Herting, Stacey Harter, Stephanie Sharuga, Kirstie Francis, Jacob Howell, Lisa Clarke, Jennifer Le, and Laughlin Siceloff.

2. Goals and Objectives

The inventory described herein includes information on MDBC species taxonomy and genetic diversity, distribution, abundance, behavior, trophic relationships, and health, as well as their biological and ecological traits. It also includes environmental and oceanographic data, information on potential contaminants in the marine environment, geological features, anthropogenic influences, and other stressors. Existing, relevant models (such as coral dispersal and population genetics) were also inventoried. This information helps identify critical needs, including data management infrastructure. It will also inform MDBC project implementation including the selection of implementation sites, prioritize data analysis, and evaluate activities to fill critical information gaps. The inventory will help establish monitoring, sampling, data collection, and data management standards for the MDBC projects.

Analysis of the datasets in the inventory will feed into a subsequent prioritization phase for the HAE project, which will identify and prioritize data gaps to direct future field surveys in coordination with the other MDBC projects. The specific objectives of the inventory were to:

- Identify data types needed by HAE
- Identify and assess existing inventories and portals
- Develop a process for data requests
- Develop an accretion plan for inventory updates and maintenance during project implementation

The objective of this report is to detail the data inventory scope and methods, such as data discovery, screening, acquisition, and extraction. The results of this inventory will be summarized with information specific to each HAE working group. An outline of the data accretion plan, a discussion of conclusions, and future uses of the data are highlighted in the conclusion of this report. This inventory report will be published as a NOAA Data Report and will be available to the public via the NOAA Institutional Repository. The accompanying inventory dataset (GIS shapefiles) will be archived with NOAA National Centers for Environmental Information (NCEI). HAE is working with NCEI to plan for revisions of this archived inventory dataset in future years as the inventory grows. An export of the data catalog will be made available to the public on the NOAA DIVER MDBC repository.

3. Inventory Structure

The HAE project was divided into five working groups: Benthic Invertebrates, Genetics, Fish Ecology, Corals, and Environmental/Oceanography. The Coral group worked in collaboration with the MDBC Coral Propagation Techniques (CPT) project team. In order to identify datasets for the inventory, data inclusion was evaluated based on certain broad and fine scale inclusion criteria within each group (Table 1).

Table 1. The working groups selected the broad inclusion criteria that were pertinent to their inventory searches. The “X” indicates the use of the broad inclusion criteria and the “NA” indicates that the criterion is not used as broad inclusion criteria per each working group.

Broad Inclusion Criteria	Benthic Invertebrates	Corals	Genetics	Fish Ecology	Environmental/Oceanography
Taxa	NA	NA	X	NA	NA
Geographic	X	NA	NA	X	X

It was beneficial to incorporate a broad temporal range so that data that may be important, despite age, were not excluded. There were different temporal considerations depending on the inventory data types. The temporal period of data to be included in the inventory was modified based on factors such as quality, availability of certain types of data, or relevance, as well as whether it was based on current scientific knowledge.

3.1. Broad Inclusion Criteria

3.1.1. Taxa

The Genetics working group was the only group with a taxonomic scope. A cross-project (CPT and HAE) genetics ‘status of knowledge’ assessment was conducted to evaluate genetic barcode information (mitochondrial and nuclear), microsatellite and restriction site-associated DNA sequencing (RAD-seq). The assessment focused on all mesophotic and deep-sea coral taxa found in the northern Gulf of Mexico, including octocorals, stony corals, black corals and sea pens. This list of taxa was generated using Etnoyer et al. (2020). Publications related to the genetic data were uploaded to the HAE EndNote library.

3.1.2. Geographic Scope

Geographic scope is the main broad inclusion criteria for the benthic invertebrate and fish ecology working groups. Geography is tiered to identify and prioritize information near the incident site and with increasing distance from the incident area. Tier 1 includes data collected within the northern GOM that corresponds with the documented injury and potential exposure zones derived from the DWH oil spill damage assessment (DWH NRDA Trustees, 2016; Figure 1).

Tier 2 represents the geographic extent that was identified as the area of interest for the MDBC projects in Open Ocean Restoration Plan 2. This area is bounded to the south by the 27 degrees north latitude line and by the 50 m isobath to the north (Figure 2). If, however, a dataset intersected with the 50 m isobath line, extending shallower than 50 m, the dataset was included in the inventory.

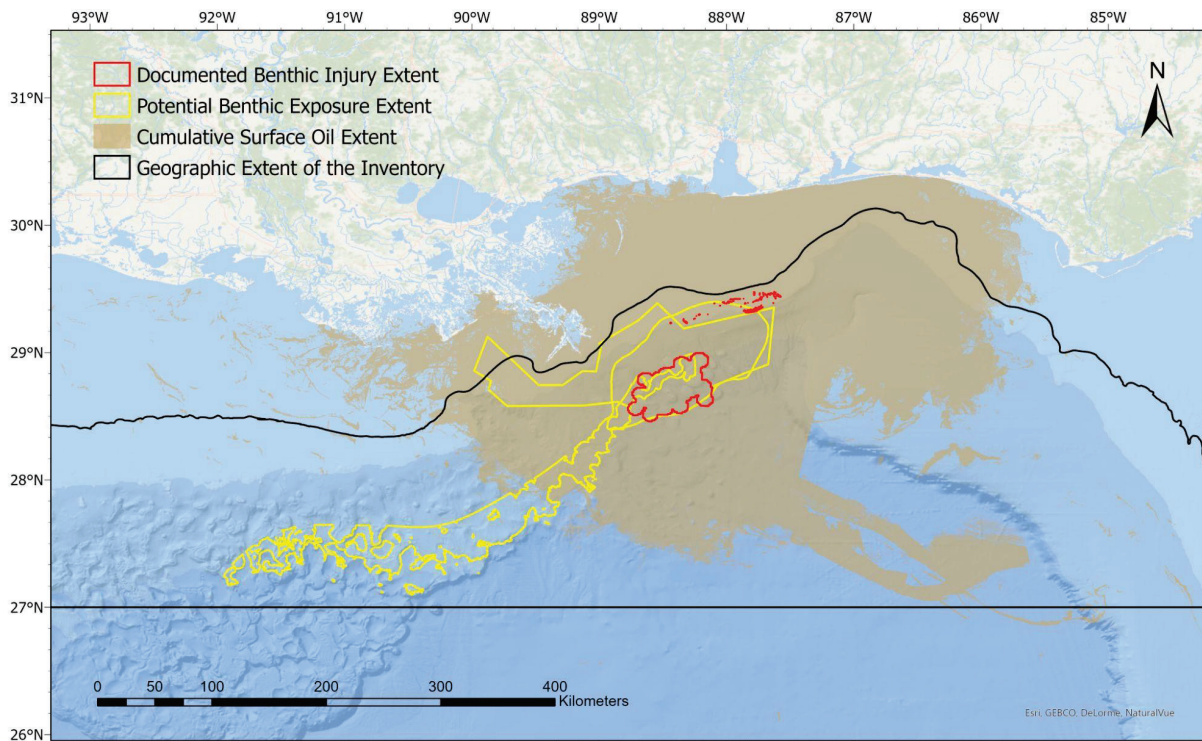


Figure 1. Tier 1 of the geographic scope for the benthic invertebrate and fish ecology working groups inventory. This is defined as the documented injury and potential exposure zones derived from the *Deepwater Horizon* (DWH) oil spill damage assessment.

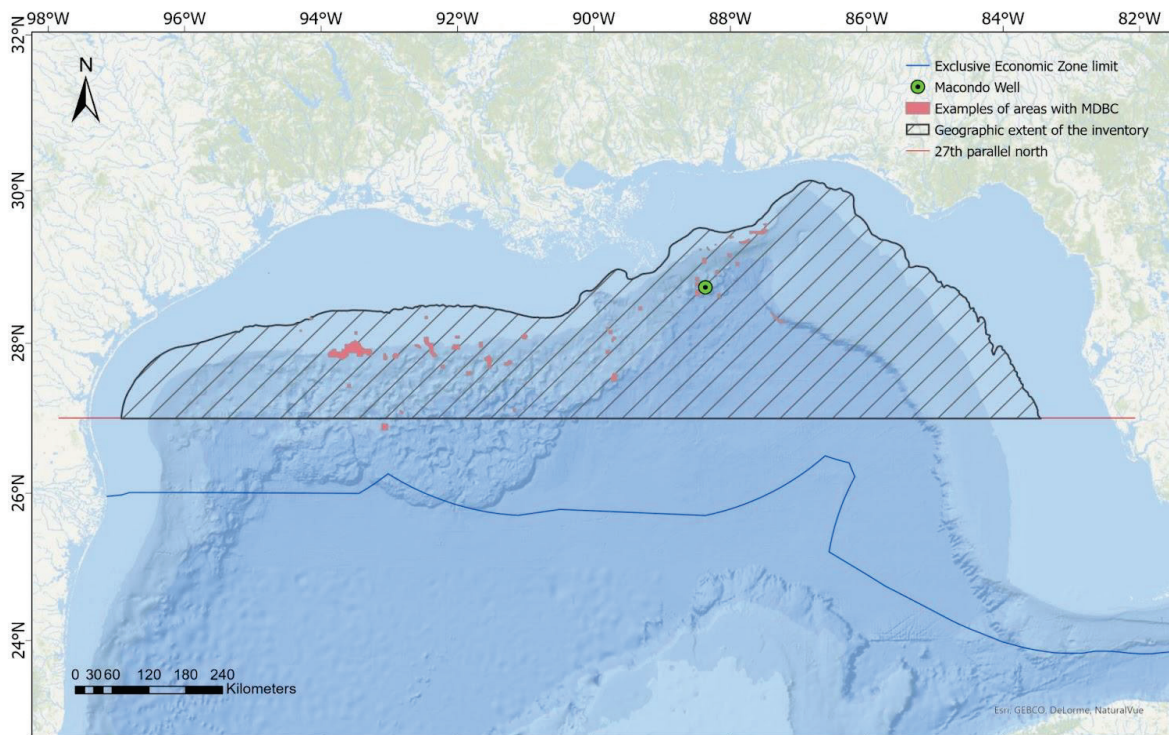


Figure 2. Tier 2 of the geographic scope for the benthic invertebrate and fish ecology working group inventory. This is defined as the northern Gulf of Mexico (GOM) north of 27 degrees latitude and generally no shallower than 50 m.

Tier 3 includes data within the GOM but outside of the areas defined by Tier 1 and Tier 2. For example, if data for a species was outside of the boundaries of the geographic extent, but the species was determined to be injured or was an indicator species, it was included in the inventory. Relevant biological or environmental information may not exist within Tiers 1 and 2 but may exist within this broader scale. This information may be useful to fill gaps within Tiers 1 and 2.

Lastly, Tier 4 were those data that were sampled outside of the GOM, again, to ensure comprehensive data on population dynamics and genetic connectivity were included in the inventory. For example, if a species in the Atlantic also occurs in the GOM in similar environmental parameters, the genetic information and community composition of the Atlantic organisms could be beneficial to our study (Frometa et al., 2021).

The corals and genetics data searches were not restricted to geographic restrictions. It was based mainly on target taxa and depth (deep sea coral species found in the northern GOM; Etnoyer et al., 2020). Additionally, datasets from the Southeast Atlantic were included due to the similarity in habitat, species presence, genetic connectivity, and community structure.

Rules applied to environmental and oceanographic data have varying geographic scopes, depending on the particular data category and type. For example, remotely sensed data has a different geographic scope than in situ sampling. Most oceanographic data (e.g., Sea Surface Temperature, currents) was included for the entire GOM region since their characteristics encompass and are often dependent on large portions of this area. Other environmental and oceanographic data are more locally-dependent and vary based on specific geographic location (e.g., water column characteristics). For these data, focus was placed on areas within the broad impact zone (Tiers 1 and 2) with the option to expand beyond this area as needed.

Additionally, other select data types (e.g., contaminants) were considered under a broader geographic scope that encompassed a global scale. For these, emphasis was placed on impact-related data such as the effects of a contaminant on fauna or the marine environment. For example, the effects of a particular contaminant on fauna or habitat quality is likely to be similar regardless of where the study had been conducted. Therefore, this type of information can be relevant to the GOM region, including the species and habitats it encompasses, even if the data itself was not collected there.

It is important to note that these tiers or other variations in defined geographic extent did not reflect importance or value, but represented how the search for datasets or portals would occur.

3.2. Fine Inclusion Criteria

3.2.1. Data Characterization Attributes

Certain attributes were used to characterize the identified datasets and differed based on data type. A table of attributes was compiled by the data management team with input from cross-project groups. These attributes were used during the data screening process and some were used as fine inclusion criteria. A comprehensive list of cross-project attributes can be found in Appendix A.

The minimum required metadata to characterize datasets are included in the table below (Table 2) and are separated by working group as each differs in requirements. Discovered data that did not have formally written metadata (e.g., paper, presentation, etc.) was screened for any existing metadata that could be utilized for screening.

Table 2. The working groups selected the minimum metadata requirements to be met with an “X” or not required with a null or “NA”. * Latitude/Longitude and depth are not listed as minimum metadata requirements for the Environmental/Oceanography category because they were not required for all data types falling under this category. Wherever possible, the aim was to have this metadata.

Minimum Metadata	Benthic Invertebrates	Corals	Genetics	Fish Ecology	Environmental/Oceanography
Sample Type	X	X	X	X	X
Data Source/Provider	X	X	X	X	X
Date and locality of Collection	X	X	X	X	X
Latitude/Longitude	X	X	X	X	NA*
Depth	X	X	X	X	NA*
Collection Procedure	X	X	X	X	NA*
Analysis Procedure	X	X	X	X	NA*

3.2.2. Data Categories

The working groups focused their data discovery on nine data category topics (Table 3). This table was a guide for data searches based on which categories were considered most applicable to each working group. It does not explicitly represent the importance of each category in relation to any specific working group. For example, genetic data is highly relevant to the benthic invertebrate working group but discovery of this type of data was the focus of the genetics working group.

Table 3. Working groups focused their data search and discovery on those data categories which have an “X”. The “NA” indicates that the category was not a criterion per each working group.

Data Categories	Benthic Invertebrates	Corals	Genetics	Fish Ecology	Environmental/Oceanography
Age/growth	X	X	NA	X	NA
Reproduction	X	X	NA	X	NA
Trophodynamics	X	X	NA	X	NA
Health/Condition	X	X	NA	X	NA
Community Metrics	X	X	NA	X	NA
Genetics	NA	X	X	NA	NA
Contaminants	NA	X	NA	NA	X
Geology and Seafloor/Sediment Characteristics	NA	X	NA	NA	X
Physical and Chemical Oceanography	NA	NA	NA	NA	X

3.2.3. Data Types

The working groups collected both raw and processed data for the inventory, including, but not limited to, direct observations from deep submergence vehicles, discrete collections of specimens, genetic sequences, sediment community data, and environmental data. Available data types varied across the working groups (Table 4). Ancillary data (e.g., conductivity, temperature, and depth [CTD] data) and derived or synthetic data products such as maps and graphs, as well as existing metadata, were also captured as part of the inventory.

Table 4. The working groups identified data types to be considered for their inventory searches. Some examples of these data types are listed for clarification. This is not an all-inclusive list of data types and more may be added during the implementation phase.

Data Types	Benthic Invertebrates	Corals	Genetics	Fish Ecology	Environmental/Oceanography	Examples
Video/imagery	X	X	X	X	X	Video; still photographs; microscopy
Skeleton/tissue samples	X	X	X	X	NA	Whole body; liver; otolith; coral fragment; shells
Sediment Samples	X	NA	X	NA	X	Surface sed.; cores; grabs
Geological and geophysical	X	X	NA	X	X	Seismic surveys
Water samples	NA	X	X	NA	X	Surface water; sediment-water interface; water column
Genetic sequences	X	X	X	X	NA	Mitochondrial and/or nuclear DNA barcodes; Environmental DNA reference libraries; multi-locus genotypes
Genetic data products	X	X	X	X	NA	Phylogenies; pop gen; genomic; STRUCTURE graphs; sequence alignments
Models	X	X	X	X	X	Habitat suitability; dispersal (e.g., biological, contaminant); population connectivity; oceanographic; age/growth
Ancillary data/supplementary information	X	X	X	X	X	Nav tracking (vessel/platform); CTDO (CTD optical), Acoustic Doppler Current Profilers (ADCP)
Literature findings	X	X	X	X	X	Published results i.e. analyzed data, habitat suitability indices/models; infaunal community composition

4. Methods

4.1. Data Sources (Inventory of Inventories)

During data discovery, information associated with published literature, known public data portals, those identified by literature review, and data accessible directly from organizations was broadly searched. Below, we detail the systematic and reproducible approach for data discovery. A list of data sources is provided in Appendix B.

Originally, the plan was to focus on species within each group, specifically those impacted by the spill, and indicator species. However, a better and more synergistic approach was decided upon using a literature review framework.

This literature review framework was developed to include the metrics and methods that were to be the focus of the search. The framework consisted of three objectives with six parameters nested within (see Section 4.1.1.). Specific metrics and methods were included in the parameters. A full list of these metrics and methods can be found in Appendix C.

4.1.1. Literature Search

Data pertinent to the HAE inventory was located in or accompanied by peer-reviewed literature published in scientific journals. Data were also reported in various types of gray literature, such as reports, theses, and other non-peer-reviewed publications. To search for these data sources, a systematic literature search was conducted using literature databases, such as the Web of Science (see Appendix B for more sources).

Using the literature review framework below, information was gathered from peer-reviewed and gray literature for fish, coral, invertebrates, and infauna for Parameters 1, 2, 3 and 6. Parameter 3 focuses on genetics so this was the focus of that working group. Parameter 4 is the only taxa-specific parameter (stony and soft corals). Parameter 5 is cross-project relevant and therefore any appropriate literature was tagged when it was encountered, but this topic is being handled by the project management team.

Objective #1: Fill data gaps and determine baselines for health and condition to guide restoration and protection

Parameter 1. Sampling sufficiency for population/communities

Parameter 2. Identified species and habitat interactions

Parameter 3. Identified metapopulation dynamics and genetic diversity

Parameter 4. Coral larvae dispersal

Objective #3: Detect and quantify trends for inference of future impacts and assess success of restoration

Parameter 6. Recovery Trajectories and restoration targets

A large catalog of literature was initially compiled by the data management team in a shared EndNote™ library (EndNote™ 20, 2020) by importing peer-reviewed papers, gray literature, and technical reports from other DWH-related bibliographies. Folders were created for each of the five HAE groups (BIN, COR, GEN, FIS, and ENV). Literature data sources that were not already in this EndNote™ library were

added to the library and all pertinent literature was marked with a code for the appropriate working group within the HAE project and with the literature review framework objectives and parameters that were addressed in the paper (e.g., “BIN OBJ 1 PAR 1” for benthic invertebrates, objective 1, parameter 1). The papers were also tagged if they contained spatial data that could be used in the HAE geographic information system (GIS) Geoplatform. The Geoplatform is an online portal used to share geographic data, maps, and online services.

4.1.2. Portals and Databases

The following steps were taken to search existing portals and data repositories:

- Compiled initial list of data portals to contribute to cross-team portal tracker
- Discovered metadata on portals
- Screened datasets using broad inclusion criteria
- Screened datasets using data characterization attributes
- Extracted screened data from portals

A comprehensive cross-team portal catalog was created by the MDBC Technical Team and includes access details, host organization, scope, data availability, and relevance. Data portals that were reviewed by the HAE working groups include the following, but this list is not exhaustive and is for illustrative purposes. See Appendix B for a full list of portals utilized by HAE team:

- NOAA Deep Sea Coral Data Portal
<https://deepseacoraldata.noaa.gov/>
- NOAA DIVER
<https://www.diver.orr.noaa.gov/>
- NCBI GenBank
<https://www.ncbi.nlm.nih.gov/genbank/>
- USGS ScienceBase
<https://www.sciencebase.gov/catalog/>
- DRYAD
<https://datadryad.org/stash>
- NOAA OER Video Portal
<https://www.ncei.noaa.gov/access/ocean-exploration/video/>
- GRIIDC
<https://data.gulfresearchinitiative.org/>

4.1.3. Organizations

The literature review resulted in the identification of organizations and individuals possessing relevant data that are not stored in online portals. In these cases, the inventory team worked directly with the organizations to include their data in the inventory. Organizations included federal agencies, state agencies, academic institutions, private industry, foundations, and non-profit organizations.

