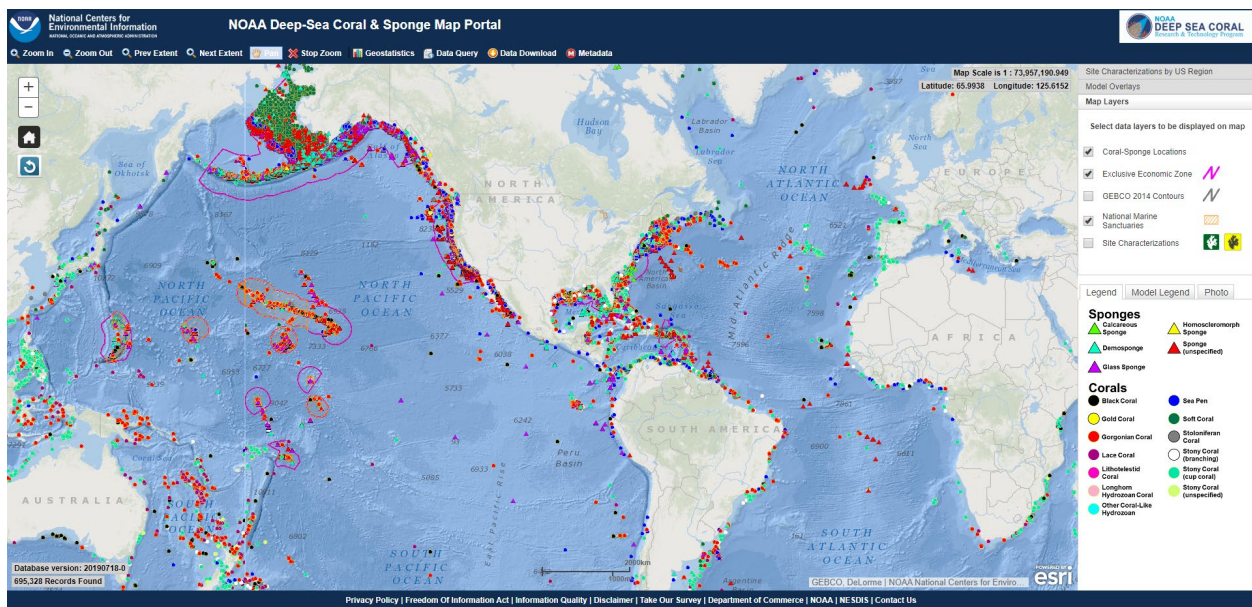# NOAA's National Database for Deep-Sea Corals and Sponges: 2020 Status Update

Robert P. McGuinn, Thomas F. Hourigan, Scott L. Cross, L. Matthew Dornback, Peter J. Etnoyer, David E. Sallis, Heather M. Coleman

NOAA's National Database for Deep-Sea Corals and Sponges: 2020 Status Update

McGuinn, Robert P. [1,2] , Thomas F. Hourigan [3], Scott L. Cross [1], L. Matthew Dornback [4,5], Peter J. Etnoyer [6], David E. Sallis [1,2], Heather M. Coleman [3]

1.   NOAA, National Environmental Satellite, Data, and Information Service, National Centers for Environmental Information

2.   Mississippi State University, Northern Gulf Institute

3.   NOAA, National Marine Fisheries Service, Office of Habitat Conservation, Deep Sea Coral Research and Technology Program

4.   NOAA, Oceanic and Atmospheric Research, Office of Exploration and Research

5.   Cherokee Nation Strategic Programs

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

•   corresponding authors (McGuinn, Robert P. [1,2], Thomas F. Hourigan [3])

Copies of this report may be obtained from:

Heather Coleman
NMFS Office of Habitat Conservation
National Oceanic and Atmospheric Administration
1315 East-West Highway, Room 14201
Silver Spring, MD 20910

Or online at:
http://spo.nmfs.noaa.gov/tech-memos/ or *http://deepseacoraldata.noaa.gov/libary/*

Table of Contents

Executive Summary

The National Oceanic and Atmospheric Administration (NOAA) Deep Sea Coral Research and Technology Program (Program) compiles, curates, and makes publically accessible a National Database of biogeographic data and information on deep-sea corals and sponges – the most important habitat-forming organisms in the deep sea. The database is accessible through the Program's Data Portal (*https://deepseacoraldata.noaa.gov/*) and provides geographically-specific information on the diversity, distribution and abundance of corals and sponges. The database focuses on records of occurrences within the United States Exclusive Economic Zone (EEZ), but the Program includes records outside of the U.S. EEZ where available and is actively encouraging partnerships worldwide to enhance the geographic scope of the data served. The database fulfills NOAA's requirements under the reauthorized Magnuson-Stevens Fishery Conservation and Management Act (MSA, 2007) to identify and map locations of deep-sea corals, and to submit this information for use by regional fishery management councils (FMCs). The database is the most comprehensive, quality-assured, national-scale, data portal for deep-sea corals and sponges available.

This report presents information on the development and enhancements to the database since it became accessible through the online portal in 2015 (Hourigan et al., 2015). Through the sustained partnership and funding activities of the Program, and the contributions of 73 unique data providers, this database now includes over 740,000 records of deep-sea corals (n = 488,876) and sponges (n = 254,877) world-wide (at database version: 20191217-0) along with extensive metadata. The database is now being used as an essential trusted input for ocean planning, resource management, and conservation activities of FMCs and beyond.

The objectives of this report are as follows.

- Provide an update on the continued development of the database through time in the form of a summary of its current contents for key variables.

- Identify enhancements and additions to the database and related data products since 2015. These include, *inter alia*, the following:

    – Summaries and metadata for each data set (238 data sets from 73 unique data providers to date);

    – Further standardization of coral taxonomy;

    – Institution of additional fields that accommodate standardized habitat descriptions, comparisons of recorded depths to modeled bathymetry, additional information on areas surveyed, morphospecies descriptions, and highlight exemplary images of taxa.

- Describe quality control and assurance processes used to build and maintain the database.

- Acknowledge the many contributors to the database (as of the date of this publication).

- Direct data providers and data users to additional online resources.

NOAA's National Database for Deep-Sea Corals and Sponges: 2020 Status Update

Background

Corals and sponges are the most important organisms creating deepwater biogenic habitats, and have become a centerpiece of deep-sea conservation in both the United States and internationally (Hourigan et al. 2017). The National Oceanic and Atmospheric Administration (NOAA) Deep Sea Coral Research and Technology Program (DSCRTP), herein referred to as the Program, compiles, curates, and makes publicly accessible a National Database of biogeographic data and information on deep-sea corals and sponges. The database contains observations (records) of corals and sponges (primarily those deeper than 50 meters), location and depth, and other relevant data that may be available, such as their size, abundance, and associated environmental conditions, as well as information on how the data were collected. Hourigan et al. (2015) provides an introduction and overview of the National Database.

The database is designed to be a resource for both scientists and natural resource managers, and is a primary tool linking the Deep Sea Coral Research & Technology Program's research to management. The database fulfills NOAA's requirements under the reauthorized *Magnuson-Stevens Fishery Conservation and Management Act* (MSA Sec. 408, 2007) to identify and map locations of deep-sea corals, and to submit this information for use by regional fishery management councils (FMCs). See *http://www.fisherycouncils.org* for an overview of the structure and function of the FMCs. The database focuses on deep-sea corals and sponges within the United States Exclusive Economic Zone (EEZ), but increasingly includes records outside of the U.S. EEZ where available. To enhance interoperability with other global biological occurrence databases, the field names, definitions, and valid values adhere to *Darwin Core standards* where possible. The Program is actively encouraging partnerships worldwide to enhance the geographic reach of the data served. To encourage these partnerships and to further data awareness and usage globally, the database is also shared with the Ocean Biogeographic Information System, USA, *OBIS-USA*, and the Global Biodiversity Information Facility, *GBIF*, on a quarterly basis.

The database was originally conceived in 2009, building upon an initial United States Geological Survey Cold-Water Coral Geographic Database that was limited to Atlantic coral records (Scanlon et al., 2010). NOAA's database was developed, and continues to be refined, in collaboration with data providers, data users from the science and resource management communities, and organizations involved in similar data efforts. Since 2012, the database has undergone rapid development experiencing periods of exponential growth during major data harvests (**Figure 1**). From 2012 to 2015 was a period of internal growth and development of the database, including the establishment of new field names, definitions, and accepted (valid) values. Additionally, a web-based mapping and data download site was established and is available at *DeepSeaCoralData.NOAA.gov*. Online access to the database, through the DSCRTP map portal, began in 2015. Growth from 2015 and onward was marked by quarterly data updates made available to the public. The quarterly database updates include the addition of new occurrences as well as improvements to existing records.

The database includes 108 publically available fields that relate to the location, taxonomy, environmental conditions, and the provenance of the observations. Not all records include values for all available fields. The completeness of the individual records depends on completeness of the incoming data set. All of the values in this report are based on version 20191217-0 of the National Database. Some of the reported values will change as the database is updated each quarter with new records and corrections to old records. There are currently over 740,000 individual records in the database composed of 238 individual data sets from 73 different data providers. There are currently 488,876 records of corals and 254,877 sponge records. See **Table 1** for a full summary of all of the data providers.
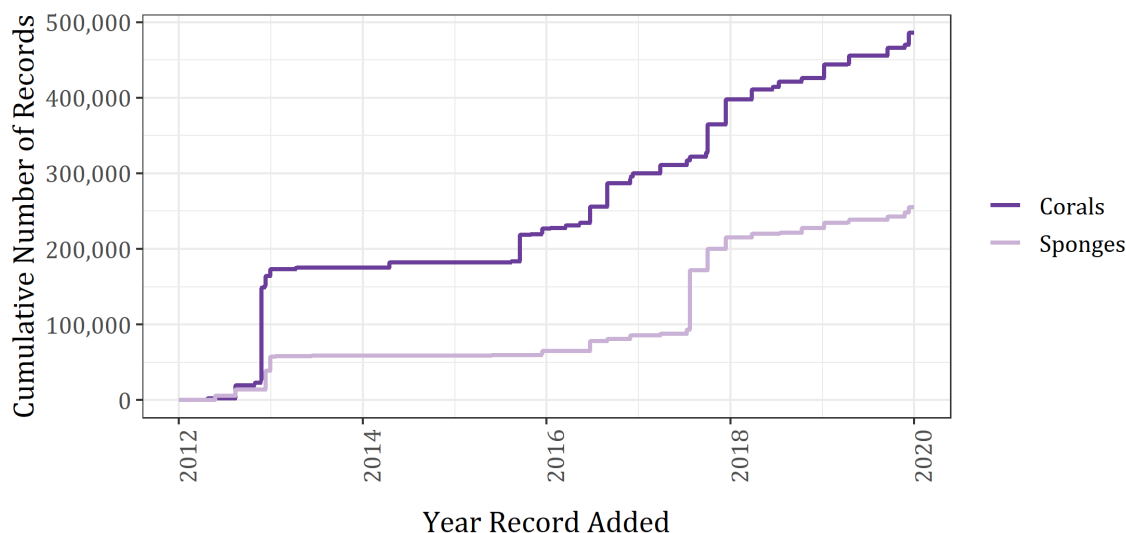
Figure 1: Growth of NOAA's National Database for Deep-Sea Corals and Sponges through time.