When approaching organizations to inquire about data, the HAE working groups followed the external point of contact engagement protocol developed under and approved by the MDBC Technical Team. Data from organizations was searched for using the steps below. In the rare case where an inquiry did not result in the acquisition of a dataset, this was noted in the tracker.

- Identified potential point of contact for a desired dataset
- The organizations and points of contact were shared in the external engagement tracker
- Directly engaged with point of contact
- Tracked communications with each organization
- Screened and extracted discovered data

4.1.4. Partner Engagement

Working group leads engaged informally with partner agencies, scientists, and stakeholders to gather information, ask questions, and garner advice during the data inventory process. This was performed in a coordinated manner using the external engagement tracker created by the MDBC Technical Team.

4.1.5. External Engagement for Purpose of Discovering Datasets

HAE staff engaged informally with external contacts who may possess or have access to datasets that were not available in an online portal or in searchable literature. This was performed in a coordinated manner using the external engagement tracker. Working groups helped to identify scientists, agencies, and stakeholders to add to project contact lists.

4.2. Data Screening

Datasets identified during the data discovery process were screened for relevance to the objectives and parameters created for the HAE project. Screening occurred initially based on broad inclusion criteria followed by fine inclusion criteria. Evaluated datasets were scrutinized to ensure the data met minimum quality/standards, availability of metadata, and level of effort for acquiring data and analysis.

4.2.1. Initial Screening

The initial screening based on broad inclusion criteria differed slightly by data source, as follows:

- Data portals: Queried data portal using metrics from the literature review framework or, for genetics, using specific species names.
- Peer-reviewed publications: Title and abstract screening followed by full text screening. If data were available, then the data was screened directly. Otherwise, the full text of the publication was screened to determine if data were relevant
- Gray literature: Title and abstract screening followed by full text screening a. If data were available, then the data was screened directly. Otherwise, the full text of the publication was screened to determine if data were relevant
- Organizations: Requests for data were made by email/phone call/other for relevant data.

4.2.2. Subsequent Screening

Following initial screening, data was reviewed for adherence to fine inclusion criteria by each HAE working group. This workflow was revised following initial screening to include a list of objectives, parameters, and qualitative attributes that data must meet. These attributes differed among subject matter and types of data (e.g., benthic invertebrates, corals, genetics, fish ecology, oceanographic). A list of all cross-project attributes is included in Appendix A.

4.2.3. Tracking the Screening Process

All screened datasets were recorded in a data catalog spreadsheet. The spreadsheet serves as a record of datasets that were:

- Excluded during the screening process
- Included for acquisition of data

Those data lacking metadata or with poor metadata, were kept and will be evaluated as to whether they can fill identified gaps in the forthcoming gap analysis. Tracking excluded datasets helped reduce the chances for duplication of effort (e.g., two inventory team members screen the same dataset). Papers, reports, and datasets that were deemed not relevant to a group were marked “deferred” and tagged by the group that deferred them. If a paper deferred by one group had possible relevance to another group, it was placed in an “of interest” folder for that group.

4.3. Downloading from Portals

After searching a portal, datasets identified as having met the broad inclusion criteria were downloaded to a Google Drive folder created by the MDBC Data Management team.

4.3.1. Material from Data Protocols (Workflow)

A workflow processes was created from initial identification of the datasets (Figure 3). All datasets are assigned unique identification and metadata are organized in a data catalog. Raw or processed spatial data sets will be added to the HAE geoplatform. All data will be evaluated to identify spatial and informational gaps. Where data not currently identified or obtained by HAE may be sought to help fill gaps or prioritized gaps may be focal points for HAE field work during implementation.

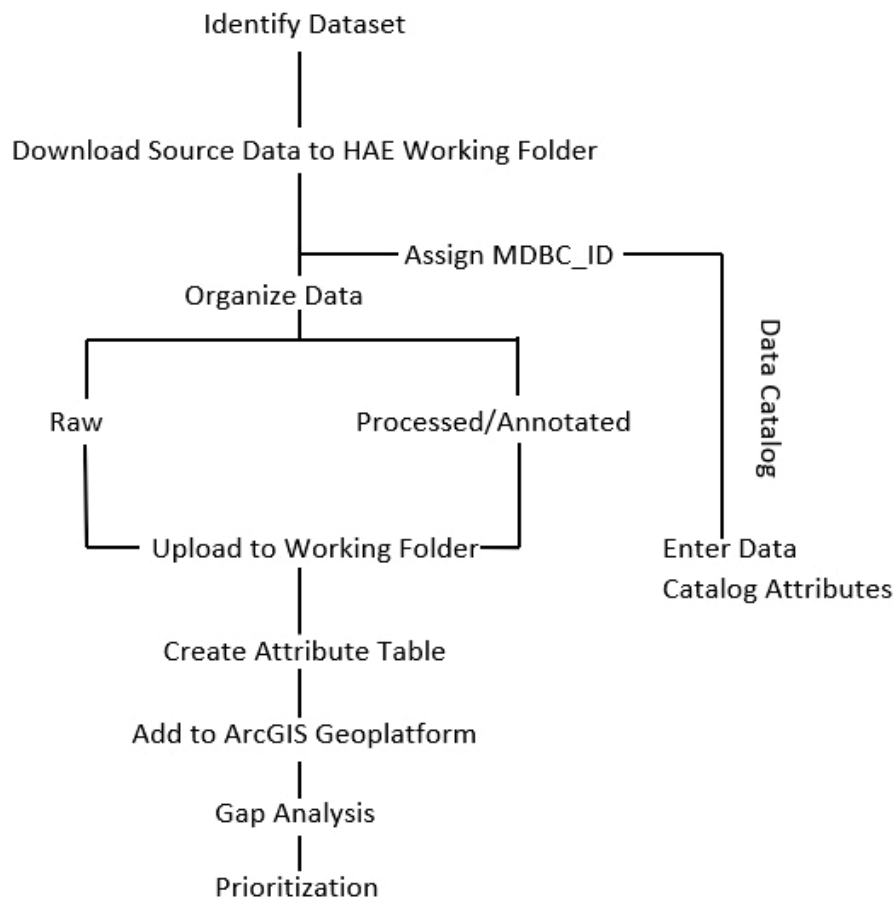


Figure 3. Workflow diagram for the identification, acquisition and prioritization of datasets.

4.3.2. Data Catalogs per working group

When a dataset was deemed relevant to the HAE data inventory, it was assigned a unique identifier for the context of the MDBC project. The unique identifier, referred to as the “MDBC_ID” is two-part code. The first part of the code is a three-letter prefix which designates the HAE working group to which it belongs (FIS, COR, BIN, ENV, GEN). The second part of the code is a six-digit numerical suffix “#####” designating the dataset number; the numbers were assigned sequentially as datasets were added to the data catalog. For example, the first BIN dataset was BIN000001; the second was BIN000002, and so forth. Once datasets were assigned an MDBC_ID, they were added to the MDBC Data Catalog, which will be archived in the NOAA DIVER MDBC repository. While the data catalog is constructed at the MDBC Portfolio level, there are spreadsheets within the catalog that are specific to each of the HAE working groups, with the attributes and column headings that are most relevant to that group. Many datasets were pertinent to more than one group. In this case, after one group entered the dataset into the data catalog, the other groups copied the information and pasted it into their tab in the data catalog.

Each group had occasion to deviate slightly from the search methodology based on the type of data for which they were searching.

- Benthic Invertebrates: these datasets came exclusively from the shared EndNote™ library (see Section 4.1.1.) and a previous NOAA and Bureau of Ocean Energy Management (BOEM) project

in the Gulf of Mexico. During the accretion phase, a more comprehensive literature search and a search of portals and databases will be completed.

- **Corals:** The majority of datasets were already known to the MDBC working group. Many datasets contain samples collected from research cruises related to MDBC work. Coral datasets added to the data catalog are also of interest to the CPT project.
- **Fish Ecology:** The majority of datasets for this inventory were identified from various data portals, cited in Section 4.1.2, then external engagement, and lastly literature review.
- **Oceanographic and Environmental:** Relevant datasets from other groups (that had oceanographic and environmental data) were pulled. Then, major data portals were searched for relevant datasets. All datasets were screened to ensure there was no duplication between those pulled from other groups and those pulled from data portals. Datasets sourced from literature in the shared EndNote™ library were not prioritized at this stage as many were older, unlikely to be in digital format, and unlikely to be easily acquired. More datasets will be added during the data accretion phase of this project by searching relevant literature both from within and outside the shared EndNote™ library.
- **Genetics:** The genetics search was based on the taxa found in the northern GOM, according to Etnoyer et al., 2020. All taxa were searched on GenBank databases for all publications and published genetic data (mitochondrial and nuclear gene barcodes, RAD-seq data, microsatellites, etc.). Data was compiled into a spreadsheet, including references (see Section 8.5).

4.4. Inventory Products

Three products were derived from the data inventory:

- A literature database (EndNote™ library)
- Topical summaries, which are assessments and summarizations of best practices for field implementation and identification of data gaps derived from the literature
- GIS data layers compiled into a geoplatform for visualization

4.4.1. Literature Database (EndNote™ Library)

Literature is an open data source for the MDBC HAE Data Inventory. While searching for data and compiling the inventory, the Inventory and Analysis team developed a literature database for all peer-reviewed and non-peer-reviewed literature. It includes publications that passed title and abstract screening. The screening process, outlined in this document, used the metrics in the literature review framework (see Section 4.1.1) as the screening phrases. Records were kept in an EndNote™ library shared across MDBC teams, and contain, at minimum, the following:

- Title of Article
- Author (s)
- Journal or Source
- Reference Type (e.g. book, journal article, thesis, etc.)
- Publication Year
- DOI (digital object identifier, if applicable)
- Tags for MDBC Parameter and spatial data presence (see Section 4.1.1)

- Brief summary of key findings as it pertains to MDBC communities and laboratory/fieldwork best practices (“research notes”)
- PDF attachment or web link

4.4.2. Topical Summaries

Summaries of the literature gathered during the literature review process were written based on the topical parameters in the literature review framework, which are listed in Table 5. Standard operating procedures, methods documentation, and collection procedures were synthesized and organized by topic along with community metrics, habitat interactions, and recovery trajectories (Table 5).

Table 5. The working groups summarized the knowledge gathered in the literature review by writing topical summaries based on the literature review framework. This table shows the summaries for each parameter by HAE group or subgroup.

Parameter	Benthic Invertebrates			Corals		Fish Ecology	Environmental/Oceanography
	Infauna	Sponges	Mobile Invertebrates	Coral Ecology	Coral Larval Dispersal		
Sampling sufficiency for population / community metrics	X	X	X	X		X	
Identified species and habitat interactions	X	X	X	X		X	X
Identified metapopulation dynamics and genetic diversity				X			
Coral larvae dispersal					X		
Recovery trajectories and restoration targets	X	X	X				

4.4.3. MDBC Geoplatform Inventory

The data gathered in the inventory were transformed into spatial GIS layers and uploaded to a shared geodatabase that will be used to support the prioritization, gap analysis and implementation efforts of MDBC projects. New data through additional discovery or generated by MDBC projects will be integrated into the geospatial database for the duration of the HAE project following the data inventory accretion plan. One subset of spatial data, the spatial bibliography, is derived from publications of interest in the literature database. When possible, coordinates and associated spatial information were extracted from the text of publications that were not otherwise represented in the Geoplatform inventory. This information was used to create GIS layers to represent and summarize the sampling locations of each publication of interest. These and all other inventory layers address the parameters included in the literature review framework (Table 5) and include, but are not limited to:

Data

- Taxa observational data
 - Remotely Operated Vehicle (ROV)/Autonomous Underwater Vehicle (AUV)/submersible stills & video transects, stationary cameras, diver surveys
- Coral & benthic invertebrate taxa sampling

- ROV, submersible, cores
- Fish taxa samples
 - Trawls, traps, angling, plankton nets
- Sediment cores (Ship/ROV/Human Occupied Vehicle [HOV])
- Location of age/growth studies
- Location of image-based monitoring of corals
- Location of genetic sampling
- Location of contaminant/petroleum compound sampling & observations
- Environmental/habitat characteristics
 - Broad scale (satellite)
 - Temp
 - Salinity
 - Currents
 - Dissolved Oxygen
 - Chlorophyll-A
 - Finer scale (in situ)
 - Location of landers, CTD, Acoustic Doppler Current Profilers (ADCP), etc.
 - Temp
 - Currents
 - Dissolved Oxygen
 - Contaminants/oil/hydrocarbons
 - Sedimentation
 - Hard bottom /Soft bottom

Analyses

- Coral larval dispersal
 - Model approach
 - Species specific Habitat Suitability Index maps
- Community metrics of injured vs non-injured sites
- Development of sentinel sites

4.5. Updating the Inventory

4.5.1. Data Accretion Plan

During HAE project implementation, the inventory and the Geoplatform will continue to grow by obtaining data that were identified but not acquired during the planning phase, newly discovered data and data generated during HAE project activities. The HAE Inventory Working Group will annually repeat the process described in this report.

4.5.2. Periodicity of Inventory Assembly

Each working group (benthic invertebrates, corals, genetics, fish ecology, environmental/oceanographic) will be required to update the inventory annually. Overall, it is anticipated that two staff members will each spend a month (320 hours) for inventory and geospatial database updating. This may be adjusted during project implementation as needed to accommodate data volume. MDBC data collections will most likely occur between May and October. Raw data may take various amounts of processing until ready for the data catalog but will be added as soon as processing and appropriate review, if any, are complete. Additionally, literature review, data portal mining, and POC, or point(s) of contact, engagement will be conducted to discover new data. Acquiring data may be the result of a simple request or a financial process to obtain data.

Monthly updates will be given to the HAE project management team on the status of data requests or inventory updates. During winter/spring, annual update addenda will be added to the inventory report with appropriate metadata. The data inventory catalog will be submitted to NCEI for archival in the NOAA DIVER MDBC repository.

5. Archiving

The collection of inventory datasets (GIS shapefiles) and their accompanying metadata will be archived at NOAA NCEI. The data inventory catalog will be archived at the NOAA DIVER MDBC repository or another archive that abides by Federal standards. If there is no appropriate, federally-managed archive then data will be archived at a repository accepted by the relevant scientific and technical communities.

6. Planned Uses of Data Inventory

6.1. Workshop

A workshop was held on December 1 and 2, 2021 with subject matter experts (SMEs) in the scientific fields relevant to HAE. The topical summaries were provided to the SMEs prior to the workshop and were presented by the authors in a more concise format during the workshop. Breakout discussions for direct feedback were held after presentations addressing each parameter in the literature review framework. In addition to the topical summaries, a concise version of the full data catalog for each HAE group and a references list of the literature that was tagged during the literature review, was provided to the SMEs that attended the workshop. This provided HAE with the opportunity to obtain external feedback from the appropriate SMEs to evaluate the products' completeness, provide guidance on missing information, and provide guidance on how and where to fill gaps.

The information gathered during the workshop will be summarized, published and will inform HAE data prioritization, gap analysis and implementation activities.

6.2. Gap Analysis

A primary goal of this comprehensive inventory was to identify data that allows the team to assess informational, spatial, and temporal gaps useful for prioritizing and conducting future MDBC HAE activities. The objective of the gap analysis is to identify full and partial data gaps in available data required to support our understanding of community metrics/composition, genetic information, population dynamics. In general, gaps will be analyzed based on depth, management type, impacted/reference site, and habitat type. Gaps may be represented by datasets that are incomplete in their analyses, raw datasets, research that has not been conducted, and/or datasets that did not meet the criteria described in the above sections. Gaps will also be identified through SME review and the topical summaries resulting from the HAE and HAE/Habitat 4 Fish workshops.

A strategy for the HAE gap analysis is currently being developed by the HAE team but the activity is anticipated to be complete in March 2023. The gap analysis will target the entire study area (Tier 2 in Figure 2), however information from Tiers 3 and 4 may be useful in gap filling as they pertain to species and habitats similar to those in the target area.

6.3. Prioritization of Topical and Spatial Data Gaps

Data from the inventory will assist in the prioritization of areas to be sampled during field operations.

All datasets will be evaluated with cost/benefit criteria to prioritize data that are most helpful to meet HAE project goals. The inventory teams will define a systematic process for data prioritization, with criteria to consider whether the dataset:

- Is available but requires processing by owner for a cost?
- Is available but requires processing by the HAE team?
- Is available but could be collected more cost effectively through MDBC fieldwork?

Dataset information that will specifically address the prioritization includes:

- Changes to structure and function of MDBC impacted by the DWH oil spill;
- Environmental baseline data on conditions and changes over time around impacted and healthy MDBC, to guide direct restoration and protection;
- Dispersal models for coral larvae species impacted by the spill;
- Data gap assessment and evaluation of sites to be selected for potential direct restoration and protection activities at both injured and reference sites;
- Data on impacts from natural and anthropogenic threats to MDBC (e.g., oil spill related impacts, invasive species, water quality anomalies, vessel anchoring, fishing impacts, marine debris, contaminant releases, marine heatwaves, and climate change);
- Data to detect and quantify trends affecting prioritized MDBC and to assess success of restoration efforts with respect to recovery, natural mortality, and growth rates; and
- Baseline data for health and condition to guide direct restoration and protection of MDBC.

6.4. Operations Planning

Fieldwork operations plans are being developed for MDBC projects using the data gathered from the data inventory, topical summaries, MDBC Geoplatform, and workshop subject matter expert input. In addition, field work will be coordinated with CPT, Mapping, Ground-truthing and Predictive Modeling (MGM), and Active Management and Protection, using objectives that were defined by similar approaches for prioritizing sampling activity and location. This collection of information from all four MDBC groups will assist in the effort to refine data requirements and define best practices for MDBC fieldwork methodologies and sample collection. Future workshops will review and update data requirements and protocols.

7. Results

During the data inventory process, hundreds of datasets were considered for inclusion. Of these, 369 datasets passed broad inclusion criteria and were further screened based on fine inclusion criteria. In total, 364 datasets passed fine inclusion criteria and were either incorporated into the inventory or are in the process of being included into the inventory. In addition to the datasets, 1,984 literature references, 10 topical summaries, and 82 newly created GIS layers were included in the inventory in addition to those created from the spatial bibliography (see Section 4.4.3). An estimated 46 more GIS layers will also be created based on the current inventory of data. See Table 6 for a breakdown of dataset results by group.

Table 6. Overview of the datasets acquired by each HAE working group. *Estimated.

Summary	Benthic Invertebrates	Corals	Fish Ecology	Environmental/Oceanography
Number of datasets	45 and 7 'in review'	50 and 10 'in review'	76	184
Date range of datasets	1971–2021	1858–2022	1982–2021	1869–2022
Number of GIS layers created from datasets	16	5	61	46*
Number of papers in the literature database	209	131	244	1400
Number of datasets that didn't pass fine Inclusion Criteria	2	2	1	NA

Combined, these data represent multiple sources, gear types, data types, genetic barcodes, collection time periods, and environmental/oceanographic data that is relevant to all MDBC project teams. The data include multi-year, multi-gear sampling surveys, as well as model prediction data.

Due to the multi-project nature of many data collection cruises, all data types (Table 4) and sampling equipment for each dataset may not pertain to all groups who chose to include the dataset in their inventory. For example, a dataset chosen by the fish ecology group may contain mapping and sediment samples and, even though the fish ecology group does not need this data, because it is part of the dataset it is identified as a data type in their inventory.

Statistical data are provided for each HAE group. There are not any combined HAE statistical representations due to the fact that many of the datasets were chosen for more than one group, creating duplication in the comprehensive data. To separate those parts of each dataset (gear type, data type, etc.) that pertain to each individual group was not feasible. For this reason, the following charts and graphs may include information that is part of the dataset but not actually needed for that group.

Ancillary data shown in the Data Types pie charts in this chapter, include data from CTDs, fluorometers and ADCPs. The “Other” category was for data types that were not defined in the inventory information. The instrumentation in the “Other” category in sampling instruments varied by group. They are as follows:

- Fish Ecology: Bandit Reel and Block Array

- Benthic Invertebrates: Bushmaster, Radiometer, Remote Sensor and Stereo Baited Remote Underwater Video
- Corals: Bandit Reel
- Environmental/Oceanographic: Bandit Reel, Block Array and Sediment Grab

7.1. Fish Ecology

Seventy-six fish ecology datasets were obtained, of which 38 were specific to fish ecology and 38 were from other HAE working groups.

7.1.1. Data Types

Most data types for fish ecology (FIS) were identified as preserved samples or imagery from video or still photos (Figure 4). Figure 4 also lists other data types collected with fish information such as mapping data and environmental data, noted by Ancillary in Figure 4.

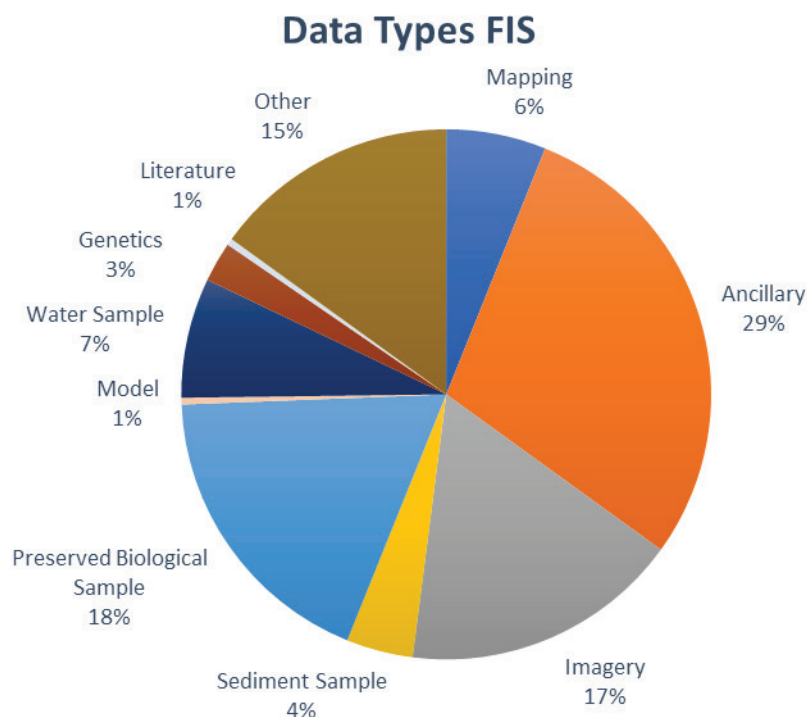


Figure 4. A pie chart displaying the percentages of data types in the datasets collected for the fish ecology (FIS) data inventory. Ancillary includes data from CTDs, fluorometers and ADCPs.

7.1.2. Data Sampling by Instrument

Many different types of sampling instruments were used in the discovered datasets (Figure 5). The most common was imagery (video/stills) collected by ROV, trawls, longline and acoustic data. Other ancillary sampling instruments are noted in Figure 5.

Sampling Instruments FIS

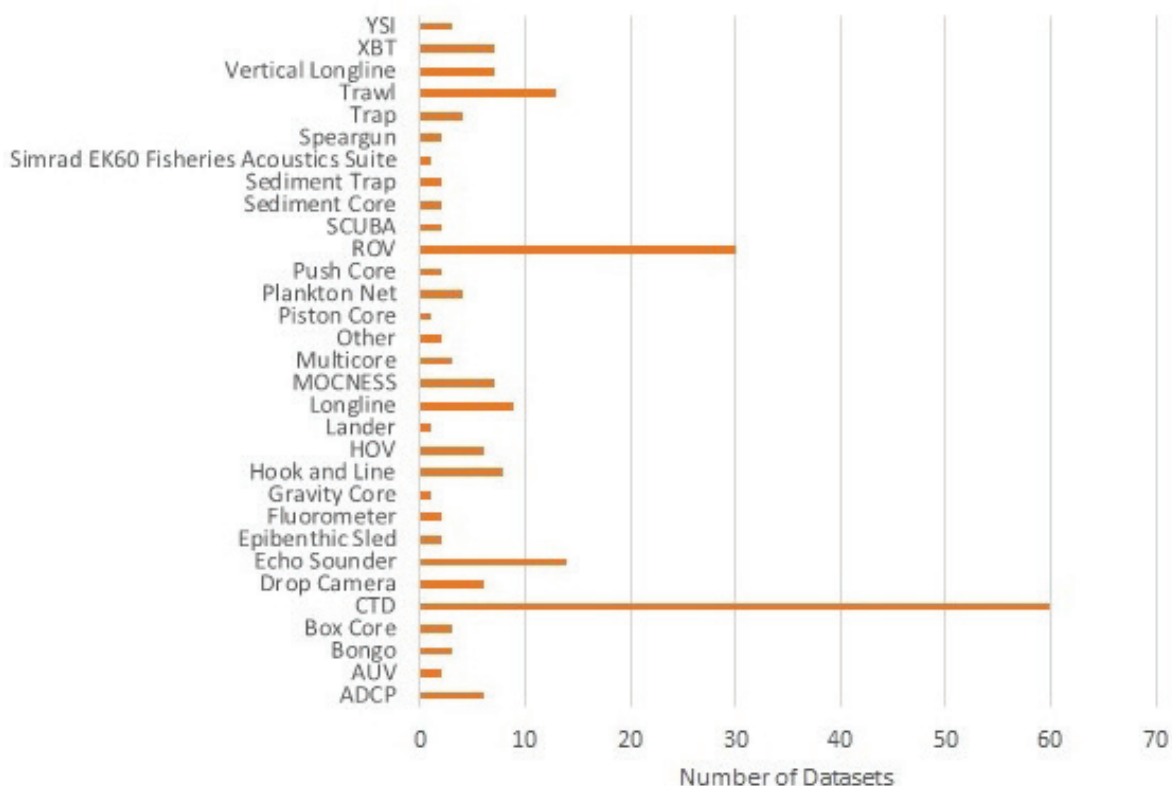


Figure 5. A bar graph displaying the number of fish ecology (FIS) datasets that used the listed sampling instruments, fish collection techniques and other associated techniques such as water quality and current measurements. “Other” includes Bandit Reel and Block Array.

7.1.3. Spatial Coverage of Data

As seen in Sections 7.1.1 and 7.1.2, there were a variety of data types and sampling instruments collected and used in the datasets acquired for the fish ecology inventory. To depict these spatially, they were plotted in ArcGIS Pro version 2.7.0 adhering to the spatial reference for geographic NAD 1983 (2011, as recommended by the MDBC MGM team’s coordinate reference systems topical summary). There was a total of 61 layers created of 73 datasets compiled for this inventory which were depicted by the sampling methods conducted during those surveys (Table 7).

Table 7. The number of layers and datasets that were represented spatially in the fish ecology data inventory by data type.

Spatial Data Type	Number of Layers	Number of Datasets
Submersible survey tracklines	13	21
Towed camera survey tracklines	1	1
Bottom video survey sites	10	10
Trap survey sites	1	1
Hook and line survey sites	14	14
Trawl surveys	12	12
Multi-Gear surveys	10	12

There were 13 layers from 21 datasets and one dataset with a layer that utilized submersible vehicles and a towed camera system, respectively to survey areas that were annotated for fishes and other attributes (Table 7; Figure 6). There were 10 datasets with layers that deployed various underwater camera arrays and were annotated for fishes and other attributes (Table 7; Figure 7). There was one dataset that utilized traps to collect fishes offshore in the northern GOM (Table 7; Figure 8). There were 14 datasets with layers that deployed various hook and line gears to collect fishes in the northern GOM (Table 7; Figure 9). There were 12 trawl survey datasets with layers that collected fishes and other data information throughout the GOM (Table 7; Figure 10). There were 10 layers from 12 datasets from surveys that utilized multiple gears (Table 7; Figure 11).

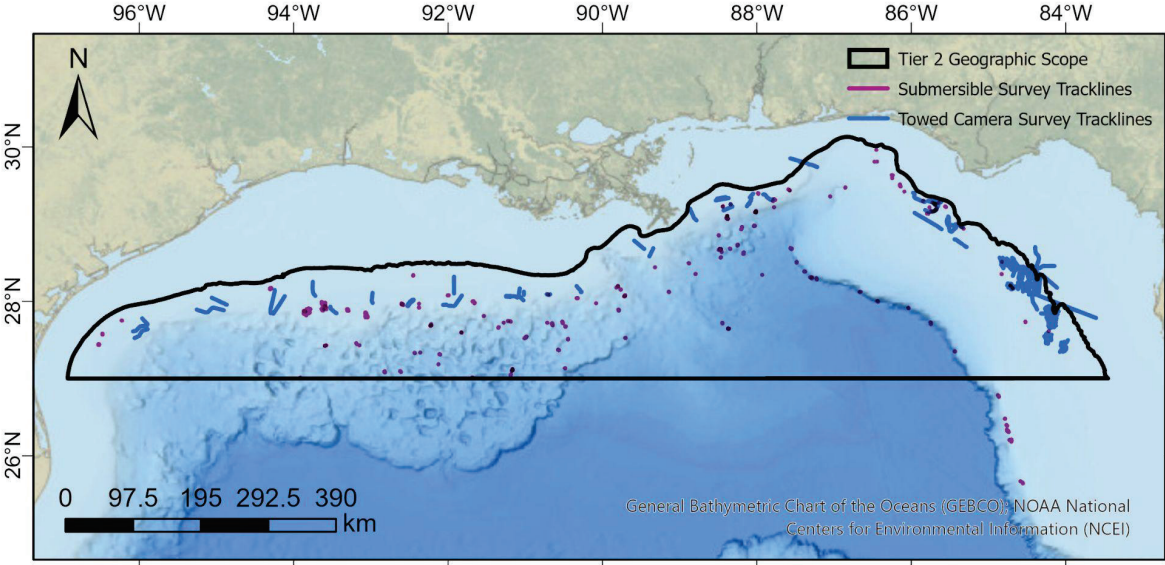


Figure 6. Submersible vehicle and towed camera system tracklines used during various surveys collected for the fish ecology data inventory.

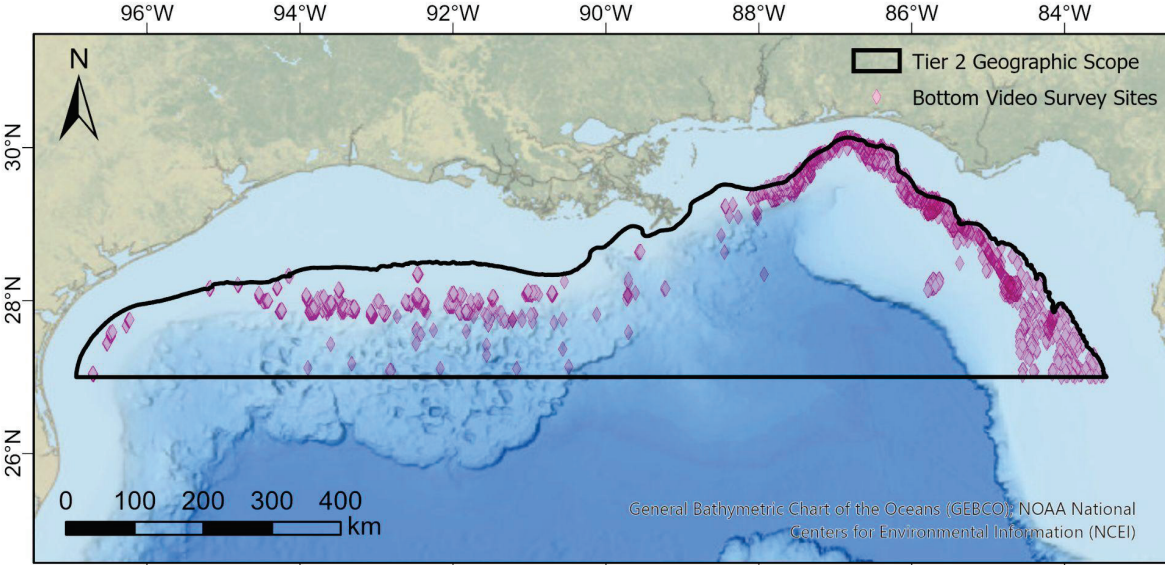


Figure 7. Deployment sites of various underwater camera gears, such as submersible vehicles and submerged baited underwater video arrays used during various surveys collected for the fish ecology data inventory.

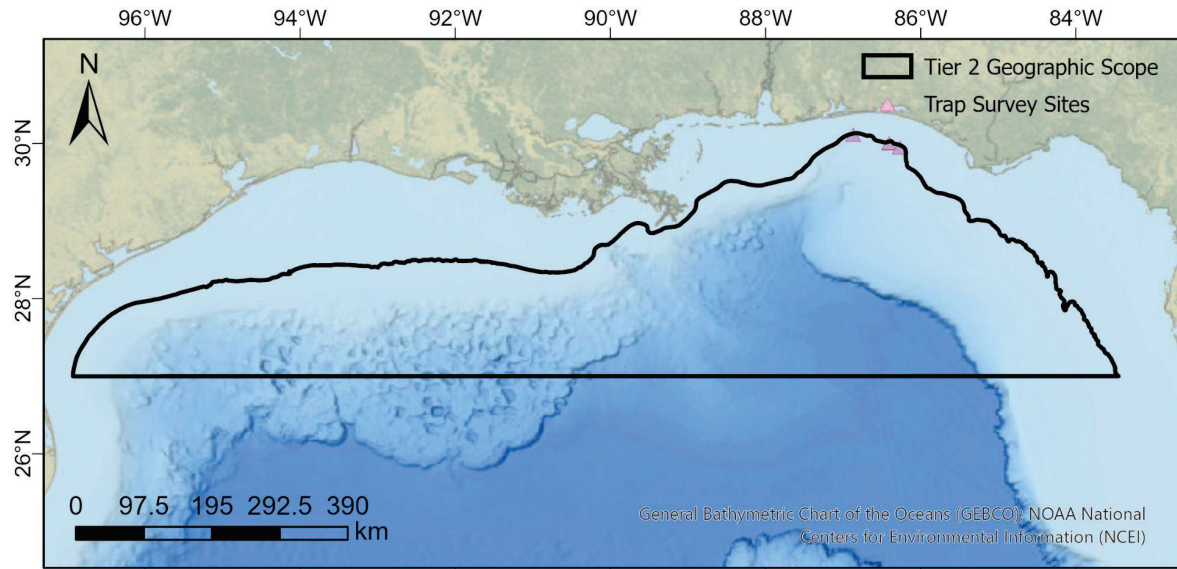


Figure 8. Deployment sites of traps by the Panama City NOAA NMFS Southeast Fisheries Science Center to identify and collect fish(es), supplementary to their bottom video survey.

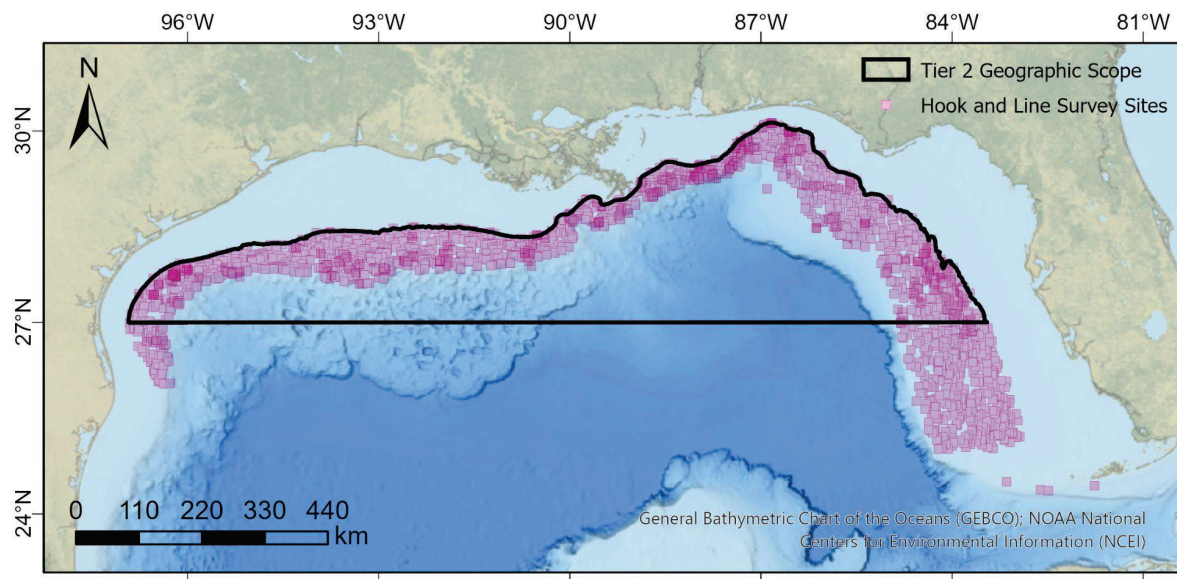


Figure 9. Deployment sites of various hook and line gears, such as bottom longline and vertical longlines, used during various surveys collected for the fish ecology data inventory.

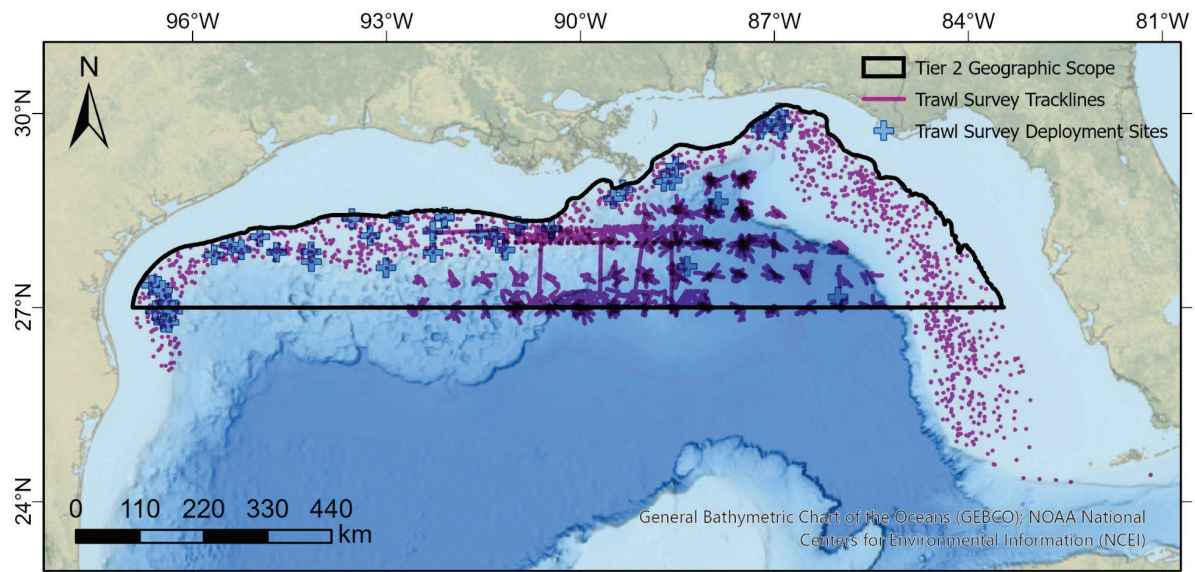


Figure 10. Trawl survey tracklines and deployment sites in the GOM used during various surveys collected for the fish ecology data inventory.