This document is a companion and update to an earlier NOAA Technical Memorandum entitled "*An Introduction to NOAA's National Database for Deep-Sea Corals and Sponges*" (Hourigan et al., 2015). Readers are referred to that document for more detail on the Program and its mandates, database structure, and on general principles that informed the design of the database and related data. The objectives of this status update include the following.

- Acknowledge contributors to the database (as of the date of this publication).

- Describe quality assurance and control (QA/QC) processes used to build and maintain the database.

- Provide an update on the continued development of the database through time in the form of a summary of its current contents for key variables.

- Direct data providers and data users to additional online resources.

Note to readers:
Throughout this document, the term 'DSCRTP web portal' is used to refer to the Program's general web portal at *https://deepseacoraldata.noaa.gov/*. The term, 'DSCRTP map portal' refers specifically to

the interactive map and data download functionality available here:
https://www.ncei.noaa.gov/maps/deep-sea-corals/mapSites.htm.

Throughout this document, when referring to a field name from the National Database, it will be shown in single quotes exactly as it appears in the database, for example, 'DataProvider'. These field names are also used in the tables and figures for consistency and to familiarize the reader with the vocabulary used in the database. For a full list of all the field names, their definitions, and the valid values see **Appendix 1**. For convenience, the most current Excel spreadsheet version of the data dictionary is located on the DSCRTP web portal at this *URL*. Please note data dictionary version number. The most recent version of the data dictionary published online at that link should be used instead of the version published in this document as changes may have occurred post-publication.

Overview of Changes to Data Working Group Workflows

Since the first online publication of NOAA's National Database for Deep-Sea Corals and Sponges in 2015, the Program has continued to develop and implement workflows to increase the number of observations, to improve the quality of existing records, and to add related data products such as Geographic Information System (GIS)-formatted data layers. The following section describes the main changes and additions to the Data Working Group workflows and processes.

Quarterly updates

The Program's Data Working Group updates the National Database with new observations and improvements to existing records on a quarterly basis. The quarterly release cycle allows for better internal coordination and external communication regarding database improvements and anticipated additions and changes.

New versions of the database are published on the DSCRTP map portal. The most recent database version number is indicated in the bottom left corner of the DSCRTP map portal screen as a date expressed in the format 'YYYYMMDD_X' where 'X' is an integer that captures additional iterations for a given version as needed. Older versions of the database are kept in perpetuity and are available upon request from the Program. The Program also uses these older versions to return individual records to an earlier time point if errors are recognized. The database is also archived annually at the NOAA National Centers for Environmental Information (NCEI) at the following location: *https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0145037*. *It is important to note* that the most current National Database is always the version available at the *DSCRTP map portal* and not at the preceding NCEI archive link.

Automated quality control

All data submissions are subject to automated quality assurance and control routines written in both R and Python scripts that check for potential errors in depth, geographic location, and taxonomy. This is a first step in a more extensive process that includes hands-on review of each data set. More details on the quality assurance/quality control (QA, QC) processes are provided below in the section entitled "Overview of Data Ingestion and Quality Assurance Processes."

Standardized taxonomy

The Program has implemented a master taxonomy table to standardize all taxonomic information in the National Database. The master taxonomy table is expanded as new taxa are acquired. All taxonomic information is based on the standards maintained at the World Register of Marine Species (WoRMS) database (Horton et al., 2018). Our taxonomic list is a subset of the WoRMS database filtered to include only the organisms that are within the taxonomic scope of the National Database for Deep-Sea Corals and Sponges. For practical purposes, the database initially deviated from the WoRMS standard taxonomy by retaining the cnidarian Order Gorgonacea (generally subsumed in the Order Alcyonacea). In the 2019 update, the database has become fully compliant with WoRMS; gorgonians have been included in the Order Alcyonacea and the structural aspects of Gorgonacea have been captured in a separate field ('VernacularNameCategory' = gorgonian coral). For a discussion on the database's taxonomic scope please refer to Hourigan et al. (2015).

Standard metadata describing component data sets

For the purposes of attribution, maintaining clear provenance, and facilitating QA/QC and end-user confidence, data are identified and tracked as discrete data sets. The Program recognizes four different categories of data set: (1) data from a specific cruise, 'cruise'; (2) data from an identifiable multiyear program, such as a fisheries survey series, 'program'; (3) records from museum collections 'repository'; and (4) records from the published literature, 'literature'. As of this writing there are 238 data sets as defined by the Program; this number grows with each quarterly update (**Table 1**). International Standards Organization (ISO) compliant metadata for each component data set are published on the DSCRTP web portal. The 'DatasetID' field is used to identify unique data sets as defined by the program. See **Appendix 3** of this document for details on accessing summary dashboards for each 'DatasetID' and downloading individual data sets and metadata.

Changes to field definitions and valid values

Changes are occasionally needed to make improvements to database fields, definitions, and valid values. The details of these changes are given below in the section, 'Important Changes and Additions to Database Fields'.

Availability of related content

Some types of content are added to the DSCRTP web portal as they become available, such as Site Characterization Reports from Program-funded cruises, as well as new and improved models of habitat suitability for various taxa of coral. Both products are available for viewing and download from the DSCRTP map portal, along with details regarding model creation. The Program is committed to supporting the development and distribution of these ever-improving models for data visualization and exploration as they become available.

Database Contributors

The National Database contains contributions from museums, research programs, research cruises, and literature (**Table 1**). The database currently contains data contributions from 73 unique data providers.

Table 1: Data providers to NOAA's National Database for Deep-Sea Corals and Sponges. There are 73 data providers and 238 individual data sets. The database version used for the creation of this table is 20191217-0.

DataProvider	Type of Data	DatasetIDs	# of Records
American Museum of Natural History	Repository	AMNH_IZ	2
Barcode of Life Data System	Repository	BOLD	8
Bermuda Museum of Natural History	Repository	BNHM	3
Bishop Museum, Invertebrate Zoology Collection	Repository	BPBM	237
Bluhm et al. (2009)	Literature	Bluhm_etal_2009	25
Brooks et al. (2016)	Program	BOEM_Lophelia_II	2,546
Bureau of Ocean Energy Management	Cruise	NOAA_FGBNMS_PSBF-I | NOAA_FGBNMS_PSBF-II | NOAA_FGBNMS_PSBF-III | USGS_SJ-09-09	53
	Program	BOEM_Lophelia_I | BOEM_Lophelia_II | BOEM_Mid-Atlantic_Canyons	3,966
California Academy of Sciences	Repository	CAS | OET_NA085	1,548
	Cruise	OET_NA085	46
Carranza et al. (2012)	Literature	Carranza_etal_2012	13
Cordeiro et al. (2015)	Repository	MNRJ | MOUFPE	22
Cordeiro et al. (2017)	Repository	NHM_UK | ZMB	4
Florida Museum of Natural History	Repository	FLMNH_IZ	162
Hall-Spencer et al. (2007)	Literature	Hall-Spencer_etal_2007	83
Harbor Branch Oceanographic Institute	Cruise	HBOI_EL-99-08 | HBOI_SJ-06-02 | HBOI_SJ-06-05-L1 | HBOI_SJ-06-05-L3 | HBOI_SJ-08-09 | HBOI_SJ-10-07 | HBOI_SJ-88-11 | HBOI_WS-17-125 | NOAA_NF-11-09 | NOAA_NF-11-09-L3 | NOAA_NF-14-08 | NOAA_PC-11-05-L1 | NOAA_PC-12-03 | NOAA_PC-13-03 | NOAA_PC-15-02 | NOAA_PC-16-02 | NOAA_PC-17-02 | NOAA_RB-03-07 | NOAA_SJ-01-08-A | NOAA_SJ-02-08 | NOAA_SJ-04-05 | NOAA_SJ-05-04 |	27,858

DataProvider	Type of Data	DatasetIDs	# of Records
		NOAA_SJ-05-08 | NOAA_SJ-05-11 | NOAA_SJ-07-05 | NOAA_SJ-07-06 | NOAA_SJ-09-08	
Hawai'i Undersea Research Laboratory	Cruise	OET_NA101	13,258
	Repository	HURL | MCZ_IZ	35,713
Herrera et al. (2012)	Literature	Herrera_etal_2012	4
Lapointe & Watling (2015)	Literature	Lapointe_Watling_2015	2
MacDonald, Ian | Silva, Mauricio	Cruise	USGS_TM-97-10 | USGS_TM-98-04 | USGS_TM-98-08 | USGS_TM-99-07	2,330
Monterey Bay Aquarium Research Institute	Repository	CAS | MBARI	217,034
Muséum national d'Histoire naturelle	Repository	MNHN	37
Museum of Comparative Zoology, Harvard University	Repository	MCZ_IZ	2,205
National Museum of Natural Science, Taichung, Taiwan	Repository	NMNS_Taiwan	1
National Undersea Research Center, University of Connecticut	Repository	NURC	17
Natural History Museum, London, UK	Repository	NHM_UK	3
New Zealand National Institute of Water and Atmospheric Research (NIWA)	Repository	NIWA	582
NOAA, Alaska Fisheries Science Center	Cruise	NOAA_AE-14-04 | NOAA_AP-13-08 | NOAA_DY-15-07 | NOAA_FMP-12-01 | NOAA_FMP-12-02 | NOAA_FMP-14-01 | NOAA_FMP-14-02 | NOAA_MD-13-08 | NOAA_RR-04-07 | NOAA_SS-10-08 | NOAA_SS-12-08 | NOAA_VA-14-08 | NOAA_VO-05-08	27,131
	Program	NOAA_AFSC_Bottom_Trawl_Survey | NOAA_AFSC_Longline_Survey	55,759
NOAA, Center for Coastal Monitoring and Assessment	Cruise	NOAA_NF-05-05 | NOAA_NF-06-03 | NOAA_NF-07-06 | NOAA_NF-08-04 | NOAA_NF-09-01 | NOAA_SH-15-03 | NOAA_SW-13-06 | NOAA_SW-16-08	5,455
	Repository	NOAA_NBI	6
NOAA, Channel Islands National Marine Sanctuary	Cruise	NOAA_SH-10-11 | NOAA_SW-15-08	1,984
NOAA, Coral Reef Conservation Program	Cruise	NOAA_NF-12-01 | NOAA_NF-13-02 | NOAA_NF-14-01	1,082