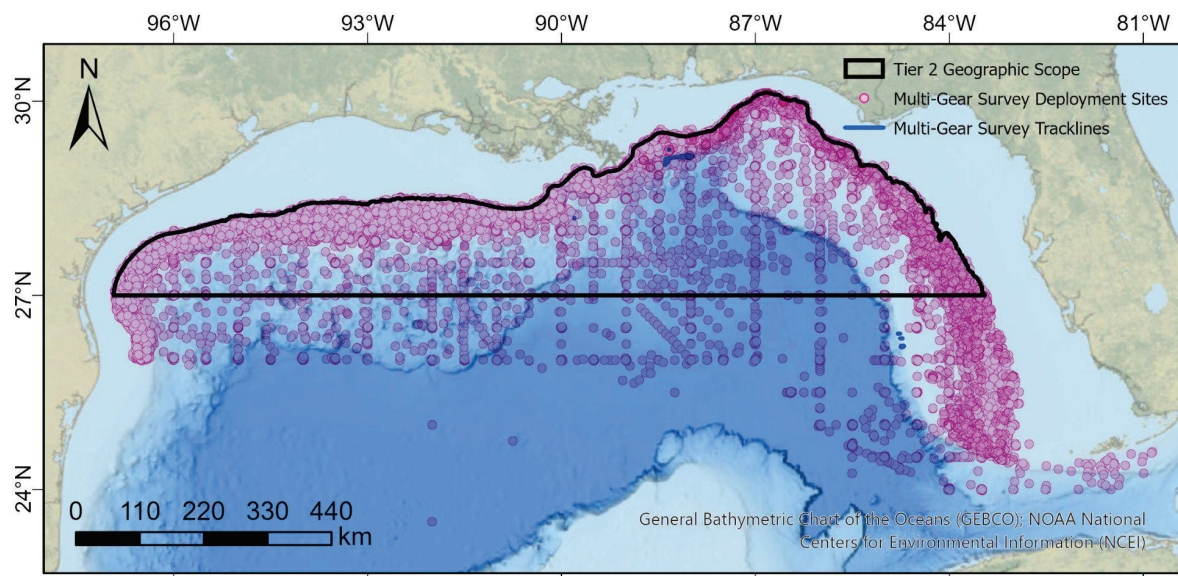


Figure 11. Survey sites per deployment of various gears, such as trawls, drop cameras, and hook and line gears used during various surveys collected for the fish ecology data inventory.

7.1.4. Data Sources

There were a variety of sources utilized to capture fish ecology datasets conducted in the GOM. These include NOAA, U.S. Geological Survey (USGS), Academia, Smithsonian Museum of Natural History, BOEM, and others. The majority of datasets were received from academic sources, with a total of 33 surveys (46.5% respectively; Table 8; Figure 12). Following that was a total of 25 surveys received from NOAA repositories (35.2% respectively; Table 8; Figure 12). USGS, BOEM, and the Smithsonian provided eight, two, and one dataset(s), respectively (11.3%, 2.8%, and 1.4% respectively; Table 8;

Figure 12). Other providers of datasets, such as state agencies, provided two datasets (2.8% respectively; Table 8; Figure 12).

Table 8. The total count as well as percentages were calculated per each source of surveys inventoried for fish ecology datasets. The other category indicates sources that were outside of those already listed, such as state agencies.

Source of Datasets	Number of Datasets	Percentage of Datasets
Academia	33	46.5 %
NOAA	25	35.2 %
USGS	8	11.3 %
BOEM	2	2.8 %
Smithsonian	1	1.4 %
Other	2	2.8 %

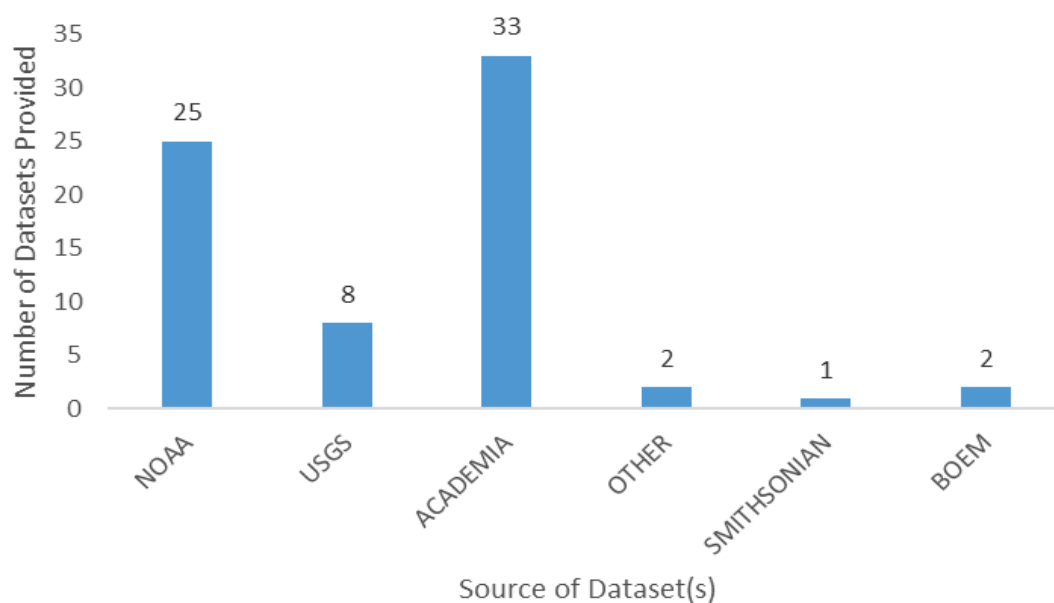


Figure 12. The total number of datasets provided by various sources for the fish ecology data inventory.

7.1.5. Data Collection per Decade

For this inventory, a temporal scope was not applied to the search for fish ecology datasets to be included. As this inventory concluded, there were no datasets conducted between 1980–1989 and 1990–1999 (Figure 13). There were 12 fish ecology surveys conducted during 2000–2009, 42 conducted during 2010 to 2019, and one survey conducted from 2020–2022 (Figure 13). There were 14 datasets that overlapped several decades, indicated as “Multi-Decade” in Figure13.

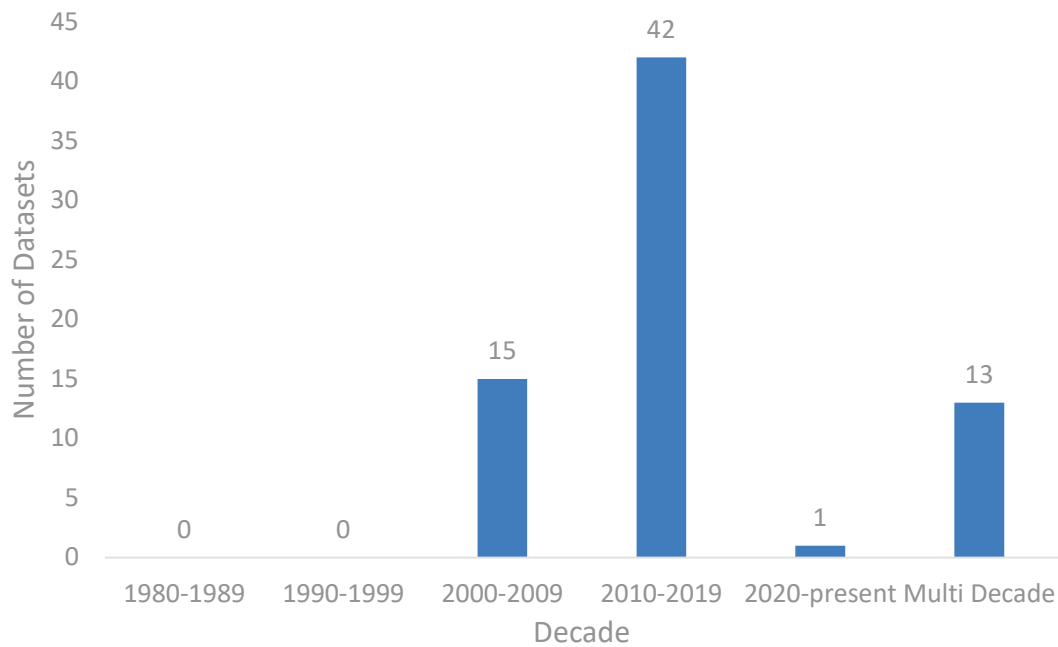


Figure 13. A plot of the total number of datasets surveyed during various decadal periods of time. Some surveys have been conducting work over several decades, SEAMAP for example, therefore are represented under the Multi-Decade category.

7.1.6. Data Collection by Depth

The majority of datasets collected for the fish ecology data inventory occurred in the mesophotic depth range, specifically from 50–300 m depth (Figure 14). There were 17 datasets that occurred in depths greater than 300 m and 21 surveys that were conducted in both mesophotic and deeper depths in the GOM (Figure 14).

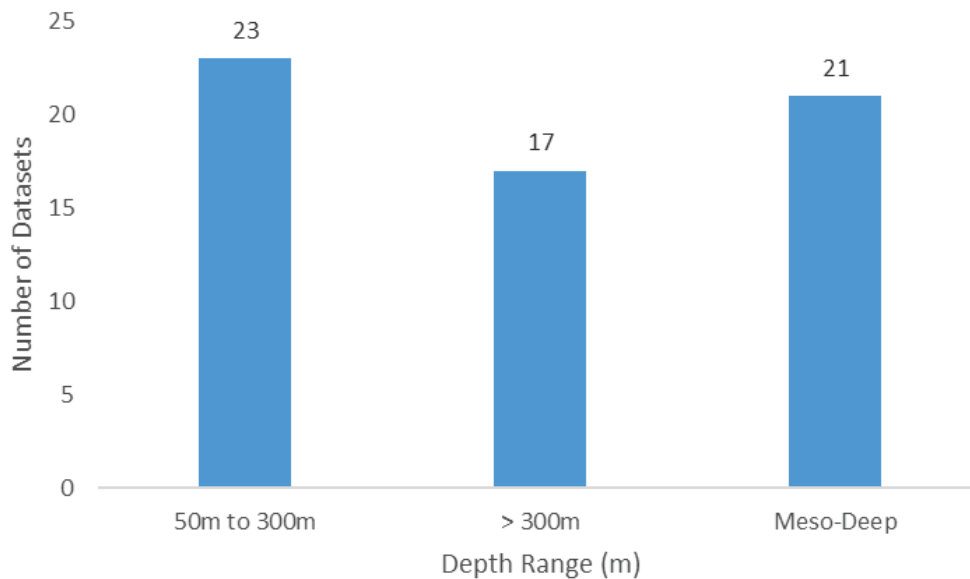


Figure 14. A plot of the total number of datasets compared to the depths in which they were surveyed. The “Meso-Deep” column were surveys that sampled in both the mesophotic and deep depth ranges.

7.2. Benthic Invertebrates

Fifty-two benthic invertebrate datasets were obtained, of which 17 were specific to benthic invertebrates and 35 were from other HAE working groups.

7.2.1. Data Types

Benthic invertebrate data types were mostly in the form of video or photographic stills, preserved samples or sediment samples. Other related data types are displayed in Figure 15.

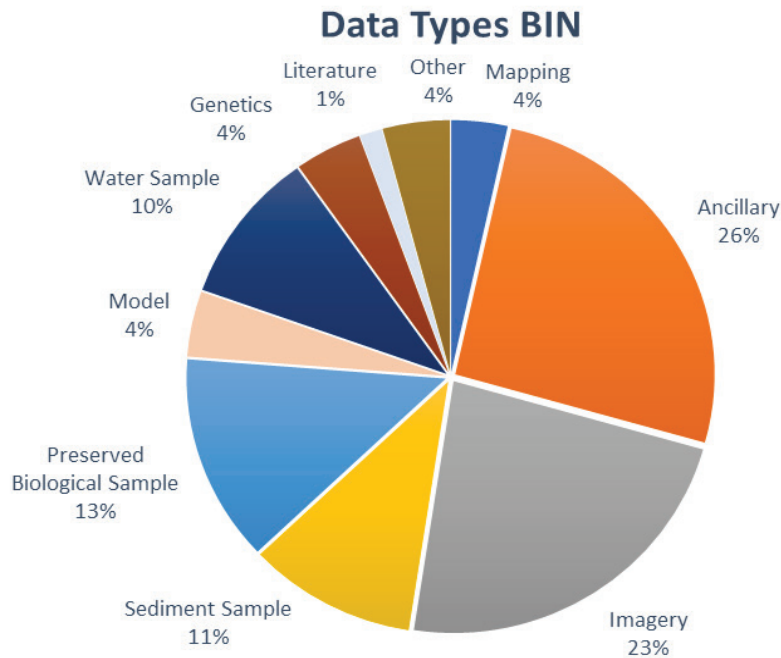


Figure 15. A pie chart displaying the percentages of data types in the datasets collected for the Benthic Invertebrates (BIN) group. Ancillary includes data from CTDs, fluorometers and ADCPs.

7.2.2. Data Sampling by Instrument

Benthic invertebrate data were mostly collected by ROV and AUV (imagery and samples). Other common instruments were trawls, traps, pushcores and drop cameras (Figure 16).

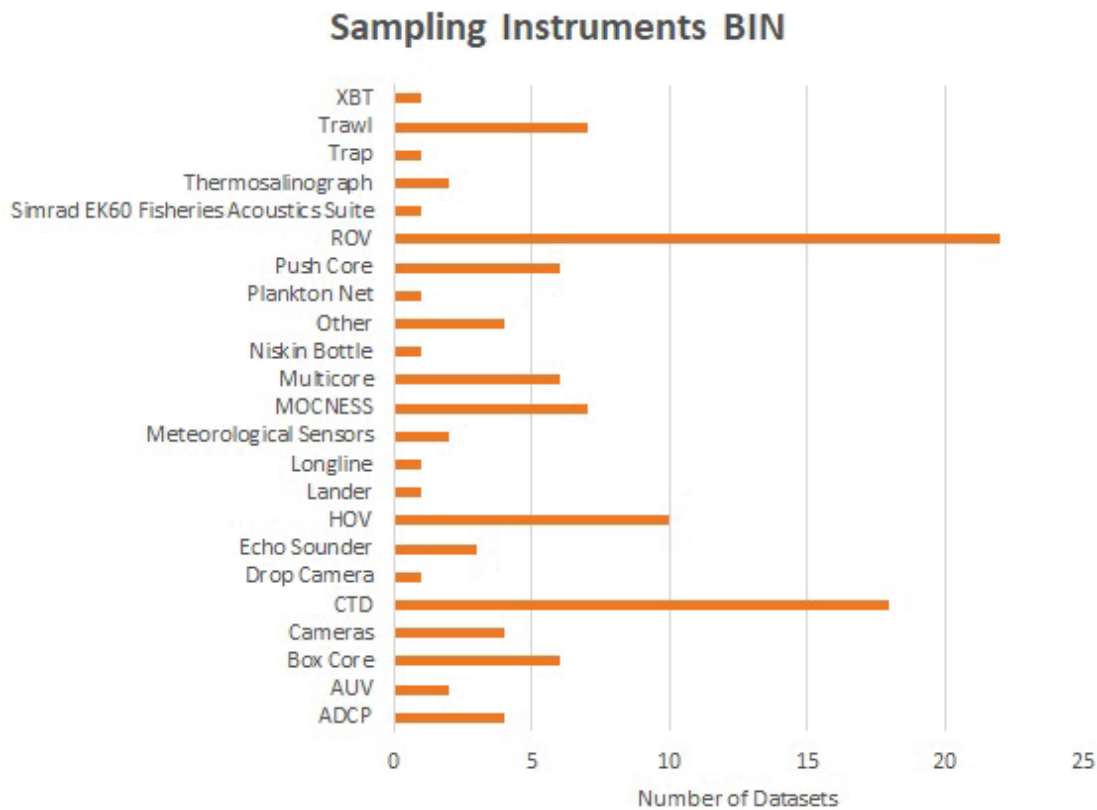


Figure 16. A bar graph with the number of benthic invertebrate (BIN) datasets that used the listed sampling instruments, sample collection techniques, and other associated techniques such as water quality and current measurements. “Other” includes Bushmasters, Radiometers, Remote Sensors and Stereo Baited Remote Underwater Video.

7.2.3. Spatial Coverage of Data

There was a total of 50 layers created from the 59 datasets compiled for the benthic invertebrate inventory which were depicted by the sampling methods conducted during those surveys (Table 9).

There were 26 layers from 35 datasets that utilized ROVs and HOVs to survey areas that were annotated for invertebrates and other attributes (Table 9; Figure 17). There were 13 datasets with layers that deployed core sampling gear and were annotated for invertebrates and other attributes (Table 9; Figure 18). There were seven datasets that utilized trawls and MOCNESS to collect invertebrates (Table 9; Figure 19). There were four datasets with layers that deployed various camera arrays (Table 9; Figure 20).

Table 9. The number of layers and datasets that were represented spatially in the benthic invertebrate data inventory by data type.

Spatial Data Type	Number of Layers	Number of Datasets
ROV/HOV dives	26	35
Core samples	13	13
Trawl and MOCNESS samples	7	7
Non-ROV underwater camera sites	4	4

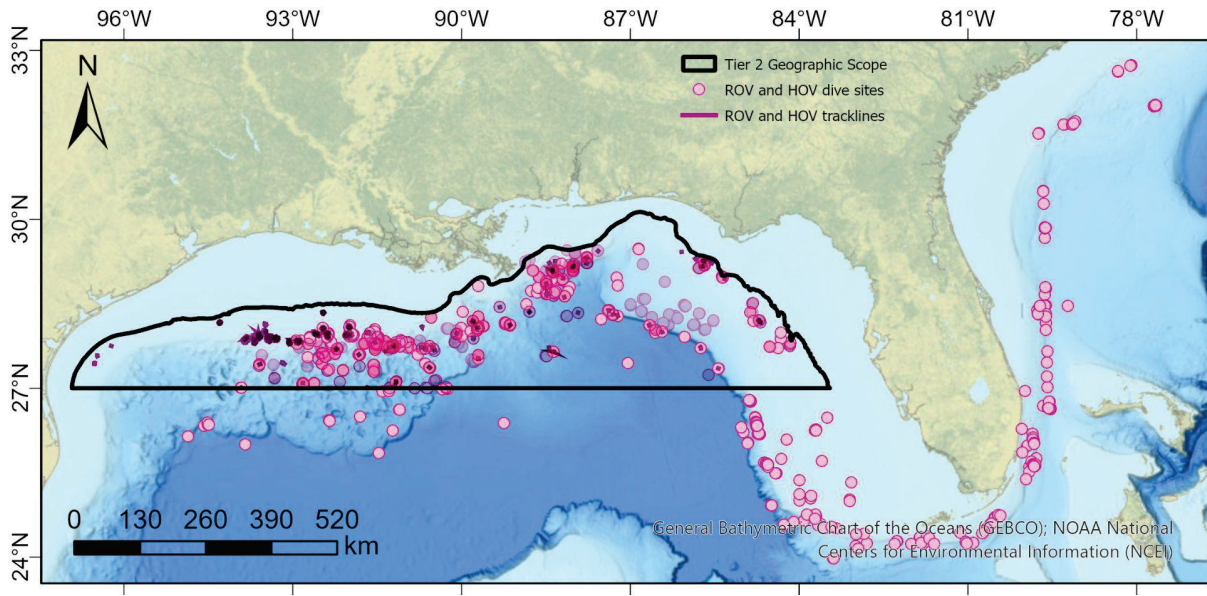


Figure 17. ROV and HOV dive sites within the datasets collected for the benthic invertebrate data inventory. The darker markers indicate dive tracklines.

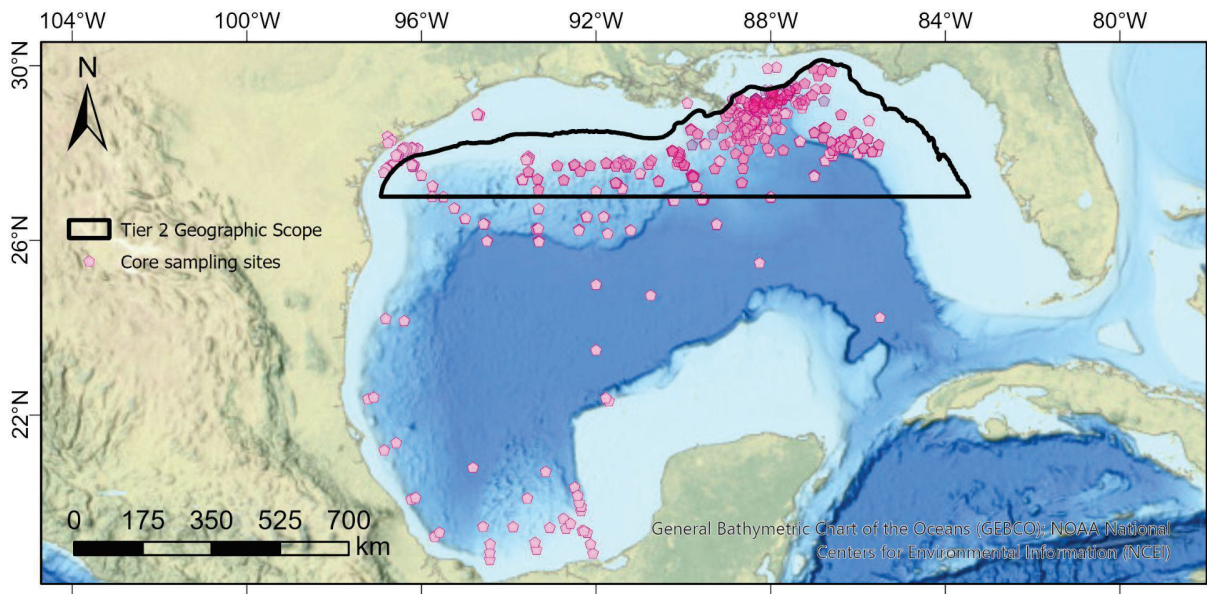


Figure 18. Core sample sites within the datasets collected for the benthic invertebrate data inventory.