DataProvider	Type of Data	DatasetIDs	# of Records
NOAA, Cordell Bank National Marine Sanctuary	Cruise	NOAA_FM-07-05 | NOAA_FM-14-09 | NOAA_M2-04-02 | NOAA_M2-10-06-L2-ROV | NOAA_VO-01-09 | NOAA_VO-02-09 | NOAA_VO-03-09 | NOAA_VO-04-08 | NOAA_VO-05-09 | OET_NA085	15,504
NOAA, Deep Sea Coral Research & Technology Program and Office of Ocean Exploration and Research	Cruise	NOAA_EX-17-03 | NOAA_EX-17-05 | NOAA_EX-17-06	168
NOAA, Deep Sea Coral Research and Technology Program	Cruise	NOAA_RB-10-08 | NOAA_SJ-09-08	4,350
NOAA, Flower Garden Banks National Marine Sanctuary	Cruise	NOAA_DFH-30 | NOAA_DFH-32-33	3,826
	Program	NOAA_FGBNMS_DFH	3,961
NOAA, Gray's Reef National Marine Sanctuary	Cruise	NOAA_PC-10-02	775
NOAA, Greater Farallones National Marine Sanctuary	Cruise	NOAA_FM-12-10	1,281
NOAA, National Centers for Coastal Ocean Science	Cruise	NOAA_NF-17-08 | NOAA_NF-18-04	6,868
NOAA, Northeast Fisheries Science Center	Program	NOAA_NEFSC_Benthic_Cruises	98
	Repository	NOAA_NEFSC_BenthicDatabase	488
NOAA, Northeast Fisheries Science Center | University of Connecticut	Cruise	NOAA_CT-13-07	2,318
NOAA, Northwest Fisheries Science Center	Cruise	NOAA_FM-11-01 | NOAA_M2-10-06-L1-AUV | NOAA_M2-10-06-L3-AUV | NOAA_PS-10-01-L1 | NOAA_PU-14-13 | NOAA_TN174	10,467
	Program	NOAA_NWFSC_Bottom_Trawl_Survey | NOAA_NWFSC_FRAM	5,817
NOAA, Office of Exploration and Research | Hawai'i Undersea Research Laboratory, University of Hawai'i	Cruise	NOAA_EX-17-08	11,381
NOAA, Office of Exploration and Research | Hawai'i Undersea Research Laboratory, University of Hawai'i	Cruise	NOAA_EX-15-04-L2 | NOAA_EX-15-04-L3 | NOAA_EX-15-04-L4 | NOAA_EX-16-03 | NOAA_EX-16-05-L1 | NOAA_EX-16-05-L3 | NOAA_EX-16-06 | NOAA_EX-17-02 | NOAA_EX-17-03 | NOAA_EX-17-05 | NOAA_EX-17-06	54,806
	Repository	NMNH_IZ	1
NOAA, Office of Ocean Exploration and Research	Cruise	NOAA_AT-02-06-L15 | NOAA_AT-04-08 | NOAA_EX-11-02 | NOAA_EX-12-02-L2 | NOAA_EX-14-02-L3 | NOAA_EX-17-11 |	11,560

DataProvider	Type of Data	DatasetIDs	# of Records
		NOAA_EX-18-03 | NOAA_EX-18-06 | NOAA_OP-02-07	
NOAA, Office of Response and Restoration	Cruise	DWH_NRDA_HC-11-09 | DWH_NRDA_NF-10-14 | DWH_NRDA_WS-14-06	2,676
NOAA, Olympic Coast National Marine Sanctuary	Cruise	NOAA_DFO_TU-08-07 | NOAA_M2-06-05 | NOAA_PS-11-08	36,279
NOAA, Southeast Fisheries Science Center	Cruise	NOAA_LS-05-01 | NOAA_PC-11-05-L1	21
NOAA, Southwest Fisheries Science Center	Cruise	NOAA_M2-10-06-L3-ROV | NOAA_OL-03-11 | NOAA_OL-04-03 | NOAA_OL-04-09 | NOAA_OL-04-10 | NOAA_OL-05-01 | NOAA_OL-05-02 | NOAA_OL-05-03 | NOAA_OL-05-04 | NOAA_OL-05-06 | NOAA_OL-05-11 | NOAA_OL-06-02 | NOAA_OL-06-05 | NOAA_OL-07-01 | NOAA_OL-07-08 | NOAA_OL-07-10 | NOAA_OL-07-12 | NOAA_OL-08-02 | NOAA_OL-08-03 | NOAA_OL-08-04 | NOAA_OL-08-06 | NOAA_OL-08-10 | NOAA_OL-10-10 | NOAA_OL-10-11 | NOAA_OL-11-04 | NOAA_OL-11-09 | NOAA_OL-11-12 | NOAA_PU-11-08 | NOAA_RL-16-06 | NOAA_SH-10-11 | NOAA_SW-07-09 | NOAA_SW-08-08 | NOAA_SW-08-09 | NOAA_SW-08-10 | NOAA_SW-09-07 | NOAA_SW-09-08 | NOAA_SW-10-07 | NOAA_SW-11-08 | NOAA_SW-12-08 | NOAA_VO-02-10 | NOAA_VO-07-09 | NOAA_VO-07-10 | NOAA_VO-08-09 | NOAA_VO-08-10 | NOAA_VO-10-10 | OET_NA072	94,520
	Program	NOAA_SWFSC_AST | NOAA_SWFSC_Submersible	35,327
Ocean Biogeographic Information System, USA	Cruise	NOAA_MF-76-09	23
	Literature	Feder_etal_2006	3
	Program	EPA_EMAP	92
	Repository	Gwaii_Haanas	44
Oceana	Cruise	Oceana | Oceana_ML-11-06	116
Oceana and Marine Applied Research & Exploration (MARE)	Cruise	Oceana_SW_16-08	3,149
Oculina Banks Geographic Information System	Cruise	NOAA_SJ-01-08-B	42

DataProvider	Type of Data	DatasetIDs	# of Records
Opresko (1972)	Literature	Opresko_1972	30
Pante & France (2010)	Repository	MNHN	2
Pante et al. (2012)	Literature	Pante_etal_2012	22
Pante et al. (2015)	Literature	Pante_etal_2015	850
Pennsylvania State University	Cruise	ECOGIG_O2-17-06 | ECOGIG_OI-16-10 | NOAA_EX-12-02-L2 | NOAA_HC-11-10 | OET_NA028 | OET_NA043 | OET_NA057 | OET_NA058 | SOI_FK006B | WHOI_AT-18	1,084
	Program	BOEM_Lophelia_II	7
Ross (2012)	Cruise	NOAA_SJ-05-10 | NOAA_SJ-05-11 | NOAA_SJ-07-05	86
Sampaio et al. (2009)	Literature	Sampaio_et_al_2009	4
Santa Barbara Museum of Natural History	Repository	SBMNH	112
Sautya et al. (2010)	Literature	Sautya_etal_2010	1
Sautya et al. (2011)	Literature	Sautya_etal_2011	30
Schmidt Ocean Institute	Cruise	SOI_FK004E | SOI_FK006C	373
SeamountsOnline (Seamount Biota)	Repository	Seamounts	42
Senckenberg Nature Research Society	Repository	SNRS	3
Sharuga (2014)	Literature	Sharuga_SM_2014	125
Smithsonian Institution, National Museum of Natural History	Cruise	NOAA_EX-16-05-L3 | NOAA_EX-17-06	3
	Repository	NMNH_IZ	25,022
Temple University	Program	BOEM_Lophelia_II	8,344
Texas A&M, Biodiversity Research and Teaching Collections	Repository	Texas_AM_BRTC	573
Thoma (2013)	Literature	Thoma_J_2013	153
	Repository	YPM_IZ	10
Tu et al. (2012)	Repository	ASIZ	3
Tu et al. (2015)	Literature	Tu_et_al_2015	2
	Repository	MNCN	3
Tu et al. (2016)	Repository	IORAS | MNHN | NBC_NL	11
United States Geological Survey	Cruise	USGS_CH-10-09 | USGS_SJ-09-09	17
	Repository	USGS_CoWCoG	744
Wagner (2011)	Literature	Wagner_D_2011	11
Washington State University	Literature	Bianchi_C_2011 | Pirtle_J_2002	85
Yale University Peabody Museum	Repository	YPM_IZ	866

How to Contribute to the National Database

The Program strives to establish direct communication with data providers prior to submission to the National Database. Individuals who wish to submit coral and sponge occurrence data are encouraged to first contact the Data Working Group directly at *deepseacoraldata@noaa.gov* to engage prior to the data assembly and submission process. The Program is available to discuss the format and structure of the database in detail and answer any questions. The Program's Data Working Group is also available to review sample products before final submission. A set of web-based resources to assist data providers is available on the DSCRTP web portal at the following link: *https://deepseacoraldata.noaa.gov/internal-documents/program-guidance/guidance-for-fieldwork-data-and-reporting*

Contributors to the National Database should consult **Appendix 2** of this document for a discussion of the required and desired fields for different types of data providers and research efforts. This information is also included on the database record submission template available at: *https://deepseacoraldata.noaa.gov/internal-documents/program-guidance/science-team-guidance-for-data-management/dscrtp-database-records-submission-template.xlsx/view*

Why Contribute Data to the National Database?

Contributing data to NOAA's National Database for Deep-Sea Corals and Sponges can accomplish the following goals:

- Provide a permanent and widely-accessible repository that serves diverse stakeholders interested in the science, management, and conservation of deep-water ecosystems, including scientists, resource managers, policy makers, economists, and educators.

- Meet the requirements of scientific journals to provide access to underlying data sets.

- Meet the mandate of U.S. federally-funded research to provide public access to research results, as well as the data-sharing requirements of numerous funding agencies.

- Allow you to easily link to and share your data sets online, helping to increase visibility of your work and increase your citation rate.

- If you have still images that are associated with your records the Program will host and link to those through the DSCRTP mapping portal, making them available for evaluation by other researchers.

- Allow your data to easily integrate with other national and international databases. The Program provides annual uploads of the database to the Ocean Biogeographic Information System (OBIS) and the Global Biodiversity Information Facility (GBIF) through the U.S. OBIS node (OBIS-USA).

- Help with the quality checking of existing database records.

- Save money by using established NOAA infrastructure to host and serve your data.

- Ensure that your data are thoroughly documented and permanently archived at NOAA's National Centers for Environmental Information.

- Stimulate funder interest and continued research in your region.

- Become part of the largest deep-sea coral and sponge data community in the world!

Overview of Data Ingestion and Quality Assurance Processes

The Program is intent on building a high-quality, value-added aggregation of deep-sea coral and sponge occurrence data that overcomes the limitations of individual incoming data sets by placing them in the context of other resources. Incoming data are checked against valid values for taxonomy, vessel names, vehicle names, and data providers. A systematic QA/QC process matches incoming values against standard value sets and thereby identifies, flags, and corrects errors where possible. All changes are documented and all previous versions of each data set, starting from the original file, are saved on NOAA computer systems for provenance tracking. The intention of the information presented below is to inform and update the reader on the Program's latest quality assurance procedures.

Initial Intake Inspection and Documentation Phase

- The original data set is received from the data provider and archived as a 'Level 0' data product. The Program's data submission template facilitates data entry and streamlines incorporating data into the database.
- Level 0 data are then saved in a uniform folder structure on NOAA systems with notes and correspondence from the data provider.
- Data sets are assigned a standard file name and assigned an 'AccessionID', which tracks each data set received, along with any changes that occur between initial submission to publication.
- Data sets are reviewed by Program staff familiar with deep-sea corals and NOAA data protocols. Additional notes on the submission are captured and filed. Additional communication with the data provider is initiated, if needed, to correct recognized errors or to clarify certain values.
- All changes from the original are documented and tracked in an issue tracking system and all subsequent versions of the data are saved so that incorrect transformations or corrections can be easily reversed.

Initial Conformation

- Data undergo a coded QA/QC process using R and Python scripts.
- Initial automated routines are performed to make sure data conform to the Program's data dictionary (see Appendix 1).
- These automated routines accomplish the following tasks:
- Check field names and transform them where necessary.
- Add any missing field names to complete the schema structure.
- Populate null values with the appropriate entries, '-999' for integer and 'NA' for text.
- Rearrange the data set in a standardized column order.

Technical Review Phase

Automated quality assurance scripts are run and manual review steps are completed that include the following:

- The data are evaluated regarding how well they conform to the specific valid values listed in the data dictionary (see **Appendix 1** for the full data dictionary).
- Taxonomic problems with any incoming 'ScientificName' are identified and corrected based on the master taxonomy table values which are synced with the World Register of Marine Species (WoRMS) database.
- A standardized master taxonomy table is joined with the 'ScientificName' values to add accepted values for Phylum, Order, Family, etc.
- Depths are checked against the best available regional bathymetry and mismatched values are flagged where they are out of the expected range.
- A visual summary referred to as a 'QA dashboard' is generated using R Markdown scripting. The dashboard visualizes the values in each field in such a way as to summarize and highlight possible issues. The dashboard report also provides maps of the positions of all occurrences.
- All issues recognized during technical review are listed and corrected if possible.
- If an issue is not correctable, then the problem records are flagged and not released to the public.
- A final automated QA dashboard is generated to inform the next stage of the process, expert review.

Expert Review Phase

- A subject matter expert conducts a final review using the information gleaned from the technical review, along with an additional inspection of the final QA dashboard, to find any outstanding issues that were not previously recognized.
- All remaining issues are listed and corrected.
- The expert reviewer then indicates that the data set is ready for publication in the National Database.

Important Changes and Additions to Database Fields

The data schema, published in 2015, has remained relatively stable, but changes have been made to enhance the utility of the database, and to improve clarity of the field definitions set forth in Hourigan et al. (2015). The full data dictionary for the National Database contains 135 fields, of which 108 are published and 27 are unpublished and used internally by the Program for data management purposes. For a discussion of the required and desired fields for submitting data to the Program, please see **Appendix 2**.

Logical groupings of fields, called "Field Category" in the data dictionary, reflect the higher level organization of the database and provide a convenient way to group related fields. Two new categories, "GIS Enhancement" and "Flagging," have been added to capture a suite of new fields, bringing the number of field categories from eight to ten.

The "Field Category" groups are defined as follows:

- **Survey Data:** Fields that pertain to a whole survey or cruise, such as vessel name, vehicle name, or cruise number.
- **Event Data:** Data that pertain to an entire dive or event (e.g., trawl, transect, ROV dive).
- **Observation Data:** Fields that that hold information about an individual record.
- **Taxonomic ID:** Fields that capture taxonomic ranks and hierarchy, along with details regarding the identification of a taxon.
- **Environment:** Fields that describe the habitat, substrate, and oceanographic properties in the same location as the observation.
- **Occurrence Detail:** Fields that capture further detail about each occurrence or observation.
- **GIS Enhancement:** Fields that originate from geospatial processing by the Program that are designed to enhance and supplement the data originally submitted. Example: 'LargeMarineEcosystem'.
- **Metadata:** Fields that describe the origin or provenance of the records.
- **Record Keeping:** Fields that capture unique identifiers for each data set and for each individual database record, 'DatasetID' and 'CatalogNumber', respectively.
- **Quality Flags:** Quality and usability flags defined by the Program.

All of the public and private fields for are listed in **Table 2a** and **Table 2b**. For a full list of current field names in the database, their definitions, and their valid values, see **Appendix 1** of this document. The data dictionary for the program is a version-controlled document (version 20191217-0 at the time of this writing). **Appendix 1** is arranged alphabetically to assist the reader when searching for specific fields. The field definitions and valid values are periodically updated by the Program's Data Working Group. **Important note:** The data dictionary presented in the previous technical memo, Hourigan et al. (2015), is superseded by the current version. For the most current version of the data dictionary, users should consult the DSCRTP web portal at: *https://deepseacoraldata.noaa.gov/internal-documents/program-guidance/science-team-guidance-for-data-management/20170707.xlsx/view*. The definitions posted at that location will always supersede all other versions (including the one published in this document).

Table 2a: Public fields of NOAA's National Database for Deep-Sea Corals and Sponges grouped by "Field Category." The single quoting convention used in this document is not used in the table below for enhanced visual clarity.

DSCRTP Category	Field Name	# of Fields
Survey Data	Ocean | SurveyID | Vessel | PI | PIAffiliation | Purpose | SurveyComments | OtherData | WebSite	9
Event Data	LargeMarineEcosystem | Country | FishCouncilRegion | Locality | DepthMethod | MinimumDepthInMeters | MaximumDepthInMeters | LocationComments | Station | EventID | SamplingEquipment | VehicleName | StartLatitude | StartLongitude | EndLatitude | EndLongitude | LocationAccuracy | NavType	18
Observation Data	SampleID | TrackingID | ImageURL | HighlightImageURL | AssociatedSequences | Latitude | Longitude | DepthInMeters | ObservationDate | ObservationYear | ObservationTime | OccurrenceComments | RecordType	13
TaxonomicID	ScientificName | VerbatimScientificName | VernacularNameCategory | VernacularName | TaxonRank | AphiaID | LifeScienceIdentifier | Phylum | Class | Subclass | Order | Suborder | Family | Subfamily | Genus | Subgenus | Species | Subspecies | ScientificNameAuthorship | TypeStatus | OperationalTaxonomicUnit | Morphospecies | CombinedNameID | Synonyms | IdentificationComments | IdentifiedBy | IdentificationDate | IdentificationQualifier | IdentificationVerificationStatus	29
Environment	Habitat | Substrate | CMECSGeoForm | CMECSSubstrate | CMECSBiotic | Temperature | Salinity | Oxygen | pH | pHscale | pCO2 | TA | DIC	13
Occurrence Detail	SampleAreaInSquareMeters | footprintWKT | footprintSRS | IndividualCount | CategoricalAbundance | Density | Cover | VerbatimSize | MinimumSize | MaximumSize | WeightInKg | Condition | AssociatedTaxa	13
GIS Enhancement	gisCRMDepth | gisGEBCODepth | gisEtopoDepth | gisMEOW	4
Metadata	Citation | Repository | DataProvider | DataContact | Modified	5
Record Keeping	DatabaseVersion | DatasetID | CatalogNumber	3
Flag	ShallowFlag	1

Table 2b: Private fields of NOAA's National Database for Deep-Sea Corals and Sponges grouped by "Field Category." The single quoting convention used in this document is not used in the table below for enhanced visual clarity.

DSCRTP Category	Field Name	# of Fields
Observation Data	VerbatimLatitude | VerbatimLongitude | ImageFilePath | HighlightImageFilePath	4
GIS Enhancement	gisLandCheck | gisNGIALocality | gisGEBCOLocality | gisNGIADist | gisGEBCODist | gisLME | gisIHOSeas | gisCaribbeanFMC | gisGulfOfMexicoFMC | gisMidAtlanticFMC | gisNewEnglandFMC | gisNorthPacificFMC | gisPacificFMC | gisSouthAtlanticFMC | gisWesternPacificFMC	15
Metadata	EntryUpdate	1
Record Keeping	EntryDate | Reporter | ReporterEmail | ReporterComments | AccessionID	5
Flag	Flag | FlagReason	2

New Database Fields

Several additions of internal and external fields, updates to field definitions, and updates to valid values have occurred since 2015. The Program recognizes the need for field definitions to remain as stable as possible through time, but the field definitions and valid values are revisited periodically to make additions or changes in order to capture critical new incoming information, or to clarify and refine the way a field should be implemented. Typical changes include adding a new field, improving the definition of an existing field, and changing the valid value set for an existing field. Highlighted below are the key changes to the database fields that have occurred since 2015, although values for all new fields have not been fully implemented for all records.

The database now has a field that indicates the current version of the database. This value is essential for reporting purposes, because it allows full reproducibility of results based on a specific version of the database.

- **'DatabaseVersion'**: Version of the entire database indicated as a date-based version in the format YYYYMMDD_. Example: '20190226_0' for the version created on February 26, 2019. The zero on the end indicates the iteration number on that day. If another version of the database was created on that same day, it would be indicated as '20190226_1'.

Four new taxonomy-related fields were added.

- **'VerbatimScientificName'**: The scientific name exactly as reported by the 'DataProvider' if different from the 'ScientificName'.
- **'TypeStatus'**: The type status (e.g., holotype, paratype, etc.), if any, of the specimen. Additional information (e.g., typified scientific name, publication) should be included in 'IdentificationComments'. The database currently contains over 700 records of holotypes.
- **'IdentificationVerificationStatus'**: IdentificationVerificationStatus - Score of the quality of the identification of a record with an associated image (see Howell et al. 2019). 1 = identified from image only, 2 = identified from image and physical specimens sampled from the same region, 3 = identified from image and that specific physical specimen.
- **'LifeScienceIdentifier'**: Life Science Identifiers are a way to name and locate pieces of information on the web. Essentially, an LSID is a unique identifier for some data, and the LSID protocol specifies a standard way to locate the data (as well as a standard way of describing that data). They are a little like DOIs used by many publishers. An LSID is represented as a uniform resource name (URN) with the following format: 'urn:lsid:Authority:Namespace:ObjectID[:Version]'. An example using WoRMS where genus = 'Thesea': 'urn:lsid:marinespecies.org:taxname:125315'. When searched on common web search services, this LSID provides a clear link to this resource *http://www.marinespecies.org/aphia.php?p=taxdetails&id=125315*. This value is assigned by DSCRTP and uses the WoRMS database as the primary authority.

Several new fields have been added that relate to capturing and reporting morphological descriptions of taxa and highlight images.

- **OperationalTaxonomicUnit**: An operational taxonomic unit (OTU) is an operational identifier used to classify groups of closely related individuals. OTU is a pragmatic identifier to group individuals by similarity, equivalent to, but not necessarily in line with classical Linnaean taxonomy or modern evolutionary taxonomy. This field does not currently have a recommended set of valid values and can be used differently by separate research groups. See related variables, 'Morphospecies' and 'CombinedNameID'.
- **Morphospecies**: A taxonomic species characteristic determined based wholly on observed morphological differences from related species. This field allows the 'DataProvider' to record the the extra detail needed to distinguish among different looking organisms of the same 'ScientificName'. For example, it is common within a given survey that a coral or sponge is clearly identifiable to the 'TaxonRank' of 'genus'. By visual inspection the analyst may also notice morphological differences that are evident, but not diagnostic enough to assigned to 'TaxonRank' of 'species. These morphological notations, e.g., msp1, msp2, msp3, or in the case of sponges encrusting, vase, fig, sponge, massive globose, etc., may be placed in 'Morphospecies'. This field does not currently have a recommended set of valid values and can be used differently by separate research groups. See related variables, 'Morphospecies' and 'CombinedNameID'.
- **CombinedNameID**: A simple concatenation of 'ScientificName' and 'Morphospecies'.
- **HighlightImageURL**: `URL that points to the highlight image hosted on the DSC portal or appropriate archive that displays the specimen of record. A highlight image is one that is an especially good representative image of the taxa as determined by the 'DataProvider'. In the

case of records submitted to the Program, this record is based on the image file indicated in the HighlightImageFilePath variable. This value is created by the Program. If this image is null, then it is assumed that the imagery associated with this record is not particularly notable. This image URL can be the same as that reported in 'ImageURL', but does not necessarily have to be the same image.