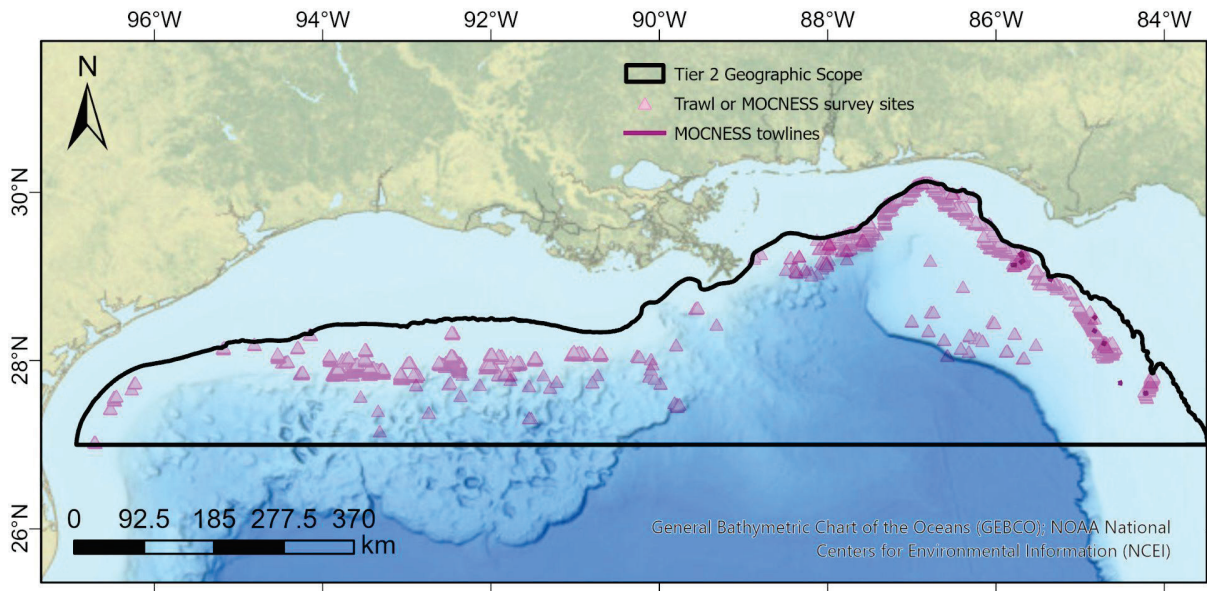


Figure 19. Trawl and MOCNESS sampling sites within the datasets collected for the benthic invertebrate data inventory. Darker markers indicate towlines.

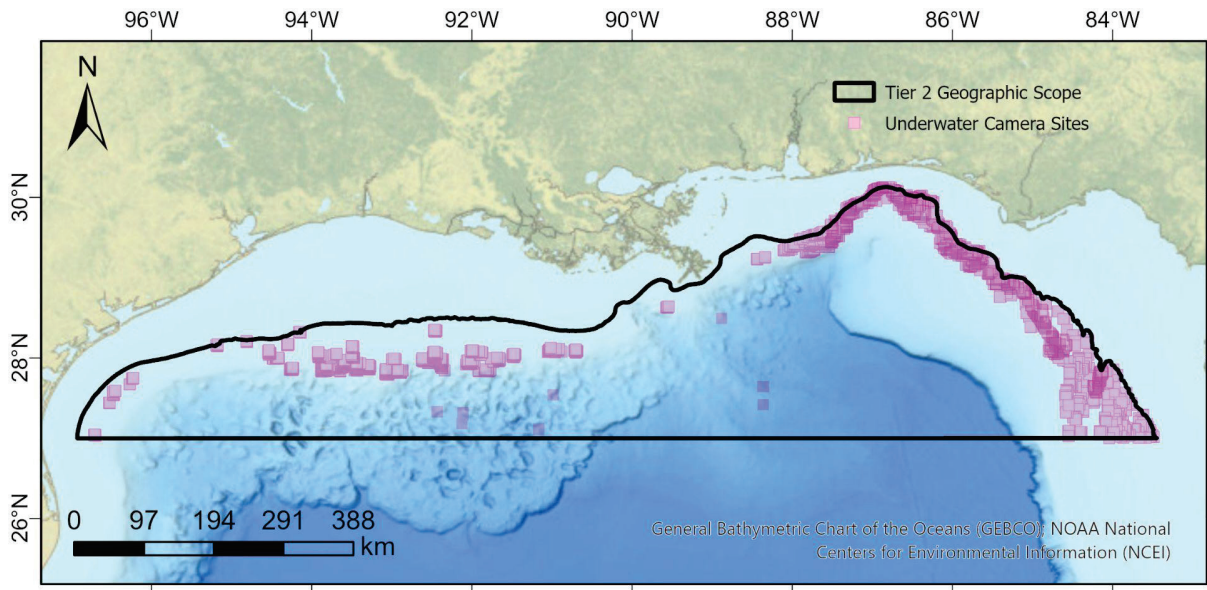


Figure 20. Non-ROV underwater camera sampling sites within the datasets collected for the benthic invertebrate data inventory.

7.2.4. Data Sources

There were a variety of sources utilized to capture benthic invertebrate datasets. These include NOAA, USGS, Academia, Smithsonian Museum of Natural History, BOEM and others. The majority of datasets were received from NOAA sources. Following that was 11 datasets received from “Other” sources, four datasets from “Academia” and USGS, and BOEM, and the Smithsonian providing two and one respectively. (Table 10; Figure 21).

Table 10. The total count as well as percentages were calculated for each source of datasets inventoried for benthic invertebrate datasets. The other category indicates sources that were outside of those already listed, such as state agencies.

Source of Datasets	Number of Datasets	Percentage of Datasets
NOAA	32	59.3%
OTHER	11	20.4%
ACADEMIA	4	7.4%
USGS	4	7.4%
BOEM	2	3.7%
Smithsonian	1	1.9%

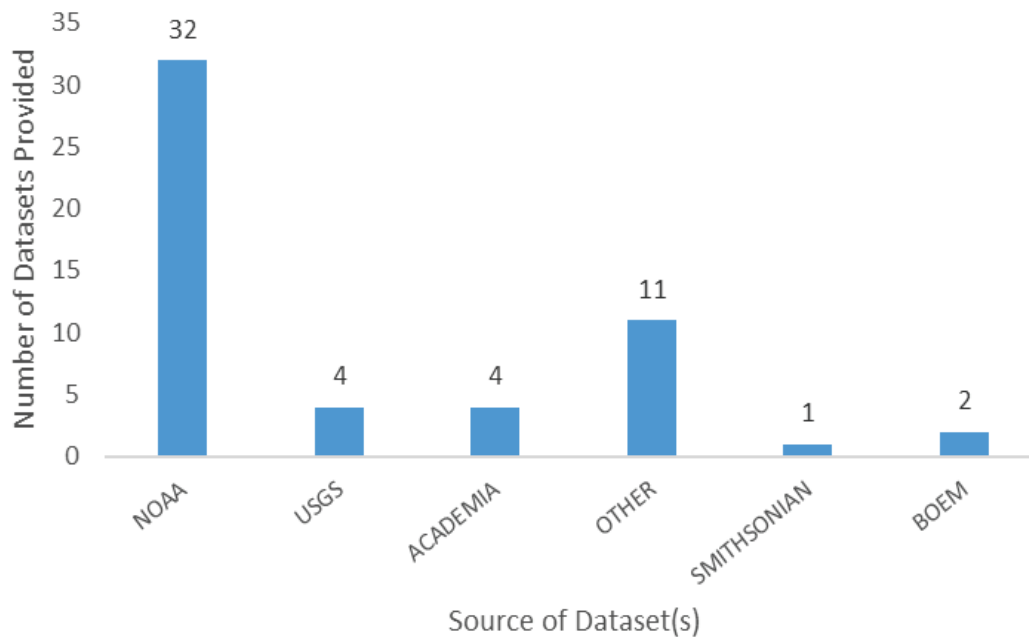


Figure 21. A plot of the total number of datasets provided by various sources for the benthic invertebrate data inventory.

7.2.5. Data Collection per Decade

For this inventory, a temporal scope was not applied to the search for benthic invertebrate datasets to be included. There were no datasets conducted between 1980–1989 or from 2020–present. There was one dataset from 1990–1999, 12 during 2000–2009, 26 conducted during 2010–2019, and 15 datasets that overlapped several decades, indicated as “Multi-Decade” in Figure 22.

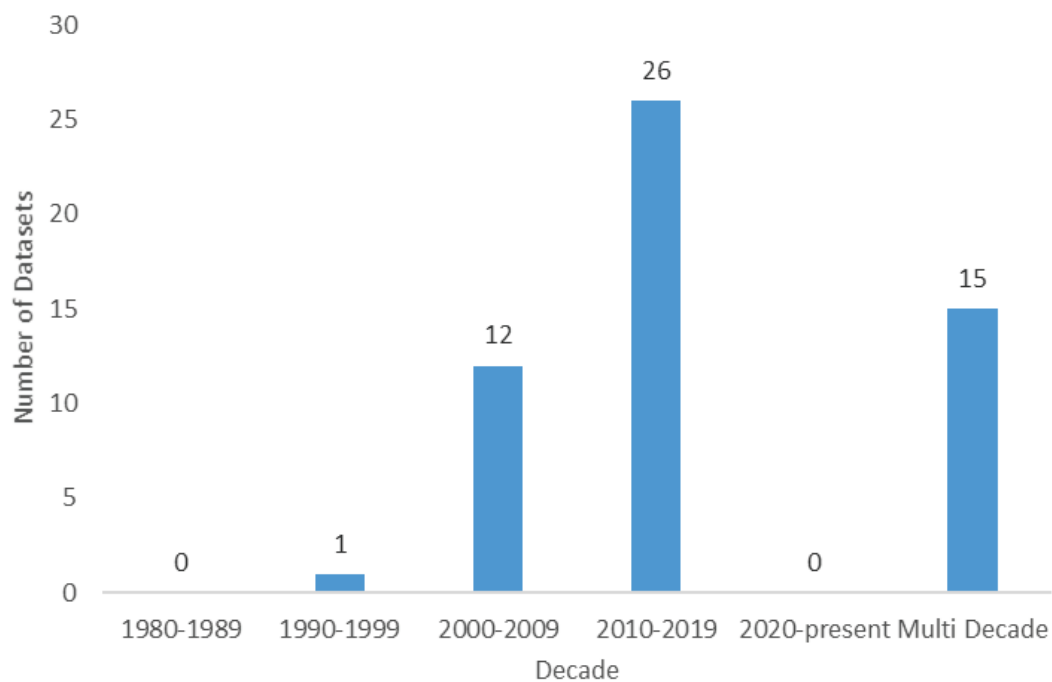


Figure 22. A plot of the total number of benthic invertebrate datasets surveyed during various decadal periods of time. Some surveys have been conducting work over several decades and are represented under the Multi-Decade category.

7.2.6. Data Collection by Depth

The majority of datasets collected for the benthic invertebrate data inventory occurred in the greater than 300 m depth range. There were 16 datasets that occurred in the 50–300 m range and 15 datasets that occurred in both mesophotic and deeper depths in the GOM (Figure 23).

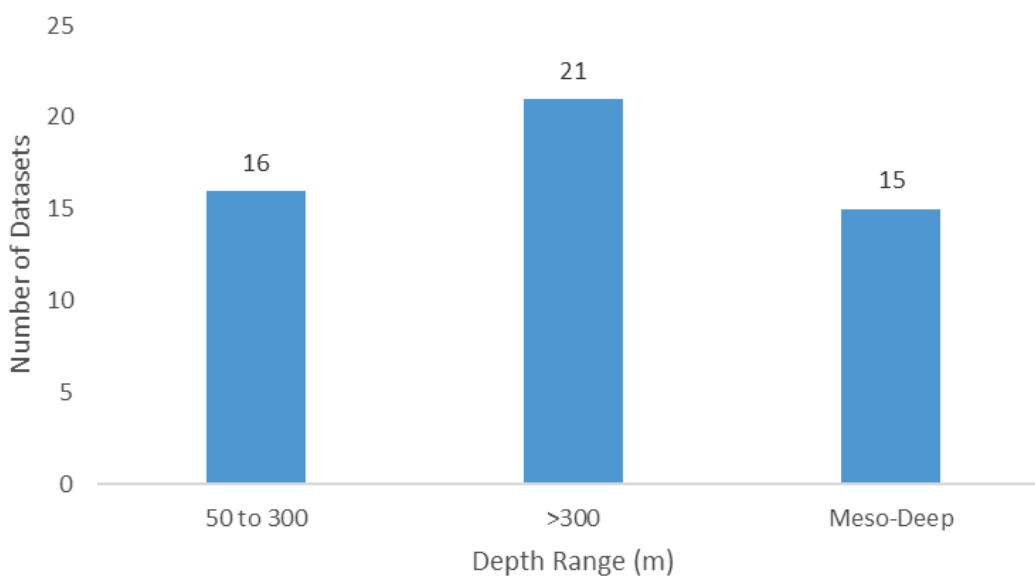


Figure 23. A plot of the total number of datasets compared to the depths in which they were surveyed. The “Meso-Deep” column were surveys that sampled in both the mesophotic and deep depth ranges.

7.3. Corals

Sixty-three coral datasets were obtained, of which eight were specific to corals and 55 were from other HAE groups.

7.3.1. Data Types

The 63 coral data sets contained a variety of data types (Figure 24). Imagery, typically from ROV surveys, accounted for almost a third while another third consists of ancillary data types (CTD, ADCP, etc).

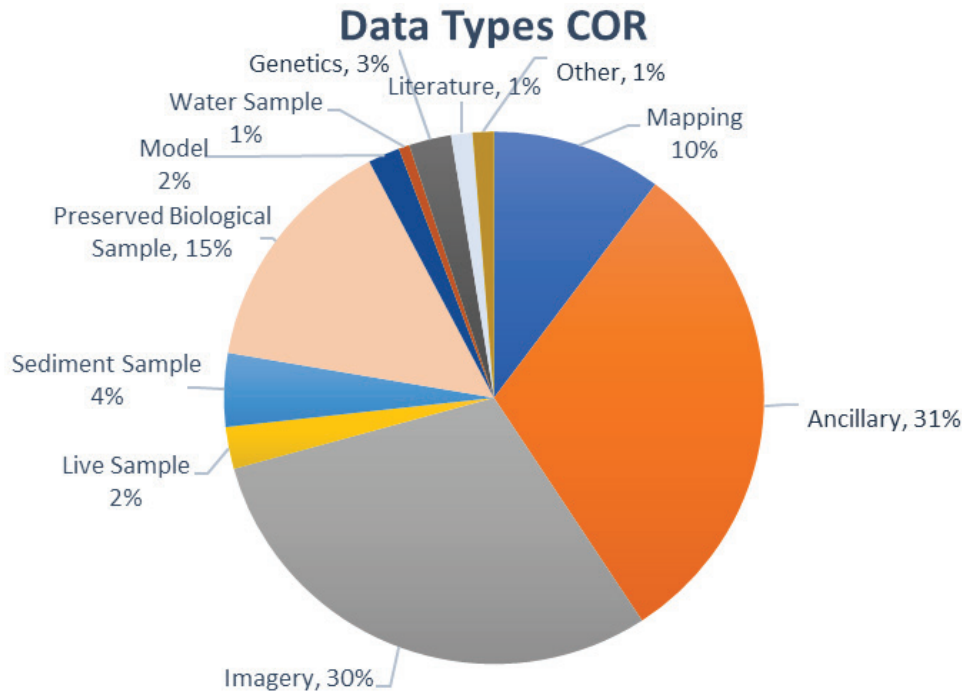


Figure 24. A pie chart displaying the percentages of data types in the datasets collected for the Coral (COR) group. Ancillary includes data from CTDs, fluorometers and ADCPs.

7.3.2. Data Sampling by Instrument

Coral data sets contained a variety of sampling instruments (Figure 25); most notably ROV and several different type of trawls. Several ancillary sampling instruments were identified including CTD and ADCP.

Sampling Instruments COR

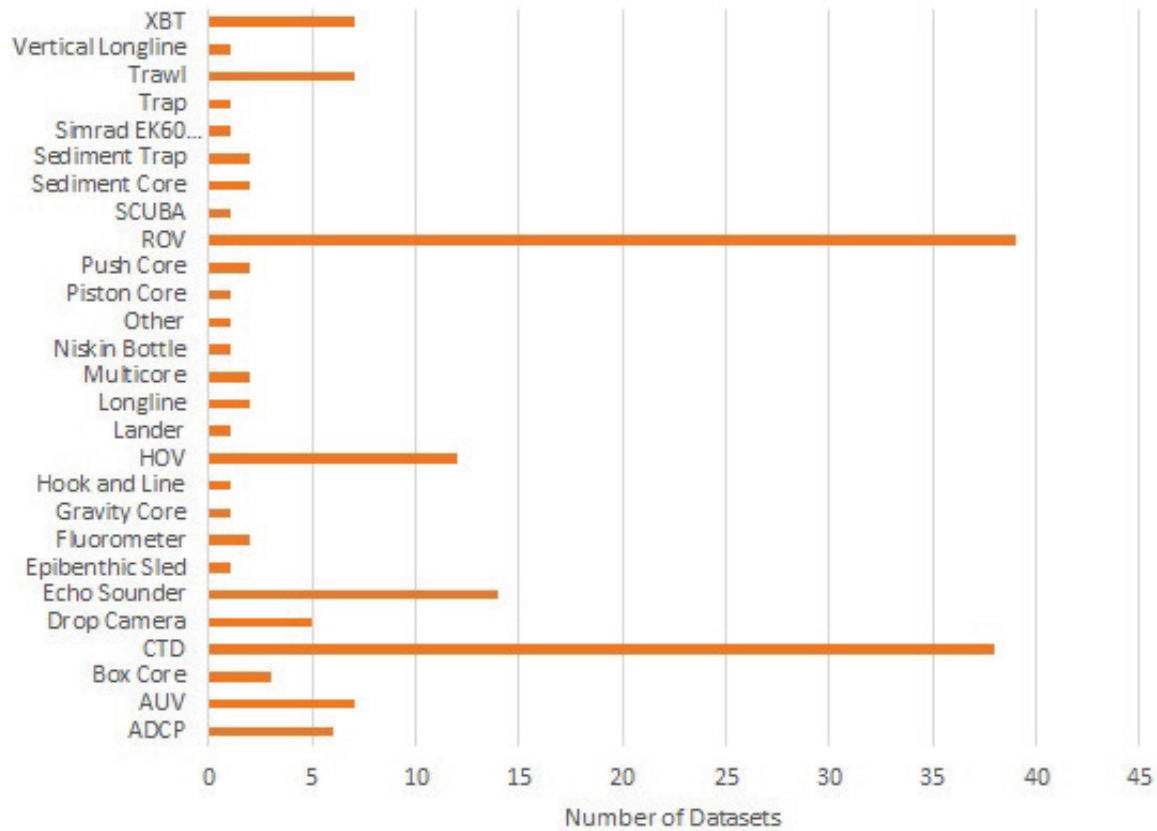


Figure 25. A bar graph displaying the number of coral (COR) datasets that used the listed sampling instruments and coral collection techniques, and other associated techniques such as water quality and current measurements. “Other” includes Bandit Reel.

7.3.3. Spatial Coverage of Data

A total of 46 layers were created from the 49 datasets compiled for the corals inventory which were depicted by the sampling methods conducted during those surveys (Table 11). There were 42 layers from 45 datasets that utilized ROVs and HOVs to survey areas that were annotated for corals and other attributes (Table 11; Figure 26). There were four datasets with layers that deployed various camera arrays (Table 11; Figure 27).

Table 11. The number of layers and datasets that were represented spatially in the coral data inventory by data type.

Spatial Data Type	Number of Layers	Number of Datasets
ROV/HOV Dives	42	45
Non-ROV Underwater Camera	4	4

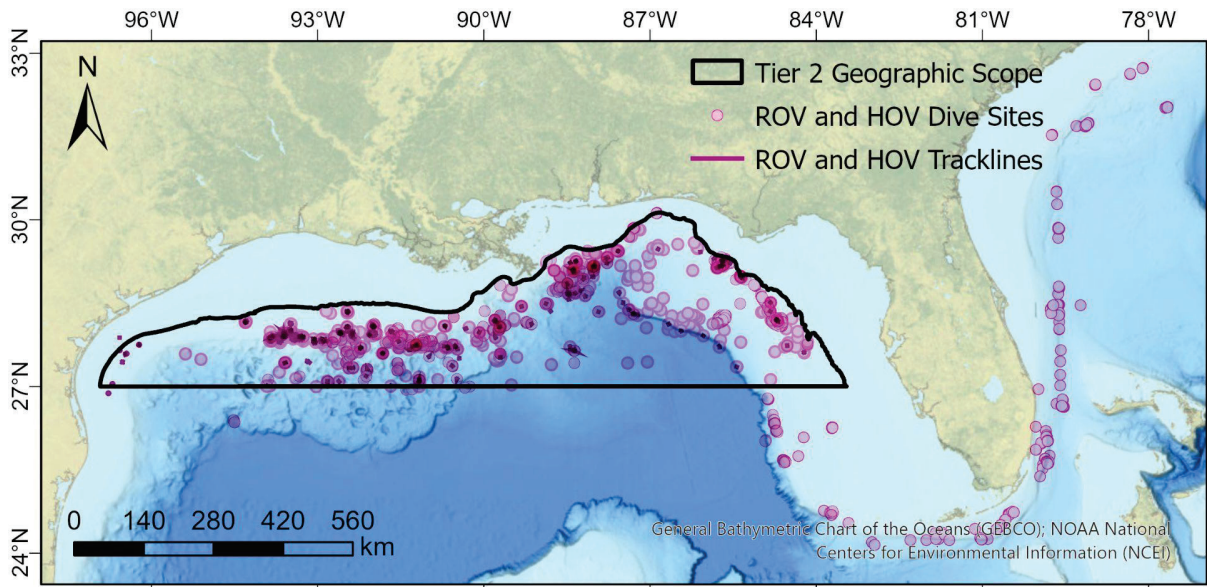


Figure 26. ROV and HOV dive sites within the datasets collected for the coral data inventory. The darker markers indicate dive tracklines.

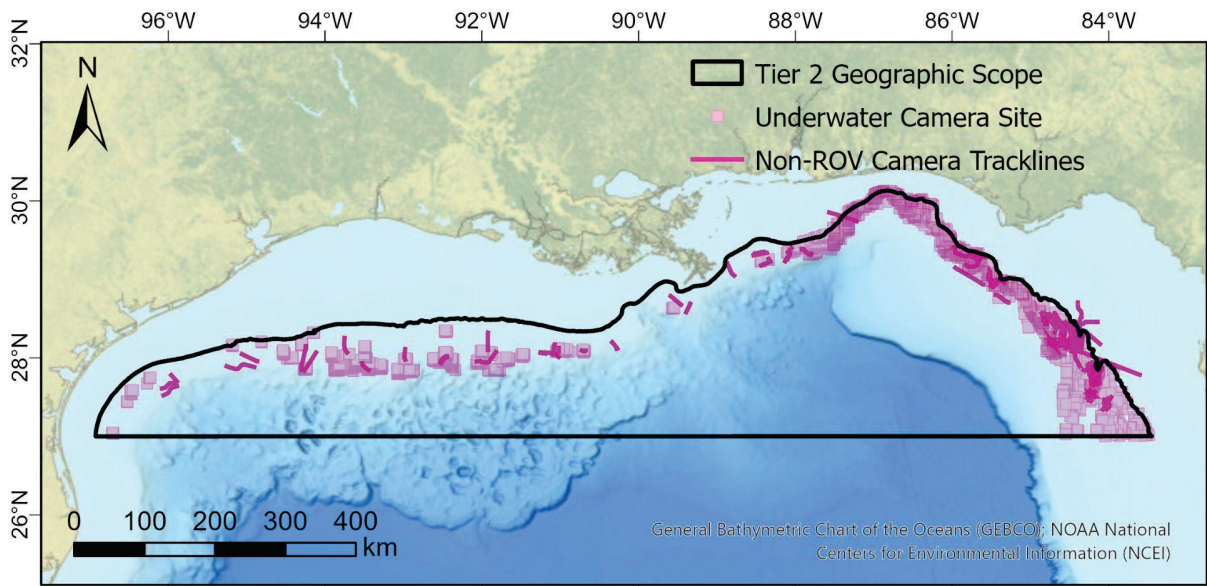


Figure 27. Non-ROV underwater camera sampling sites within the datasets collected for the coral data inventory.

7.3.4. Data Sources

There were a variety of sources utilized to capture coral datasets. These include NOAA, USGS, Academia, BOEM and others. The majority of datasets were received from NOAA sources. Following that were nine datasets received from “Other” sources, three datasets from “Academia”, and USGS and BOEM both providing one dataset each. (Table 12; Figure 28).

Table 12. The total count as well as percentages were calculated for each source of datasets inventoried for coral datasets. The other category indicates sources that were outside of those already listed, such as state agencies.

Source of Datasets	Number of Datasets	Percentage of Datasets
NOAA	48	77.4%
OTHER	9	14.5%
ACADEMIA	3	4.8%
USGS	1	1.6%
BOEM	1	1.6%

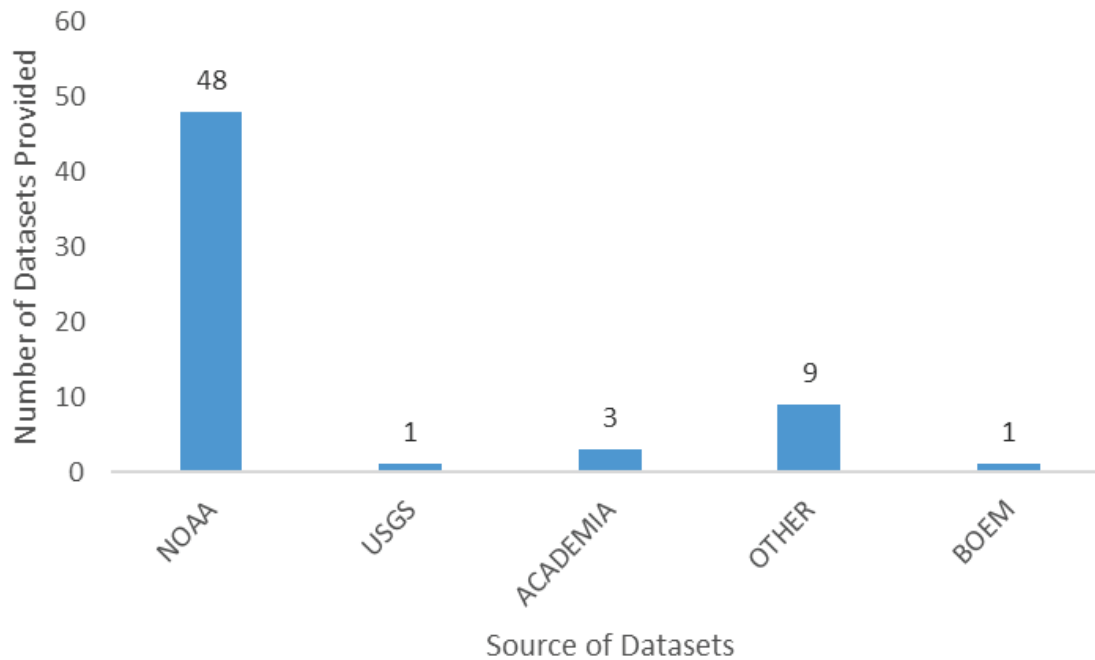


Figure 28. A plot of the total number of datasets provided by various sources for the coral data inventory.

7.3.5. Data Collection per Decade

For this inventory, a temporal scope was not applied to the search for coral datasets to be included. There were no datasets conducted between 1980–1989 or from 2020–present. There were four datasets from 1990–1999, 23 during 2000–2009, 32 conducted during 2010–2019, and 17 datasets that overlapped several decades, as indicated as Multi-Decade in Figure 29.

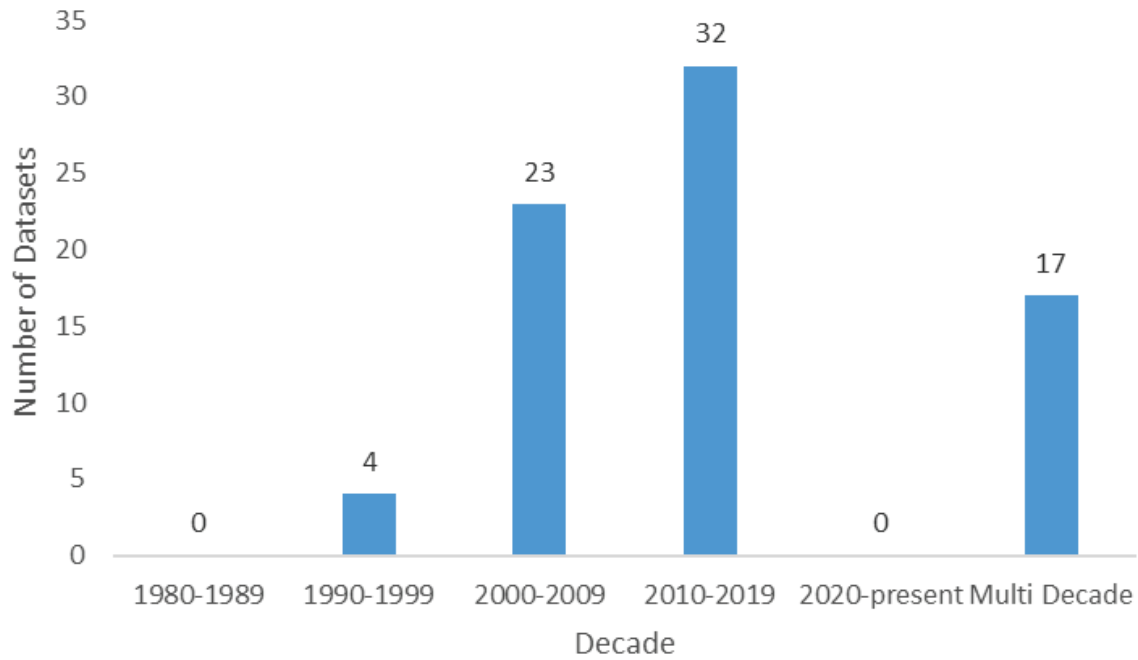


Figure 29. A plot of the total number of datasets surveyed during various decadal periods of time. Some surveys have been conducting work over several decades and are represented under the Multi-Decade category.

7.3.6. Data Collection by Depth

The majority of datasets collected for the corals data inventory occurred in the 50–300 m depth range. There were 18 datasets that occurred in the greater than 300 m range and 12 datasets that occurred in both mesophotic and deeper depths in the GOM (Figure 30).

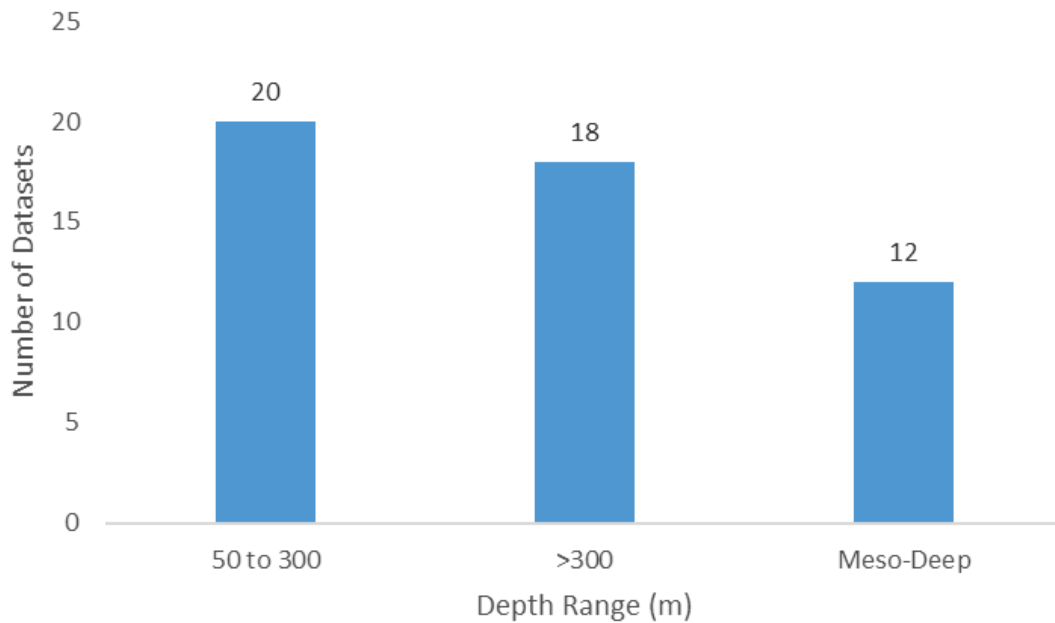


Figure 30. A plot of the total number of datasets compared to the depths in which they were surveyed. The “Meso-Deep” column were surveys that sampled in both the mesophotic and deep depth ranges.

7.4. Environmental and Oceanographic

A wide variety of data types associated with the environmental/oceanographic group were collected, with much of the data being ancillary data. A total of 184 environmental/oceanographic datasets were obtained, of which 67 were specific to environmental/oceanographic and 117 were from other HAE working groups.

7.4.1. Data Types

Aside from ancillary data types, the most frequent activities collecting environmental data were ROV surveys (imagery and preserved samples), sediment sampling and mapping (Figure 31).

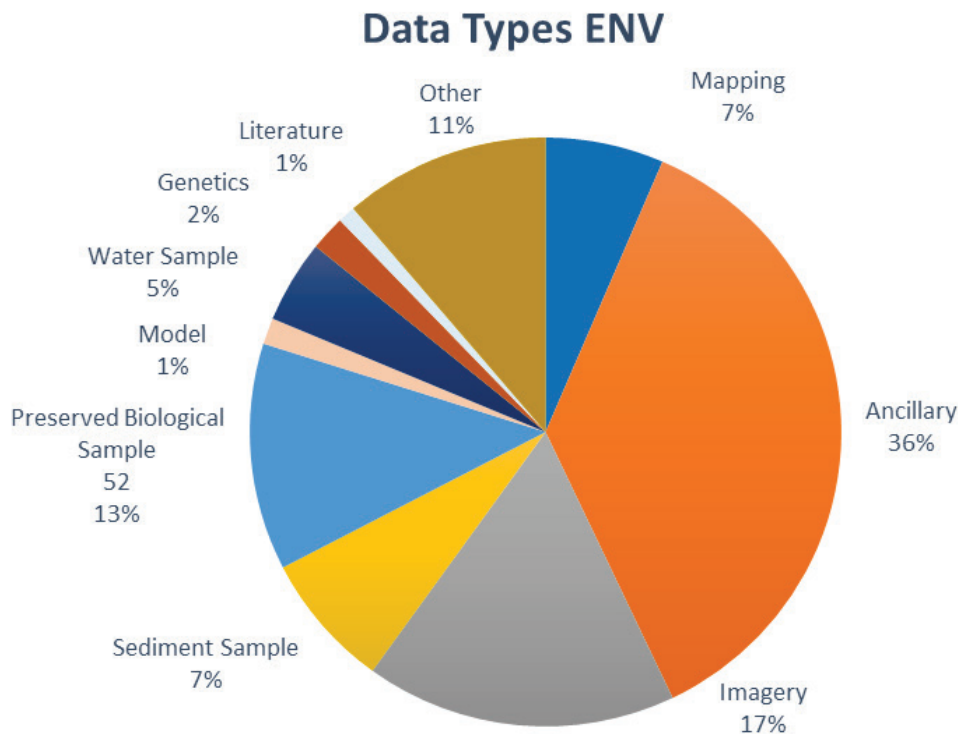


Figure 31. A pie chart displaying the percentages of data types in the datasets collected for the Environmental/Oceanographic (ENV) group. Ancillary includes data from CTDs, fluorometers and ADCPs.

7.4.2. Data Sampling by Instrument

Environmental sampling (Figure 32) was mostly conducted by CTD casts, sensors or CTDs on ROVs, trawls and Expendable Bathythermograph (XBT). Many other instruments and methods were used to collect environmental data.

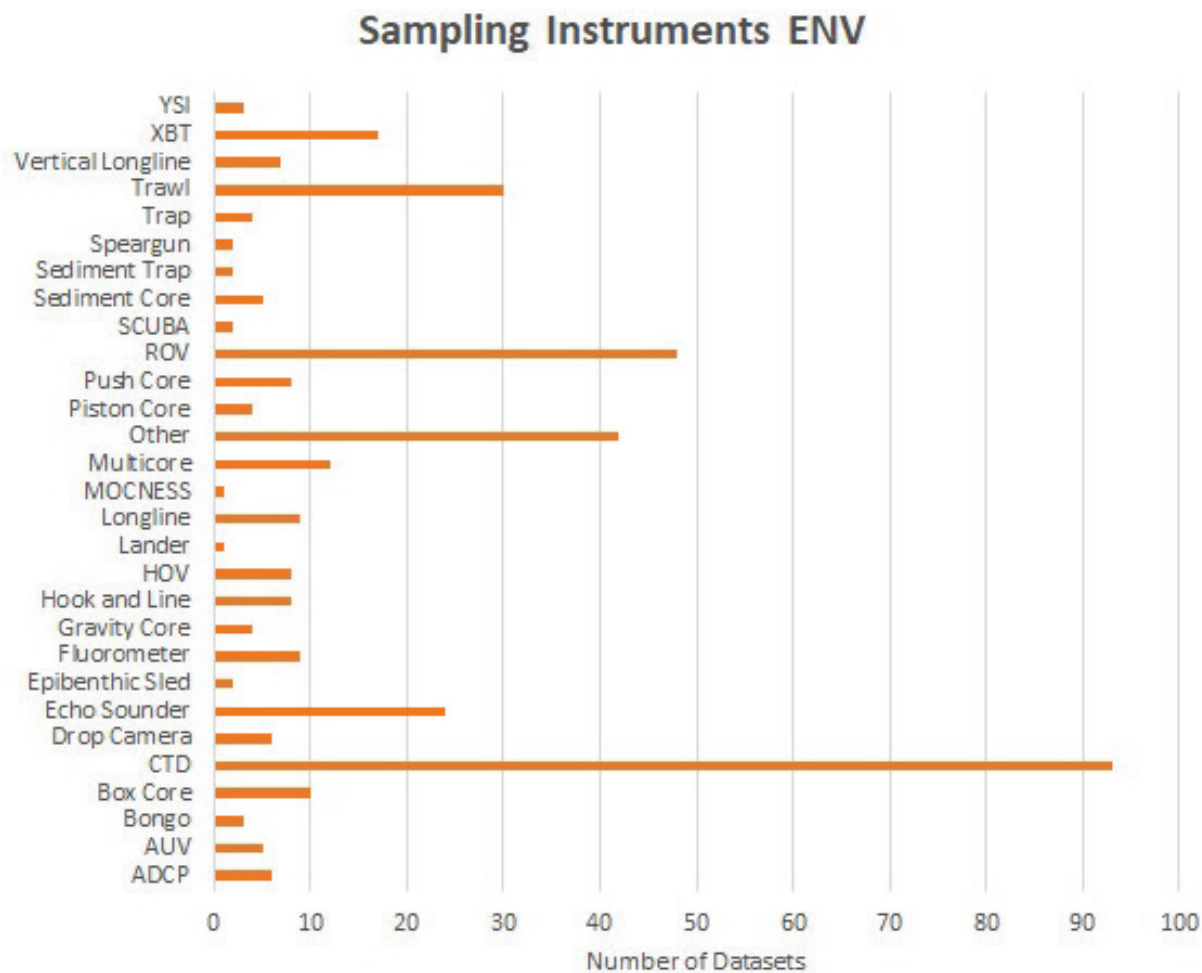


Figure 32. A bar graph displaying the number of environmental/oceanographic datasets that used the listed sampling instruments, sediment collection techniques, and other associated techniques such as water quality and current measurements. “Other” includes Bandit Reel, Block Array and Sediment Grab.