- **HighlightImageFilePath**: `Full file path and file name pointing to an image of the occurrence. This field should only be populated if the image is an especially good representative image for the taxa of interest. The image can be the same or different than that reported in ImageFilePath. The file path and name will provide the Program with enough information to link the image to the record. Two or more images can be submitted per observation, separated by the pipe character with a leading a trailing space, ' | '. For example, an in-situ and on-deck image of the same organism. If multiple images are submitted, the first image in the list should be the one that you want displayed on the web mapping portal. This primary image will be displayed when a web mapping portal user clicks on a coral or sponge point in the map.

The previous 'Size' field was changed to 'VerbatimSize' and two new fields were added to capture the minimum and maximum size of the individual corals and sponges where these measurements are provided.

- **VerbatimSize**: This field is used to store the originally reported Size value if they were not in our required format. If the data provider is available to provide values in our required format, then that is preferable.
- **MinimumSize**: Minimum size of the taxon observed. Methods for determining size (e.g., height or maximum diameter) should be noted in 'OccurrenceComments'.
- **MaximumSize**: Maximum size of the taxon observed. Methods for determining size (e.g., height or maximum diameter) should be noted in 'OccurrenceComments'.

The following fields were added to capture key habitat characteristics using standardized terminology from the Coastal and Marine Ecological Classification Standard (CMECS) *https://www.cmecscatalog.org/cmecs/index.jsp*.

- **CMECSGeoForm**: Tectonic or geologic features, such as shelf, canyon, or seamount. See the Coastal and Marine Ecological Classification Standard catalog for description *https://www.cmecscatalog.org/cmecs/index.jsp*.
- **CMECSSubstrate**: Geologic or biogenic substrate. See the Coastal and Marine Ecological Classification Standard catalog for description *https://www.cmecscatalog.org/cmecs/index.jsp*.
- **CMECSBiotic**: Biotic community present at the observation site. See the Coastal and Marine Ecological Classification Standard catalog for description *https://www.cmecscatalog.org/cmecs/index.jsp*.

The 'EntryUpdate' field was added to capture the most current date that the Program modified the record. This is in addition to the field called 'Modified', which captures the last date that the data provider modified the record.

- **EntryUpdate**: The most recent date the DSCRTP Program updated or verified the record. Format: YYYY-MM-DD. This variable will only contain a value if the DSCRTP Program has updated a record which has previously been published in the database. Also see the related variables that inform record versioning: 'Modified' and 'EntryDate'.

We added the field 'ShallowFlag' to flag occurrences that are less than 50 meters deep. These occurrences are included in our database because they represent deeper taxa that can also occur at depths shallower than our typical 50 meter cutoff.

- **ShallowFlag**: A value of 1 denotes a coral or sponge record where DepthInMeters is correct but it is less than 50 meters. A value of 0 denotes a coral that has a correct DepthInMeters greater than or equal to 50 meters.

Several new spatially calculated depth fields were added to help ground truth the observed depths within the database.

- **gisCRMDepth**: GIS derived depth (bathymetry is Coastal Relief Model).
- **gisGEBCODepth**: GIS derived gisGEBCODepth. The GEBCO_2019 grid is used. It is a global terrain model for ocean and land at 15 arc-second intervals. https://www.gebco.net/data_and_products/gridded_bathymetry_data/
- **gisEtopoDepth**: GIS derived depth (bathymetry is ETOPO1).

The following location-related internal fields were added to enhance the ability to describe and standardize place names.

- **gisNGIALocality**: GIS derived Locality (based on NGIA point placenames) - 30km max search distance for matching.
- **gisNGIADist**: Distance to nearest official NGIA locality (km).
- **gisGEBCOLocality**: GIS derived Locality (based on GEBCO point placenames) - 30km max search distance for matching.
- **gisGEBCODist**: Distance to nearest official GEBCO locality (km).

The following fields have been added to capture the actual spatially explicit boundary of the sample area footprint for each record. These fields are mostly unpopulated in the current version (20191217-0) of the database.

- **SampleAreaInSquareMeters**: The area in square meters, if projected vertically to a plane representing the seafloor surface, of the sample space in which the observation was made. Use 'footprintWKT' field to specify the exact shape of the sampling space.
- **footprintWKT**: A Well-Known Text (WKT) representation of the shape (footprint, geometry) that defines the sample area. As an example, for downward looking or oblique still images, this would be a polygon representing the footprint of the viewable area (an orthogonal projection using the sea floor as the projection plane). For a video transect, this would be a polygon that represents the footprint of the entire viewable area throughout the length of the video. To report the actual area of the sample, see 'SampleAreaInSquareMeters'. Also please report 'footprintSRS' if you report 'footprintWKT'. Example: the one-degree bounding box with opposite corners at (longitude=10, latitude=20) and (longitude=11, latitude=21) would be expressed in well-known text as POLYGON ((10 20, 11 20, 11 21, 10 21, 10 20)). For discussion see http://terms.tdwg.org/wiki/dwc:footprintWKT.
- **footprintSRS**: A Well-Known Text (WKT) representation of the Spatial Reference System (SRS) for the footprintWKT of the Location. Do not use this term to describe the SRS of the Latitude and Longitude, even if it is the same as for the footprintWKT. 'footprintSRS' must always be populated if 'footprintWKT' is given. Example: The WKT for the standard WGS84

SRS (EPSG:4326) is
"GEOGCS["GCS_WGS_1984",DATUM["D_WGS_1984",SPHEROID["WGS_1984",6378137,298.257
223563]],PRIMEM["Greenwich",0],UNIT["Degree",0.0174532925199433]]" without the
enclosing quotes. For discussion see *http://terms.tdwg.org/wiki/dwc:footprintSRS*.

The 'ScientificName' field definition has been updated. From this writing forward, the Program will
no longer require the addition of 'sp.' to incoming 'ScientificName' entries where 'TaxonRank' is at
the genus level. This change was made to make our taxonomic naming conform to community
standards.

Key Notes on Database Fields for Data Users and Providers

This section provides additional details regarding the database structure to assist both users and
data providers in the proper creation and interpretation of the occurrence records within the
National Database.

SampleID vs. TrackingID vs. CatalogNumber

The 'SampleID' field is meant to be used by the data provider as an identifier to specify the exact
observation, image, or physical specimen from which the occurrence record was created.
'SampleID' is a key value that allows Program records to be traced back to the data provider's
original observation. For example, for museum specimens, the 'SampleID' would be the museum's
catalog number. The purpose of the 'SampleID' field is for end-users of the database to have a link
back to the original data provider's observation records. Therefore the 'SampleID' must be a
persistent identifier, internal to the data provider's institution or working group, that ensures
individual physical samples or imagery records will be located within their own data storage
systems. If appropriate, this 'SampleID' field can be repeated within the National Database records.
For example, a single still image that contains two taxa can be properly represented as two separate
records in the database with the same SampleID. The 'TrackingID' can be used as a secondary
identification field, e.g., for a field ID number. The 'TrackingID' field can be used, in addition to the
'SampleID' field, when two separate IDs have been used for a single sample. If only one identifier is
available, use only 'SampleID'. In contrast, the 'CatalogNumber' field is a completely unique
identifier created and used by the Program to locate individual records within the National
Database. The 'CatalogNumber' is assigned by the Program itself, and could number in the
hundreds of thousands, or millions one day. These numbers are used to uniquely identify and
reference specific occurrence records only within the context of the National Database.

Geographic Position

Published database records have positional information in the 'Latitude' and 'Longitude' fields. The
positional information can reference a point or a line. When only a single coordinate is given to
represent a linear sampling event (e.g., a transect), that point is typically the start point or
midpoint, with details provided in the 'LocationComments' field. When available, some records
collected along transects or during trawl surveys will also have information in the following fields:
'StartLatitude', 'StartLongitude', 'EndLatitude', and 'EndLongitude'. Other fields in the database
describe details about the quality of the location data: 'NavType', 'LocationAccuracy', and
'LocationComments'. 'LocationAccuracy' specifically, is important to understand uncertainty in
spatial modeling and guide users on the spatial resolution of the occurrence record. Additionally,

fields about the named locality of each occurrence are included based on automated spatial analysis routines: 'Ocean', 'LargeMarineEcosystem', 'gisMEOW', 'Country', and 'FishCouncilRegion'. 'Locality' is the non-standardized named location that the data provider listed in the initial data submission. Well Known Text (WKT) representations of the geographic position of occurrences can be recorded in the pair of variables: 'footprintWKT' and 'footprintSRS' (for a deeper discussion of these variables and the quantification of the sample area, see the discussion below in the section: The Sample Area).

The Sample Area

The area sampled by a survey provides important information relevant to understanding the occurrence records in the database. Each record in the database represents an observation of one or more occurrences of a single taxon within a given sample area (in square meters). The sample area will vary based on whether the observation is an actual physical sample or imagery (both still images and video), or other types of direct observation (e.g., directly observed from a submersible or by trawl sampling). Sampling area means the viewable area (or trawled area), orthogonally projected onto the sea floor, represented by a polygon. The sample area could denote the area observed in a single still image, multiple still images, a video segment, or, in the case of physical samples, sometimes a single discrete point in space. In the case of trawl records, the sample area is the area trawled, which is typically determined by trawl width and length of the trawl haul. All details regarding the methods used for the determination of horizontal position and sample area will be reported in the variable 'LocationComments'.

The Program has instituted several new database fields to better capture the sample area. The sample area can be reported in the field 'SampleAreaInSquareMeters'. If the exact georeferenced footprint of the polygon representing the sample area is known, it can be reported in the pair of fields called 'footprintWKT' and 'footprintSRS'. Most major GIS software provides functionality to convert and export Well-Known Text representations of sampling polygons. The Program would like to obtain these explicit polygonal representations of the sampling area for entire video transects or for single images where possible (example: footprints of georeferenced, downward-looking photos). This type of explicit representation of the sample area will allow a better quantification of sampling effort in a given location.

For convenience and simplicity, each record in the database, regardless of sample area shape or size, is represented on the web mapping portal as a single geographically-located point. Geographic position of this single point is indicated in the 'Latitude' and 'Longitude' fields.

The Program places a high value on detailed quantification of the sample area, as it helps quantify total sampling effort and is essential for analysis of the density of coral and sponge communities.

Connection between the Ocean Biogeographic Information System, USA (OBIS-USA), the Global Biodiversity Information Facility (GBIF), and the National Database

The Program is committed to contributing to national and international aggregations of biological occurrence data. To this end, the Program contributes updated coral and sponge data quarterly to *OBIS-USA*, the U.S. node of OBIS, which is part of the International Oceanographic Data and

Information Exchange (IODE) of the Intergovernmental Oceanographic Commission (IOC) of the United Nations Educational, Scientific and Cultural Organization (UNESCO). OBIS-USA is coordinated by the Science Analytics and Synthesis (SAS) Program of the United States Geological Survey (USGS). The National Database contributes a particular subset of fields from its database to OBIS-USA during quarterly submissions, and these fields align to the *Darwin Core standard*, the standard used by OBIS-USA and GBIF. OBIS-USA is also a participant node of the Global Biodiversity Information Facility, *GBIF*. **Table 3** shows the National Database fields that are contributed to these global databases. The data that is contributed to OBIS-USA is synchronized at GBIF as well. These data sharing efforts are intended to increase the global awareness, availability, and interoperability of the National Database.

- Direct links to the National Database on *OBIS* and *GBIF.*

*Table 3: Fields from the National Database for Deep-Sea Corals and Sponges that are contributed to the Ocean Biogeographic Information System, USA (OBIS-USA) and the Global Biodiversity Information Facility (GBIF). The corresponding DarwinCore Term is given that matches the term used within the National Database (DSCRTP Field Name). In addition to the fields listed below, a **'references'** field is created, which is a URL that provides a direct link to the full record within the National Database.*

DSCRTP Field Name	DarwinCore Term
ScientificName	scientificName
AphiaID	scientificNameID
CatalogNumber	occurrenceID
IndividualCount	individualCount
ObservationDate	eventDate
Latitude	decimalLatitude
Longitude	decimalLongitude
LocationAccuracy	coordinateUncertaintyInMeters
MinimumDepthInMeters	minimumDepthInMeters
MaximumDepthInMeters	maximumDepthInMeters
RecordType	basisOfRecord
ImageURL	associatedMedia

Summary of Database Contents

The following figures and statistics provide insight into the completeness and quality status of data sets included within the National Database. These results are meant to inform and caution data users about the adequacy of using the database for various analytical objectives. **Appendix 3** includes a link to an online resource (data set dashboards) that describes each unique data set in detail as well as providing an interactive map to explore those data in isolation. The data set dashboards also have a link to download the data set as well as International Standards Organization compliant metadata (ISO 19115-Parts 1 and 2).

The figures and tables that are presented in the following section are designed to give readers greater insight into the current contents and status of a few key fields in the database. For clarity and brevity, not all variables within the database are covered here. We elected to summarize fields and highlight relationships that clarify the distribution of data holdings and well as identify issues that data users may find useful for their analytical goals.

Analysis Details

The analyses presented here use version 20191217-0 of the database. In some cases, the data have been filtered to enhance readability or interpretation of a figure or table. These filters will be called out explicitly in the text or captions when used. The purpose of this analysis is to succinctly summarize the current contents and quality of the database, so that any major limitations are apparent to users. The numerical and graphical summaries provided in this document were created using R and R-Markdown used within the RStudio development environment (R Core Team, 2017; RStudio Team, 2016). All R code is shared on GitHub at *https://github.com/RobertMcGuinn/deepseatools*.

Distribution of Records by U.S. Fishery Management Council Region

The primary geographic focus of NOAA's Deep Sea Coral Research & Technology Program is in the U.S. EEZ. Primary clients for the National Database are the United States regional fishery management councils. The 'FishCouncilRegion' field identifies the council region (and adjacent state waters) within which a record occurs. **Figure 2** shows the distribution of records by region. The large number of records from the Pacific Council Region (i.e., U.S. West Coast) reflects our partnership with the Monterey Bay Aquarium Research Institute (MBARI). We also recognize the scientific value of data from around the world and hope that our database can serve as an international resource. Currently, around 7% of records occur outside U.S. fishery management council regions, either in international waters (areas beyond national jurisdiction) or within the boundaries of other countries' EEZs (**Figure 2**).
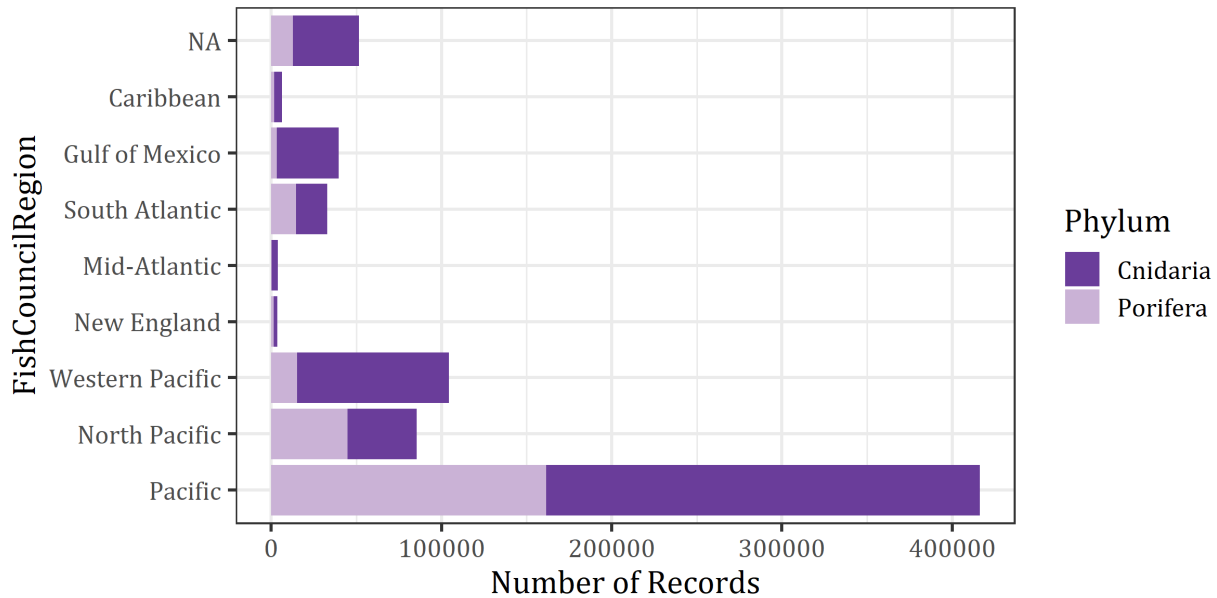
Figure 2: Distribution of known coral and sponge occurrences grouped by fishery management council region ('FishCouncilRegion'). The number of records in a given region is primarily a measure of sampling effort and data reporting to the National Database. The counts are not meant as a reflection of the abundance of deep-sea corals or sponges within a region. 'NA' = records outside the U.S. EEZ.

Images Associated with Records

As of the current version, approximately 10% or 75,014 records have full resolution still images available. There are 56,591 unique images (i.e., some images contain multiple taxa). These images are viewable directly in the DSCRTP map portal by clicking on coral or sponge points. The 'ImageURL' field also contains the link to the image if available, and this field is available in data downloaded from the DSCRTP map portal.

Depth Distribution

'DepthInMeters' is the primary depth field in the database with all of the published records including depth information. This field records either the actual reported depth, the average depth (when separate 'MinimumDepthInMeters' and 'MaximumDepthInMeters' are reported), or, in cases where no depth is reported, a GIS-derived depth for their position is assigned, relative to the best available bathymetry. These spatially derived depths are available in the following fields: 'gisCRMDepth', 'gisGEBCODepth', and 'gisEtopoDepth' and provide an important means of checking depth. Averaged and assigned depths should be treated as approximations.

The 'DepthMethod' and 'LocationComments' fields contain additional details on the source and methodology for depth values present in the 'DepthInMeters' field. For full definitions and valid values of these fields, see **Appendix 1**. See **Figure 3** below for a distribution of depth values in the National Database by fishery management council region ('FishCouncilRegion'). It is important to note here again: Distribution of depth values represented in **Figure 3** are a result of sampling effort

and subsequent reporting and not necessarily a reflection of regional depth distributions of corals and sponges.
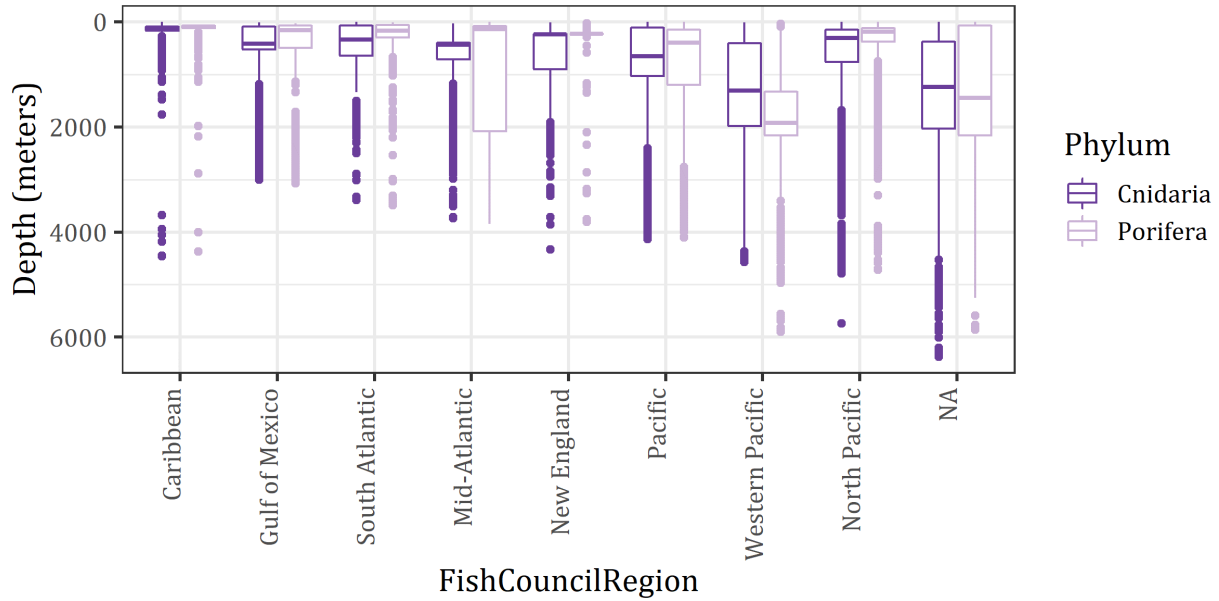

Figure 3: Distribution of depth values by fishery management council region ('FishCouncilRegion'). The 'NA' values are records that are outside of the U.S. EEZ.