7.4.3. Spatial Coverage of Data

GIS layers for unique environmental/oceanographic datasets will be created in the near future and added to the HAE Geoplatform.

7.4.4. Data Sources

There were a variety of sources utilized to capture environmental/oceanographic datasets. These include NOAA, USGS, academic sources, BOEM, Smithsonian Museum of Natural History and others. The majority of datasets were received from NOAA sources. Following that were 66 datasets received from “Other” sources, 12 from USGS, 11 from “Academia”, and BOEM and Smithsonian both providing two datasets each. (Table 13; Figure 33).

Table 13. The total count as well as percentages were calculated for each source of datasets inventoried for environmental/oceanographic datasets. The other category indicates sources that were outside of those already listed, such as state agencies.

Source of Datasets	Number of Datasets	Percentage of Datasets
NOAA	91	49.5%
OTHER	66	35.9%
USGS	12	6.5%
ACADEMIA	11	6.0%
BOEM	2	1.1%
SMITHSONIAN	2	1.1%

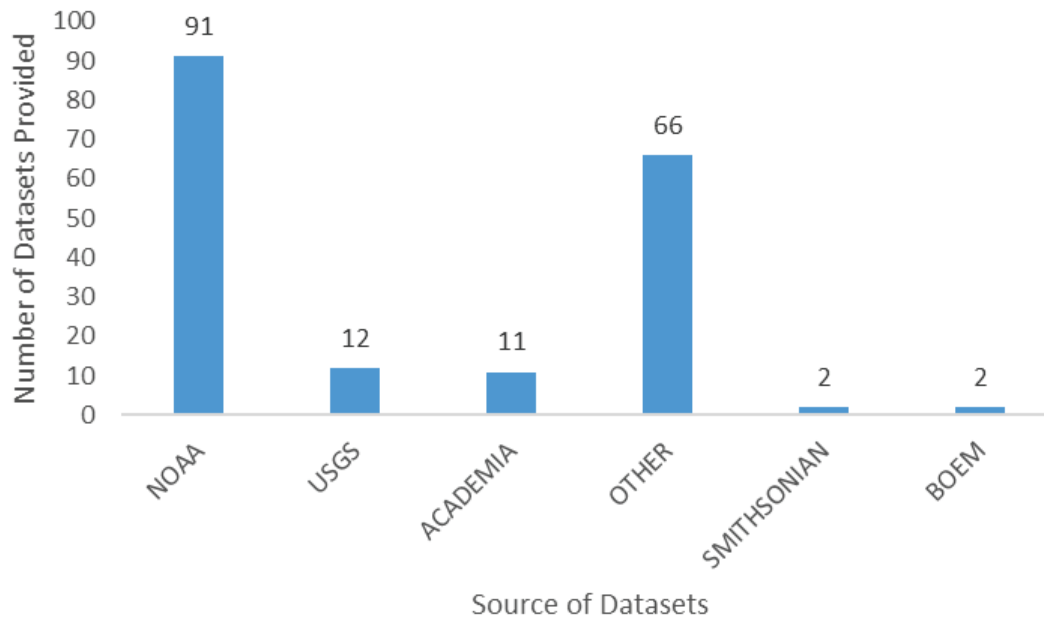


Figure 33. A plot of the total number of datasets provided by various sources for the environmental/oceanographic data inventory.

7.4.5. Data Collection per Decade

For this inventory, a temporal scope was not applied to the search for environmental/oceanographic datasets to be included. There were no datasets discovered between 1980–1989. There was one dataset identified from 2020–present. There were five datasets from 1990–1999, 25 during 2000–2009, 117 conducted during 2010–2019, and 33 datasets that overlapped several decades, as indicated as Multi-Decade in Figure 34.

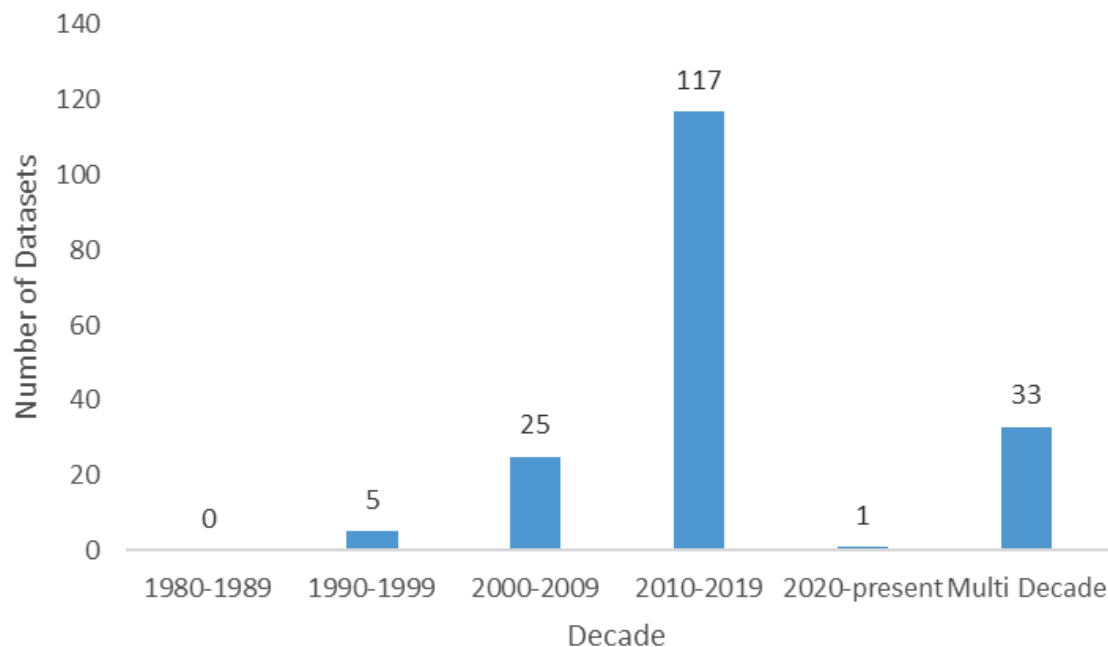


Figure 34. A plot of the total number of datasets surveyed during various decadal periods of time. Some surveys have been conducting work over several decades and are represented under the Multi-Decade category.

7.5. Genetics

The search for genetics data was different than for the other working groups. As a cross-project deliverable for the HAE and CPT projects, genetic data for all northern GOM deep-sea coral species was found using the National Center for Biotechnology Information (NCBI) GenBank databases. A spreadsheet which includes each species in rows, provides numbers of sequences associated with each species, and details on which molecular barcodes were available for each.

A total of 150 references were gathered related to the genetic data, including 100 peer-reviewed publications and 10 Bio-Projects (high-throughput sequencing e.g. RAD-seq). Over 3,000 mitochondrial and nuclear sequences were found, not including larger sequences used for single nucleotide polymorphisms (SNP)-calling (e.g. target capture, RAD-seq sequences). Of the 228 species researched, only 30 species had high-throughput sequences publicly available in the GenBank Sequence Read Archive. Table 14 provides a list of all taxa found to be missing publicly available genetic data in GenBank, and Figure 35 shows the percentage of each order with and without genetic data available in GenBank.

The following information was gathered for each deep-sea coral species:

- General information
 - Scientific Name
 - Order
 - Genus
 - Known depth zone (mesophotic, deep, or both)

- Mitochondrial data
 - Mitochondrial barcode availability (yes/no)
 - MutS sequence availability (yes/no)
 - COI sequence availability (yes/no)
 - Mitochondrial genome availability (yes/no)
 - Other mitochondrial barcodes available (names of genes, e.g. 16S, NAD5)
 - Number of mitochondrial sequences available in GenBank (total #)
 - References for mitochondrial data (citations for publications, direct submissions, or BioProjects)
- Nuclear data
 - Nuclear barcode availability (yes/no)
 - 28S sequence availability (yes/no)
 - Other nuclear barcodes available (names of genes, e.g. 18S, ITS1)
 - Number of nuclear sequences available in GenBank (total #)
 - References for nuclear data (citations for publications, direct submissions, or BioProjects)
- High-throughput sequence data (e.g. SNPs)
 - Sequence availability in GenBank Sequence Read Archive database (yes/no)
 - References (citations for publications, direct submissions, or BioProjects)

Table 14. Mesophotic and deep-sea coral taxa with no existing publicly available genetic data. Taxa are organized by taxonomic group: octocorals (pink), stony corals (blue), black corals (gray), and sea pens (yellow).

Group	Taxa name		
Octocorals	<i>Acanella aurelia</i> <i>Acanthogorgia schrammi</i> <i>Anthothela quattrinae</i> <i>Anthothela tropicalis</i> <i>Bebryce cinerea</i> <i>Bebryce grandis</i> <i>Bebryce parastellata</i> <i>Callipodium rubens</i> <i>Callogorgia linguimaris</i> <i>Chironephthya agassizii</i> (= <i>Siphonogorgia agassizii</i>) <i>Chrysogorgia desbonni</i> <i>Chrysogorgia elegans</i> <i>Chrysogorgia fewkesii</i> <i>Chrysogorgia multiflora</i> <i>Chrysogorgia spiculosa</i> <i>Ellisella atlantica</i> (= <i>Ctenocella (Viminella) atlantica</i>) <i>Ellisella funiculina</i>	<i>Epiphaxum breve</i> <i>Iridogorgia pourtalesii</i> <i>Leptogorgia cardinalis</i> <i>Leptogorgia medusa</i> <i>Metallogorgia</i> sp. j <i>Muriceides kukenthali</i> <i>Narella spectabilis</i> <i>Nicella deichmannae</i> <i>Nicella flagellum</i> <i>Nicella guadalupensis</i> <i>Nicella hebes</i> <i>Nicella robusta</i> <i>Nicella spicula</i> <i>Nidalia deichmannae</i> <i>Nidalia occidentalis</i> <i>Placogorgia mirabilis</i> <i>Placogorgia rudis</i> <i>Plumarella dichotoma</i> <i>Riisea paniculata</i>	<i>Scleracis petrosa</i> <i>Scleranthelia rugosa</i> var. <i>musiva</i> <i>Scleranthelia rugosa</i> var. <i>rugosa</i> <i>Stenisis humilis</i> <i>Stereotelesto corallina</i> <i>Telesto flavula</i> <i>Telesto fruticulosa</i> <i>Telesto sanguinea</i> <i>Thesea citrina</i> <i>Thesea grandiflora</i> <i>Thesea granulosa</i> <i>Thesea guadalupensis</i> <i>Thesea hebes</i> <i>Thesea nivea</i> <i>Thesea parviflora</i> <i>Thesea rubra</i> <i>Thesea rugosa</i> <i>Thesea solitaria</i> <i>Villogorgia nigrescens</i>
Stony corals	<i>Anomocora fecunda</i> <i>Anomocora marchadi</i> <i>Anomocora prolifera</i> <i>Astrangia solitaria</i> <i>Balanophyllia floridana</i> <i>Balanophyllia palifera</i> <i>Bathypsammia tintinnabulum</i> <i>Caryophyllia ambrosia caribbeana</i> <i>Caryophyllia antillarum</i> <i>Caryophyllia barbadensis</i> <i>Caryophyllia berteriana</i> <i>Caryophyllia horologium</i> <i>Caryophyllia polygona</i> <i>Cladocora debilis</i> <i>Coenocyathus caribbeana</i> <i>Coenocyathus parvulus</i> <i>Concentrotheca laevigata</i>	<i>Madrepora carolina</i> <i>Dasmosmilia variegata</i> <i>Deltocyathoides stimpsonii</i> (= <i>Peponocyathus stimpsonii</i>) <i>Deltocyathus calcar</i> <i>Deltocyathus eccentricus</i> <i>Deltocyathus italicus</i> <i>Eguchipsammia cornucopia</i> (= <i>Dendrophyllia cornucopia</i>) <i>Flabellum (Ulocyathus) moseleyi</i> <i>Flabellum floridanum</i> (= <i>Flabellum fragile</i>) <i>Fungiacyathus (Bathyactis) crispus</i> <i>Labyrinthocyathus langae</i> <i>Madracis asperula</i> <i>Madracis brueggemanni</i>	<i>Oculina tenella</i> <i>Oxysmilia rotundiflora</i> <i>Phacelocyathus flos</i> <i>Phyllangia americana americana</i> <i>Phyllangia pequegnatae</i> <i>Polycyathus senegalensis</i> <i>Polymyces fragilis</i> (= <i>Rhizotrochus fragilis</i>) <i>Pourtalocyathus hispidus</i> <i>Pourtalosmilia conferta</i> <i>Premocyathus cornuformis</i> <i>Rhizopsammia goesi</i> <i>Schizocyathus fissilis</i> <i>Stephanocyathus paliferus</i> <i>Tethocyathus cylindraceus</i> <i>Trochocyathus rawsonii</i>
Black corals	<i>Acanthopathes humilis</i> <i>Antipathes gracilis</i> <i>Antipathes lenta</i> <i>Aphanipathes salix</i>	<i>Parantipathes tetrasticha</i> <i>Stichopathes pourtalesi</i> <i>Stylopathes Americana</i> (= <i>Antipathes americana</i>)	<i>Stylopathes columnaris</i> <i>Tanacetipathes hirta</i>
Sea pens	<i>Acanthoptilum agassizii</i> <i>Acanthoptilum oligacis</i>	<i>Protoptilum thomsoni</i> <i>Stylatula antillarum</i>	<i>Umbellula guentheri</i> <i>Virgularia presbytes</i>

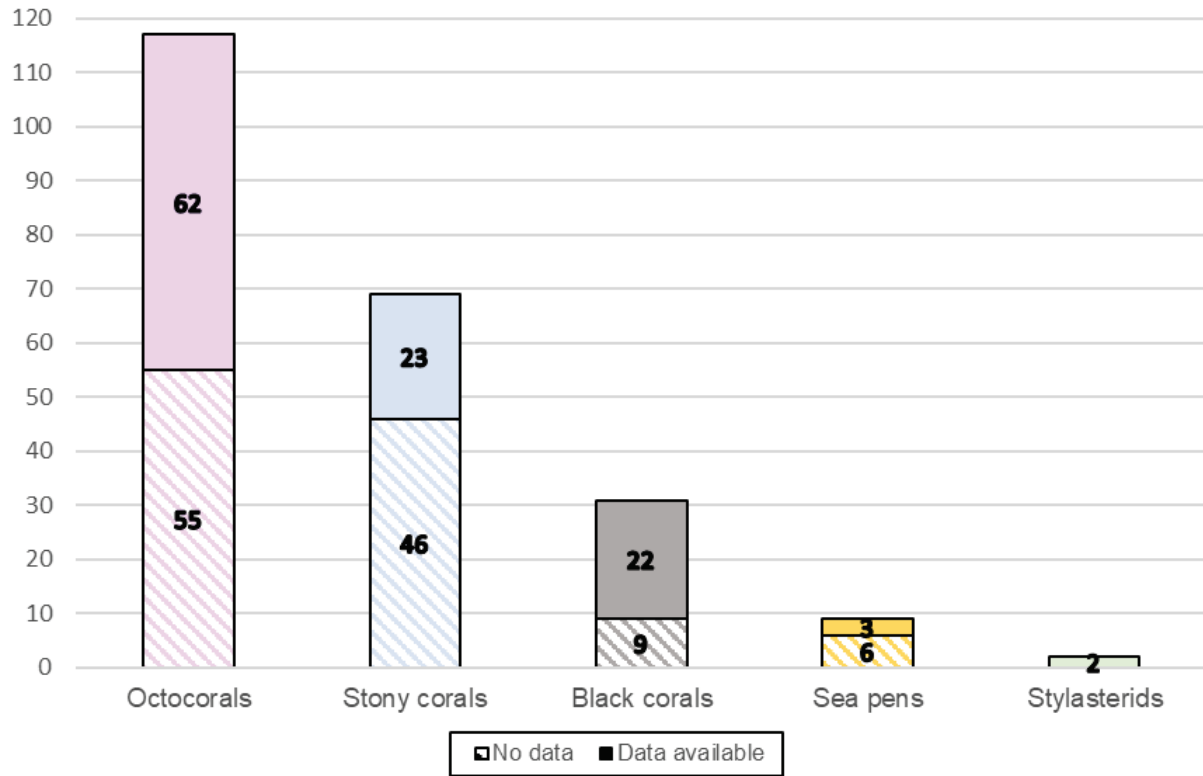


Figure 35. Total number of deep-sea coral species with genetic data available (solid fill) and no data available (pattern fill), organized by scientific taxonomic group.

8. Conclusions

The literature search and external engagement involved a large time investment and facilitated discovery of some datasets but most had already been discovered from broad sources, like NCEI archives or data portals. For much of the literature it was difficult to acquire processed annotations as most datasets included spatial data, real-time annotations, still images or video, tissue samples or visual representations of the data (i.e., charts and graphs).

There are large parts of the study area with little to no data in the inventory, particularly when datasets are subset by technology, time period, etc. The vast amount of data to be discovered necessitates more time and, if possible, more team members. The inventory provides a framework for identifying these data gaps and directing future data collection and literature reviewing accordingly.

This MDBC HAE inventory is the most comprehensive assembly of biological and oceanographic data for the northern GOM. Not only does this information support DWH restoration activities but can be a valuable resource for academic and management entities. The datasets (GIS shapefiles) and their metadata will be archived at NOAA NCEI and the data inventory catalog will be archived in the NOAA DIVER MDBC repository for public accessibility.

As previously stated, the inventory will be a continuous activity throughout project implementation. Data generated by MDBC as well as data generated by external sources will be added as described in the accretion process. There is still a large amount of literature and data that need to be evaluated. The current inventory provides the foundation for gap analysis and operations planning for the MDBC projects. The gap analysis will assist the HAE group in determining where to focus our efforts for future literature searches and data gathering which will aid in planning operations and project activities.

Literature Cited

- BOEM. 2021. Biological Environmental Background Report for the Gulf of Mexico OCS Region. Bureau of Ocean Energy Management, Gulf of Mexico Regional Office. OCS Report BOEM 2021-15. 283 pp.
- Deepwater Horizon Natural Resource Damage Assessment Trustees (DWH NRDA Trustees). 2016. Deepwater Horizon oil spill: Final Programmatic Damage Assessment and Restoration Plan and Final Programmatic Environmental Impact Statement. Available at <http://www.gulfspillrestoration.noaa.gov/restoration-planning/gulf-plan> (Accessed 14 February 2023)
- Etnoyer, P.J., A. Shuler, and S.D. Cairns. 2020. Deep-Sea Coral Taxa in the U.S. Gulf of Mexico: Depth and Geographical Distribution (v. 2020). 14 pp. Available at <https://deepseacoraldata.noaa.gov/library/2020-regional-deep-sea-coral-species-list> (Accessed 14 February 2023)
- Frometa, J., P.J. Etnoyer, A.M. Quattrini, S. Herrera, and T.W. Greig. 2021. Genetic Divergence and Polyphyly in the Octocoral Genus *Swiftia* [Cnidaria: Octocorallia], Including a Species Impacted by the DWH Oil Spill. *Diversity* 13(4): 172. <https://doi.org/10.3390/d13040172>
- Hourigan, T.F., P.J. Etnoyer, R.P. McGuinn, C. Whitmire, D.S. Dorfman, M. Dornback, S. Cross, and D. Sallis. 2015. An Introduction to NOAA's National Database for Deep Sea Corals and Sponges. NOAA Technical Memorandum NOS NCCOS 191. Silver Spring, MD. 27 pp.
- Kahng, S.E., J.R. Garcia-Sais, H.L. Spalding, E. Brokovich, D. Wagner, E. Weil, L. Hinderstein, R.J. Toonen. 2010. Community ecology of mesophotic coral reef systems. *Coral Reefs* 29(2): 255-275.
- Minerals Management Service (MMS). 2009. Notice To Lessees No. 2009-G39 Biologically-Sensitive Underwater Features and Areas. Effective Date: January 27, 2010 Expiration Date: January 26, 2015. New Orleans, LA. Available at <https://www.boem.gov/sites/default/files/regulations/Notices-To-Lessees/2009/09-G39.pdf> (Accessed 14 February 2023)
- Open Ocean Trustee Implementation Group (OOTIG). (2019). Deepwater Horizon Oil Spill Natural Resource Damage Assessment, Open Ocean Trustee Implementation Group, Draft Restoration Plan 2/ Environmental Assessment: Fish, Sea Turtles, Marine Mammals, and Mesophotic and Deep Benthic Communities.
- Puglise, K.A., L.M. Hinderstein, J.C.A. Marr, M.J. Dowgiallo, and F.A. Martinez. 2009. Mesophotic Coral Ecosystems Research Strategy: International Workshop to Prioritize Research and Management Needs for Mesophotic Coral Ecosystems, Jupiter, Florida, 12-15 July 2008. NOAA Technical Memorandum NOS NCCOS 98 and OAR OER 2. Silver Spring, MD. 24 pp.
- Pyle, R.I. 1996. The twilight zone. *Natural History* 105: 59-62.