Observations through Time

The year in which an observation was made, 'ObservationYear', provides some insights for understanding deep sea exploration activities through time. **Figure 4** provides a visualization of the distribution of the 'ObservationYear' field. The database includes records from 1833 to 2018, with the median year of observation being 2008. Most records prior to 1980 represent museum specimens. The large increase in 21st century records reflects the increasing importance of records annotated from still images and video.
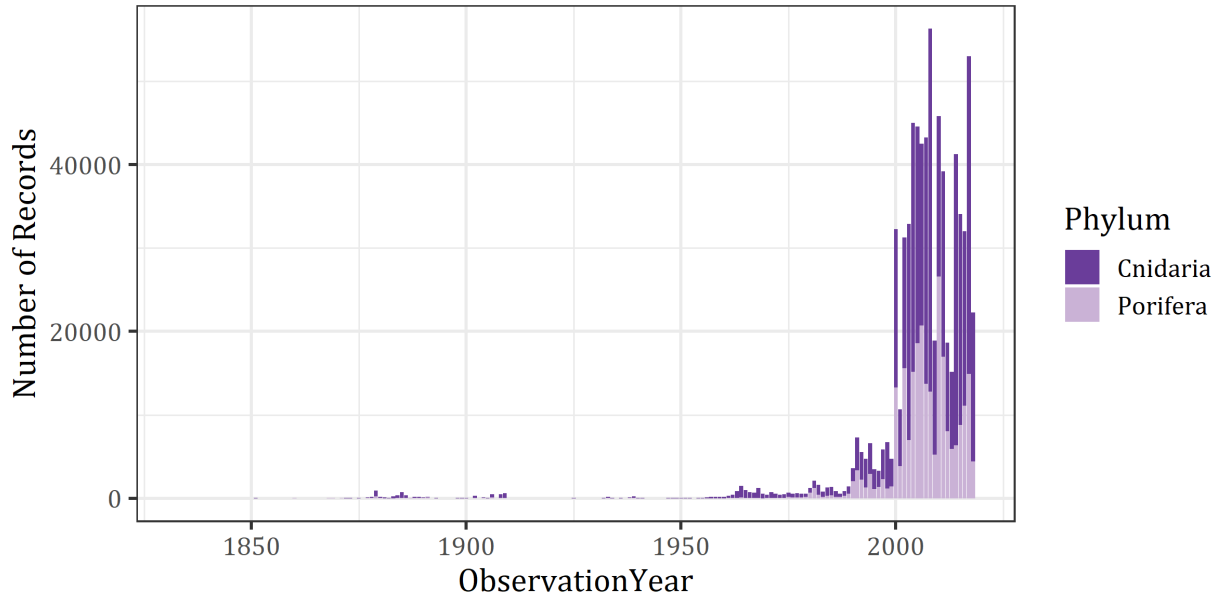
Figure 4: The temporal distribution of coral and sponge observations by 'ObservationYear'.

Individual Count

The 'IndividualCount' field indicates the number individual corals or sponges per record; 66 % or 492,441 records have 'IndividualCount' values. Most of the records fall within the range of 1-50 individuals per observation frame, but the maximum is over 16,000 individuals. The median value of 'IndividualCount' is 1. See **Figure 5** for the distribution of individual count between 0 and 50.

When interpreting individual count values ('IndividualCount'), it is important to consider the geographic extent of the sampling events. The database includes sampling frames that are as small as a single specimen or photograph and as large as a video transect or the area sampled during a trawl haul. For many records, the size of the sampling frame is unknown. The 'SampleAreaInSquareMeters' is a new field designed to capture the area of the sampling frame, but at the current database version (20191217-0) only 1% of the records are populated for this variable. Where populated, this value can be used to spatially normalize the count values by area for the purpose of making calculations of density. Also, when 'IndividualCount' is not populated, species abundance values may be reported in 'CategoricalAbundance' instead, which is simply a categorical version of the count of the number of individuals counted within the sampling frame. The valid values for 'CategoricalAbundance' are: present; 1; 2-10; 11-20; 21-50; 51-100; >100; NA. Additional fields accommodate other measures of taxon abundance can be expressed as 'WeightInKg' (used most often for reporting trawl survey results) and 'Cover' (often used for framework-forming corals such as *Lophelia pertusa* where individual colonies cannot be distinguished and counted), and 'Density'.
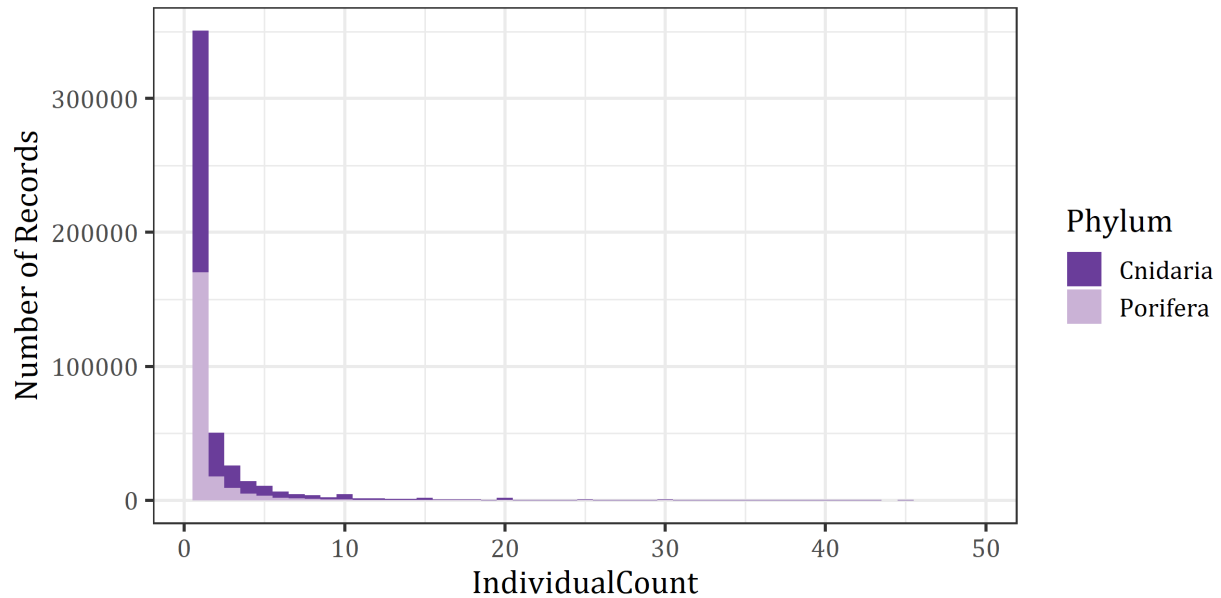
27

Figure 5: Distribution of 'IndividualCount' values between 0 and 50 corals or sponges. The figure's x-axis is truncated at 50 individuals for visual clarity.

Size

Size of coral colonies and sponges is an important measure that may be associated with a structure-forming organism's value as habitat for other species. About 24% of all records have reported size values (n = 178,526) In the database, size values are reported in fields called 'MinimumSize' and 'MaximumSize' fields in order to capture a range of sizes that might be observed within given sampling area. In the case of a video transect, for example, multiple sizes may be observed for a single taxon, so reporting a range is appropriate. Even in a single still image, you may observe a range of sizes for a single taxon. See **Figure 6** below for distribution of maximum size values. The original size values, exactly as reported by the data provider, are captured in a field called 'VerbatimSize'. The maximum size reported in the database for corals is 762 cm and for sponges is 600 cm.
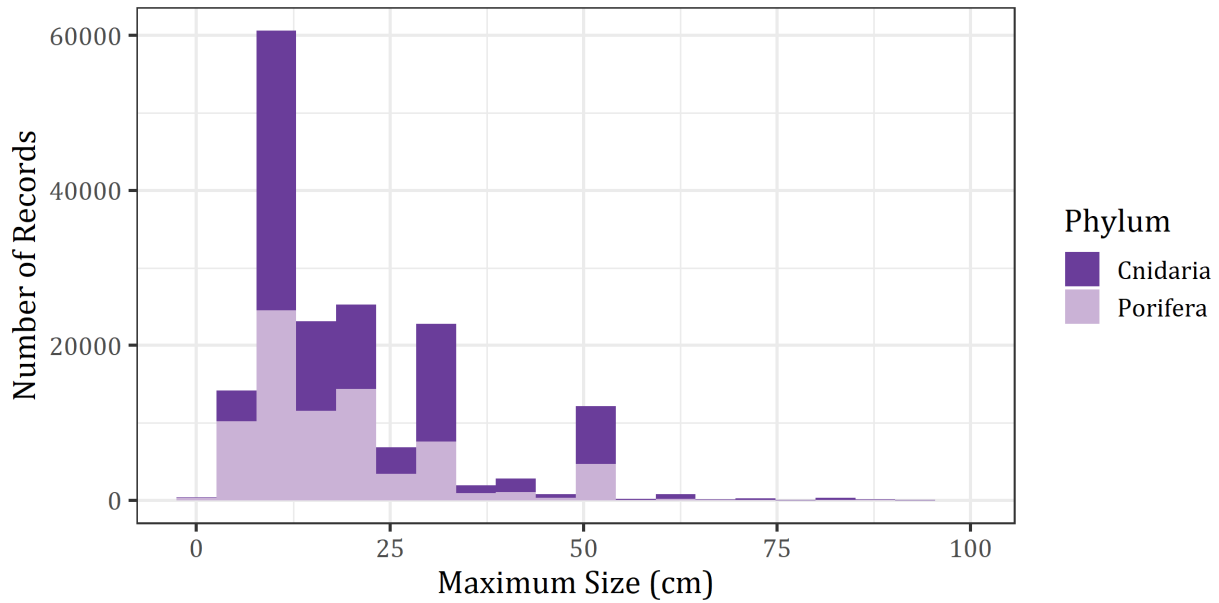
Figure 6: Histogram of 'MaximumSize' values. For clarity the x-axis limit has been set to 100 cm.

Condition

Approximately 61% of all records have 'Condition' values reported. The currently valid values for Condition are: Live; Damaged; Dead; NA. Most of the records that have 'Condition' reported have been reported as 'Live'. It is likely that this is influenced by a strong sampling and observation bias for living coral and sponges, along with a difficulty in discerning, or inconsistency in reporting, damage (**Figure 7**).

Figure 7: Bar plot of 'Condition' values, showing most records in the database are of live coral and sponges.