Sulak, K.J., and P.M. Dixon. 2015. Change in Gulf of Mexico mesophotic reef fish community structure across the time threshold of the Deepwater Horizon oil spill event in 2010: Mesophotic reef fish community impacts of the DWH oil spill. DWH Deepwater Benthic NRDA Technical Working Group Report.

The EndNote Team. 2020. EndNote™ (Version EndNote 20) [64 bit]. Philadelphia, PA: Clarivate Analytics.

Appendices

Appendix A: Data Characterization Attributes

Table A1. Cross-project Attribute Table

Attribute	Type	Description
MDBC_ID	core	unique identifier for dataset in the context of MDBC. Three letter working group acronym followed by 6 digits
child_ID	core	unique identifier for child of dataset, referenced to the MDBC_ID. Prefix is MDBC_ID, followed by "C" and a two-digit numeral. For example, if MAP000001 is the parent, then MAP000001C01 is the first child of that parent. A child always consists of a subset of a parent.
orig_ID	core	If a dataset is already entered in catalog and will be used by another team, then it should be referenced in this column on the second team's catalog by the MDBC_ID from the original entry.
data_type	core	MDBC controlled vocabulary of broad data types: (Mapping; Ancillary; Imagery; Live Sample; Sediment Sample; Preserved Biological Sample; Model; Water Sample; Genetics; Literature; Other) Can be semicolon separated list if more than one type applies.
sub_type	core	MDBC controlled vocabulary of more specific data types. The complete list can be viewed through the dropdown menu in the sub_type column in the catalog. Can be semicolon separated list if more than one type applies.
MDBC_reviewer	core	MDBC team member who vetted dataset
PCS	core	Did it Pass Coarse Screening (PCS) Attributes? Yes/No/In Review
PFS	core	Did it Pass Fine Screening (PFS) Attribute? Yes/No/In Review
last_modified	core	last time record was updated. YYYY-MM-DD. Automatically updated when row is modified.
date_created	core	date entry was created in data catalog. YYYY-MM-DD
MDBC_data_location	core	location of data within MDBC work space. Can have multiple links, but please include filepath to Google Drive ingest folder as one link (so it is available to entire portfolio). Can also be a location (e.g. hard drives in Steve Formel's office), NA if the data was not ingested.
archive_status	core	Has this data been archived at a bonafide archive like NCEI/USGS/NCBI. Yes, No, "In Prep" if we have plans to archive it
comments	core	Open field for comments about dataset
geo_keywords	core	controlled vocabulary of geographic keywords to give a sense of location for dataset. See image of map in Dictionary. Can select multiple keywords. (Atlantic Ocean, Edisto Reef, Gulf of Mexico, Eastern Gulf of Mexico, Central Gulf of Mexico, Western Gulf of Mexico, MDBC Footprint, South of 27N, 30m-300m, 301m-800m, 801m-2500m). The MDBC footprint is bounded by a 50m isobath in the north and 27N latitude on the south.
WKT	core	Well-known text markup, for describing footprints. See wikipedia for more info.
title	core	Title of dataset, your title should include one or two primary variables, not more than one or two ship/platform names, the sea area and date range when data were collected. A formal ISO field that can be scraped from the XML metadata at NCEI.
abstract	core	Abstract of dataset (a dataset description) A formal ISO field that can be scraped from the XML metadata at NCEI.
start_date	core	start date of data collection YYYY-MM-DD
end_date	core	end date of data collection YYYY-MM-DD

keywords	core	open field for keywords of any type. please separate with a semicolon
orig_data_location	core	Location from which data was originally sourced. Can be a link to where original data can be downloaded, landing page
related_references	core	a semi-colon separated listing of relevant existing reports, publications, user guides, web sites, or other supporting docs
repository_name	core	Name of repository from which data or sample was sourced.
identifier	core	equivalent to DOI, Accession number. Must include type in identifier_type column (DOI, NECI Accessions, NGDC ID, etc.)
identifier_type	core	Type of identifier: Digital Object Identifier (DOI), NCEI Accession, NGDC Accession, etc. If unsure with ISO metadata at NCEI use text inside of <gmi:MI_Metadata><gmd:fileIdentifier>
analysis_method	universal	a brief description of analysis method
annotator	universal	name of person(s) who annotated dataset
arrival_date	universal	date vessel returned to port. YYYY-MM-DD
arrive_port_fullname	universal	Port of Arrival, equivalent to Arrival Port in CruisePack
author_list	universal	for data that we plan on archiving, a list of those who should be given credit, with each author's name separated by a semi-colon. Organizations may also be listed as authors.
CMECS_keywords	universal	Keywords from controlled CMECS vocabulary
cruise_segment	universal	segment or leg of cruise
depart_port_fullname	universal	Port of Departure, equivalent to Departure Port in CruisePack
departure_date	universal	Date vessel left port. YYYY-MM-DD
east_bound_longitude	universal	eastern boundary of collection polygon in decimal degrees
funding_agency	universal	semicolon separated list of funding agencies
funding_program	universal	semicolon separated list of funding programs
general_source	universal	General source of data, reflecting broad agency or institution level (current list is: BOEM, NOAA, USGS, ACADEMIA, INDUSTRY, OTHER). If more descriptors are needed, please contact stephen.formel@noaa.gov
lineage	universal	semicolon-separated list of source datasets for a model, product, table listing injuries. Helps us stay aware of the influence of individual source datasets upon products that we use in downstream synthesis.
metadata_dialect	universal	Dialect of metadata associated with dataset (e.g., ISO 19115)
metadata_file_name	universal	Name of source metadata file. This is a string, for example 'B00004.xml'
NCCOS_data_types	universal	controlled vocabulary needed to submit data to NCCOS data managers in prep for geoportal and archiving
NCCOS_locations	universal	controlled vocabulary needed to submit data to NCCOS data managers in prep for geoportal and archiving
NCCOS_priorities	universal	controlled vocabulary needed to submit data to NCCOS data managers in prep for geoportal and archiving
NCCOS_topics	universal	controlled vocabulary needed to submit data to NCCOS data managers in prep for geoportal and archiving
north_bound_latitude	universal	northern boundary of collection polygon in decimal degrees
parameters_list	universal	semicolon separated list of parameters that might be of interest at the sample level but whose individual values aren't needed to screen the dataset. (e.g. pH ; chlorophyll)

POC	universal	primary point of contact for dataset (not within MDBC). The person who entered the data set into the catalog and serves as the MDBC POC is listed under MDBC_reviewer
POC_affiliation	universal	affiliation of the POC
POC_email	universal	email address for POC
POC_phone	universal	phone number for POC
Processing_level_native	universal	the processing level of data as they were received by the MDBC team. Possible choices include: Level 0, Level 1a, Level 1b, Level 2, Level 3, Level 4. See link for definitions of levels.
Processing_level_output	universal	the processing level of data after they were processed by the MDBC team. Possible choices include: Level 0, Level 1a, Level 1b, Level 2, Level 3, Level 4. See link for definitions of levels.
Processing_level_description	universal	description of data processing level if more detail is needed than processing_level, or if processing_level categories do not apply
project_name	universal	Equivalent to "Project" (AKA cruise_name) in R2R or "Project"(AKA Project Name/Project Number) in H93 survey metadata
purpose	universal	A brief summary of why these data were collected or the research goals intended for these data.
QAQC_method	universal	a brief description of QA/QC methods
sample_event_ID	universal	ID of the survey event (e.g., dive number, transect, or trawl haul number) on which the sample(s)and data was(were) collected
sample_preservation	universal	Specific methods for preserving a sample (e.g., dry, in water, in formaldehyde or other preservative solution)
sampling_instrument	universal	Name of sampling instrument. Can include any instrument or sampling tool. Some suggestions include: drop camera; towed camera; trawl; net; dredge; longline; pot; hook and line; grab; corer; SCUBA; other; epibenthic sled
sampling_method	universal	a brief description of sampling method
scientist	universal	Equivalent to PI, lead scientist, chief scientist
scientist_affiliation	universal	Institutional affiliation of the chief scientist
sea_name	universal	IHO Sea Area, or controlled NODC Sea Area name or GCMD sea name.
secondary_POC	universal	secondary point of contact for dataset (not within MDBC). In case the POC is no longer available.
secondary_POC_email	universal	email address for secondary_POC
secondary_POC_phone	universal	phone number for secondary_POC
south_bound_latitude	universal	southern boundary of collection polygon in decimal degrees, must be negative
survey_ID	universal	Name or ID of the cruise or project in which the sample or observation was collected. SEDCI uses Survey ID. Equivalent to Cruise ID in CruisePack, CruiseID in R2R and if relevant to NCEI accession, <gmd:MD_Identifier> <gmd:code> in NCEI XML. If we are confident that multiple survey_IDs exists for one dataset, this can be recorded as a set of semicolon-separated values.
vehicle_name	universal	name of non-ship platform
vehicle_type	universal	type of non-ship platform: glider, buoy station, ROV, AUV, etc. NCEI does have a controlled vocabulary we could use as starting point.
vessel_name	universal	equivalent to Ship Name, Vessel Name, platform_name
west_bound_longitude	universal	western boundary of collection polygon in decimal degrees, must be negative
mod_framework	predictive habitat modeling	Modeling framework(s) used e.g. Non-statistical functional/mechanistic, Maximum Entropy, Generalized Additive Model (GAM), Boosted Regression Tree (BRT), Occupancy, Ensemble of multiple frameworks, Bayesian/frequentist

mod_perf_method	predictive habitat modeling	General method(s) used to assess model performance e.g. Fit to training data (estimation), Fit to validation or test data (prediction)
mod_perf_metric	predictive habitat modeling	Metric(s) used to assess model performance e.g. Area under the receiver operating characteristic curve (AUC)
mod_software	predictive habitat modeling	Software used to implement modeling framework e.g. MaxEnt, R, JAGS
mod_tuning_pars	predictive habitat modeling	Default or custom values of any tuning parameters for the modeling framework
prediction_ID	predictive habitat modeling	Identity of primary quantity predicted by the model e.g. Habitat class, Taxon
prediction_sp_ext	predictive habitat modeling	Geographic boundaries of model predictions
prediction_sp_res	predictive habitat modeling	Unit of distance indicating spatial precision of model predictions
prediction_t_ext	predictive habitat modeling	Temporal boundaries of model predictions
prediction_uncertainty	predictive habitat modeling	Was uncertainty/precision of model predictions quantified and if so how?
prediction_var	predictive habitat modeling	Primary quantity predicted by the model e.g. Habitat suitability, Binary/multinomial, presence/absence, Probability of occurrence, Abundance/density, Area/percent, cover, Size
predictor_avail	predictive habitat modeling	Are predictor data input to the model accessible for subsequent use?
predictor_importance	predictive habitat modeling	Was relative importance of predictor variables quantified and if so how?
predictor_info	predictive habitat modeling	Link to Google spreadsheet with supplementary table providing detailed information about each predictor variable (e.g., source, spatial and temporal scope, native and processed resolution, processing methods)
predictor_response	predictive habitat modeling	Were relationships between predictor and response variables examined and presented?
predictor_sp_res	predictive habitat modeling	Unit of distance indicating range of spatial precision of predictor data input to model (minimum and maximum)
predictor_uncertainty	predictive habitat modeling	Was uncertainty/precision of predictor data input to model quantified and accounted for in model predictions?
predictor_var	predictive habitat modeling	Predictor variables and categories of predictor variables
response_avail	predictive habitat modeling	Are response data input to the model accessible for subsequent use?
response_collect	predictive habitat modeling	Method used to collect response data input to model e.g. ROV, AUV, Drop Camera, Physical collection
response_source	predictive habitat modeling	Source(s) of the response data, separated with semicolons
response_sp_ext	predictive habitat modeling	Geographic boundaries of response data input to model
response_sp_res	predictive habitat modeling	Unit of distance indicating spatial precision of response data input to model
response_t_ext	predictive habitat modeling	Temporal boundaries of response data input to model
response_type	predictive habitat modeling	Type of response data input to model e.g. Presence, background/pseudo-absence, Presence, absence, Abundance/density/percent cover, Size

response_uncertainty	predictive habitat modeling	Was uncertainty/precision of response data input to model quantified and accounted for in model predictions? E.g., location precision or taxonomic uncertainty.
ancillary_data_format	mapping	file format of Ancillary data available from data source (asvp, .0001, .txt)
ancillary_data_type	mapping	type of Ancillary Data available from data source (CTD, XBT, POSMV)
final_data_format	mapping	format of Final Data layers available from data source (type of survey tracklines, type of raster surface, videos, type of in situ samples)
final_data_type	mapping	type of Final Data layers available from data source (type of survey tracklines, type of raster surface, videos, type of in situ samples)
horizontal_datum	universal	Geographic reference, projected coordinate system, or UTM zone (e.g., NAD83 UTM 16N) [source - NBS]
operator	mapping	seems to be equivalent to "operator" in R2R, or "responsible party - organization" in NCEI. Not actually under core_att or universal_att yet, needs more discussion
project_name	mapping	Equivalent to "Project" (AKA cruise_name) in R2R or "Project"(AKA Project Name/Project Number) in H93 survey metadata
raw_data_format	mapping	file format of Raw Data from data source (ex. ALL, .wcd, .xls)
raw_data_type	mapping	type of Raw Data available from data source (survey tracklines, videos, or sample records)
spatial_resolution	mapping	Pixel cell size or node distance in square meters; this is the resolution of the provided data for which coverage information is accurate [source - NBS]
vertical_datum	mapping	Ellipsoidal reference, tidal datum, or vertical projection (e.g., MMLW or ellipsoid) [source - NBS]
qps_workflow_status	mapping	Status of raw seafloor mapping data processing in QPS software [dropdown options: N/A, complete, incomplete, problem]
arc_workflow_status	mapping	Status of seafloor mapping data processing in ArcPro software [dropdown options: N/A, complete, incomplete, problem]
network_transfer_status	mapping	Status of seafloor mapping data transfer to network [dropdown options: N/A, complete, incomplete, problem]
backscatter_workflow_status	mapping	Status of seafloor mapping backscatter data processing. Options: N/A, complete, incomplete, problem
MAP_priority_level	mapping	Priority level of a dataset for MAP team to move data through the inventory compilation workflow. Levels include: 1 - Processed data, 2 - Point data, 3 - Raw data, 4 - Transit data, 5 - Synthetic shapefiles, 6 - Other
camera_resolution	imagery	In case the image resolution is different from camera resolution. If given as one value, that indicates there is one consistent resolution across the data sets. Two semicolon-separated values indicate a min and max for the dataset. 3-5 values indicate the discrete values that represent the entire dataset.
imagery_resolution	imagery	Number of displayed pixels and lines; higher resolution allows better identification of taxa and surrounding habitat. If given as one value, that indicates there is one consistent resolution across the data sets. Two semicolon-separated values indicate a min and max for the dataset. 3-5 values indicate the discrete values that represent the entire dataset.
media_type	imagery	e.g. hard drives, DVD, mini DV, Hi8. Can be controlled list if helpful.

Appendix B: Data Sources

A list of data sources used by the HAE inventory and analysis working groups is below. A comprehensive cross-project data portal catalog was created by the technical team.

- NOAA Deep Sea Coral Data Portal
<https://deepseacoraldata.noaa.gov/>
- NOAA DIVER
<https://www.diver.orr.noaa.gov/>
- NOAA GOM Data Atlas
<https://gulfatlas.noaa.gov/>
- NOAA OER Video Portal
<https://www.ncei.noaa.gov/access/ocean-exploration/video/>
- NCEI Ocean Carbon and Acidification Data Portal
<https://www.ncei.noaa.gov/access/oads/>
- NCEI Ocean Carbon Data System
<https://www.ncei.noaa.gov/access/ocean-carbon-data-system/>
- NCEI World Ocean Database
<https://www.ncei.noaa.gov/access/world-ocean-database-select/dbsearch.html>
- EPA Science Topics Data and Tools
<https://www.epa.gov/environmental-topics/science-topics>
- EPA EnviroFACTS
<https://www.epa.gov/environmental-topics/science-topics>
- OceanReports
<https://coast.noaa.gov/digitalcoast/tools/ort.html>
- Southeast Deep Sea Coral Digital Atlas
- NCBI GenBank
<https://www.ncbi.nlm.nih.gov/genbank/>
- BOLD Systems
<http://www.boldsystems.org/>
- USGS ScienceBase
<https://www.sciencebase.gov/catalog/>
- USGS GOM GIS Data Catalog
- USGS Coastal and Marine Geology Program Interactive Map
- Smithsonian Invertebrate Zoology Collections
<https://collections.nmnh.si.edu/search/iz/>
- UK Natural History Museum Collections
<https://data.nhm.ac.uk/>
<https://www.ncbi.nlm.nih.gov/genbank/>
- DRYAD
<https://datadryad.org/stash>
- GeOMe Database
- GRIIDC
<https://data.gulfresearchinitiative.org/>

- BOEM ESPIS
<https://marinecadastre.gov/espis/#/>
- BOEM GOM SCHEMA (limited access)
- BSEE Data Center
<https://www.data.bsee.gov/>
- OBIS
<https://obis.org/>
- MarineCadastre
<https://marinecadastre.gov/>
- The Coral Portal
<http://portal.gulfcouncil.org>
- FIN Commercial Landings Database
<https://www.gsmfc.org/>
- SEAMAP Gulf Monitoring database
<https://www.gsmfc.org/seamap.php>

Appendix C: Literature Review Framework

Objective #1: Fill data gaps and determine baselines for health and condition to guide restoration and protection

Parameter 1. Sampling sufficiency for population/communities

- Key Metrics
 - Dominant, managed or indicator species identified
 - Age/size distributions
 - Biomass
 - Density
 - Abundance
 - Community composition
 - Diversity
 - Richness
 - % cover
 - Ontogenetic changes
- Methods
 - Assess for each metric
- Structure/Function
 - Documented community structure/function
 - Compare changes in reference communities vs. injured communities

Parameter 2. Identified species and habitat interactions

- Key Activities
 - Feeding/predation
 - Trophic Flow
 - Reproductive strategies
 - Use of substrate or structure (habitat use)
 - Water column
 - Diel cycles
 - Nursery habitat
 - Ontogenetic changes
 - Suitability
 - Environmental variables (temp, Dissolved Oxygen, etc.)
- Methods
 - Assess for each metric

Parameter 3. Identified metapopulation dynamics and genetic diversity

- Genetics
 - Molecular barcodes
 - Population structure
 - Genetic diversity

- Source/sink population demographics
- Rate and direction of genetic exchange
- Growth rates
- Connectivity
 - Mesophotic to deep
 - Pelagic to benthic
 - Within mesophotic
 - Within deep
- Methods
 - Assess for each metric

Parameter 4. Coral larvae dispersal

- Model development
- Model testing and validation

Objective #2: Identify and assess natural and anthropogenic threats

Parameter 5. Documented threats and impacts

- Natural
 - Water quality anomalies
 - Climate change
- Anthropogenic
 - Oil spills
 - Invasive species
 - Anchoring
 - Fishing activities
 - Marine debris
- Methods
 - Assess for each metric

Objective #3: Detect and quantify trends for inference of future impacts and assess success of restoration

Parameter 6. Recovery Trajectories and restoration targets

- Key Hard bottom/soft bottom metrics
 - Age/size distribution
 - Biomass
 - Density
 - Abundance
 - Growth
 - Recovery
 - Mortality
 - Recovery trajectories
- Methods
 - Assess for each metric