Habitat and Substrate

Habitat and substrate values are in the database, verbatim, as reported by the data provider. No standard valid values have been adopted or implemented for these fields. Three newly-established fields using the Coastal and Marine Ecological Classification Standard (CMECS) classifications (CMECSBiotic, CMECSGeoForm, CMECSSubstrate) are intended to help standardize reporting of habitat measures in the fu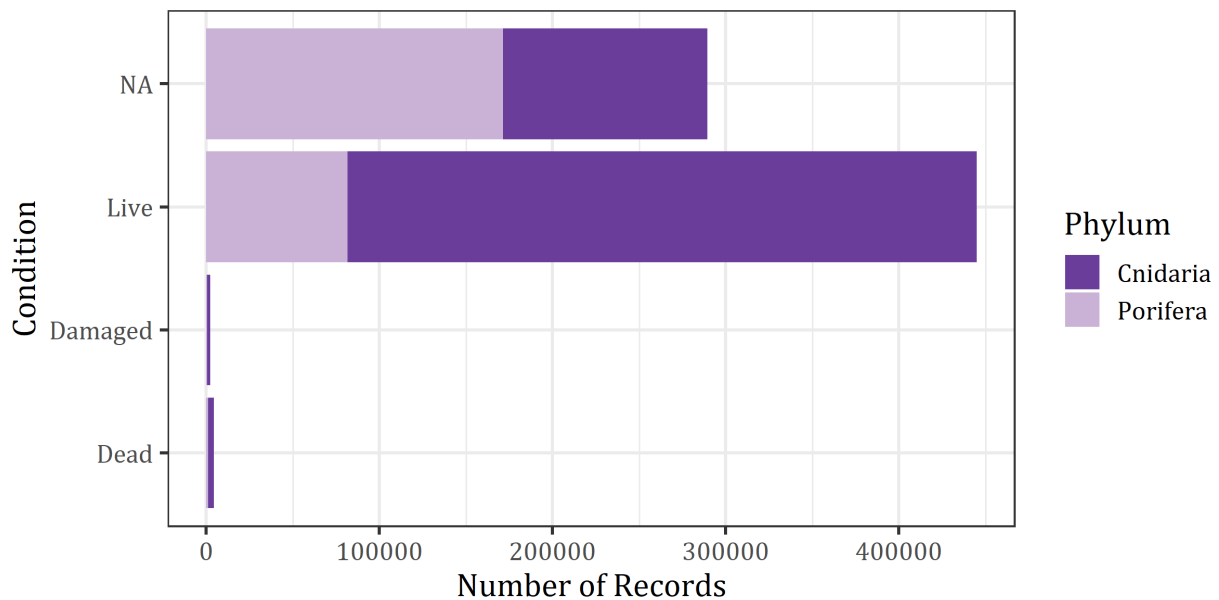ture. See the following URL for more information about CMECS. *https://www.natureserve.org/conservation-tools/coastal-and-marine-ecological-classification-standard-cmecs*

- Records with 'Substrate' values: 21% or 152,613
- Records with 'Habitat' values: 43% or 319,361

Taxonomy

All records in the database are populated with a complete taxonomic string based on the *World Register of Marine Species (WoRMS)*. In addition, greater information about taxonomic identifications are available in the following descriptive fields: IdentificationComments, IdentifiedBy, IdentificationDate, IdentificationQualifier, IdentificationVerificationStatus. See **Appendix 1** for the full definitions of these fields. Taxonomic identifications can change through time, either through refinements in identification by taxonomists or through changes to underlying group taxonomy. The Program attempts to keep up with these changes, but these fields should be checked carefully before using the data for analysis.

'TaxonRank' provides taxonomic resolution of the identifications in the database (**Figure 8**) Most corals and sponges are identified to the taxon rank ('TaxonRank') of genus or species.
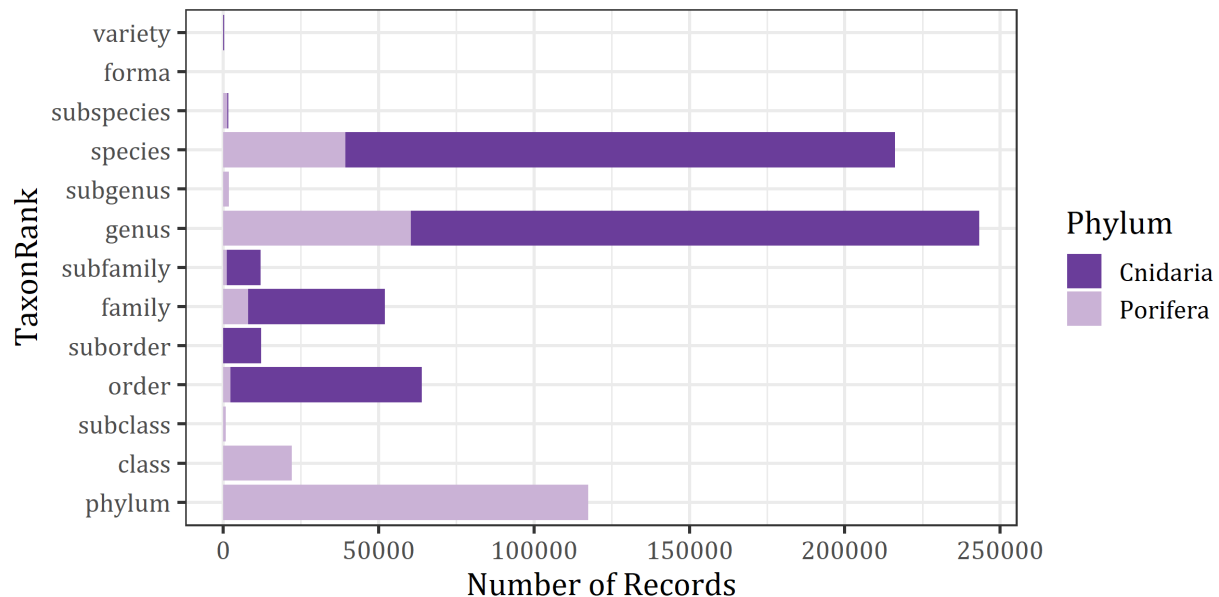
Figure 8: Distribution of 'TaxonRank' across all records. Corals are only included in the database if they are identified at least to Order, as higher taxonomic levels contain cnidarians that are not considered to be corals. Sponge records are accepted at the phylum level and below.

Record Types

100% of all records (n = 743,758) have 'RecordType' values reported. **Figure 9** shows the distribution of these values across the entire database.
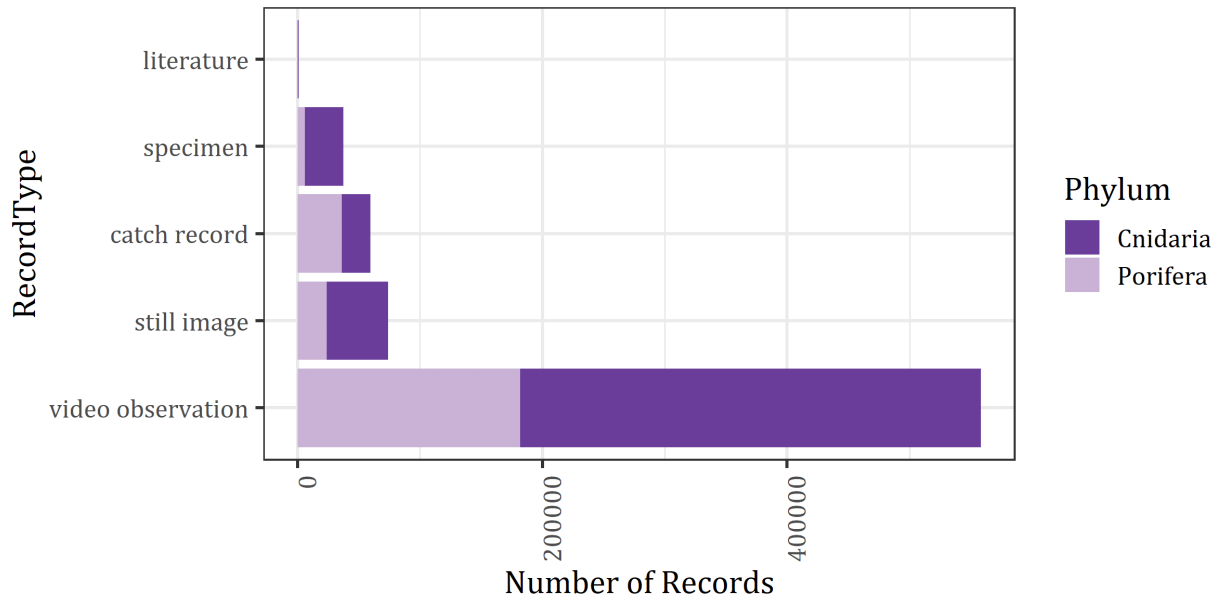
Figure 9: The distribution of valid 'RecordType' values across the entire database.

Sampling Equipment

Approximately 98% of all records (n = 727,029) have 'SamplingEquipment' values reported. Additional information on the sampling equipment (e.g., specific type of trawl) and other methods are often included in the 'SurveyComments' or in the data set metadata. See **Figure 10** below for the makeup of the entire database by 'SamplingEquipment'.
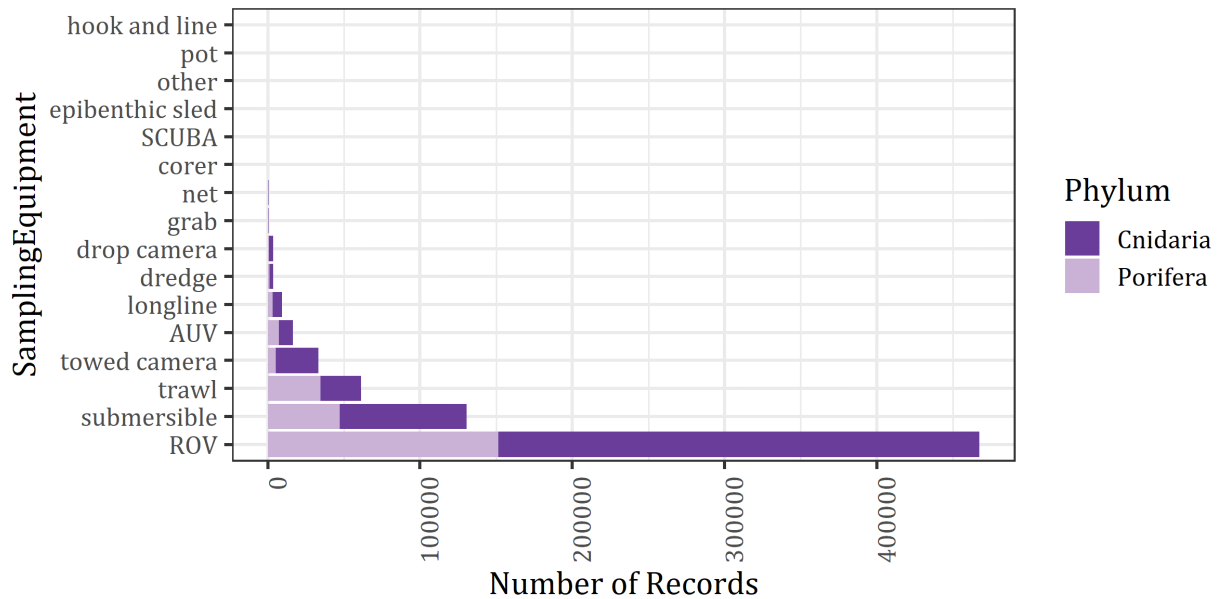
Figure 10: The distribution of 'SamplingEquipment' values across the entire database.

Caveats Regarding Use of the Data

Data users should be cautious when using these data for specific analytical purposes, especially when using data from more than one sampling effort at a time. As with all data aggregations, meta-analyses of disparate data are complex and many pitfalls can arise when combining data from sources that use different sampling methods. Users of the data are encouraged to carefully consider which data sets (and even which records within a data set) are appropriate to use for the analytical purpose at hand, especially in regard to evaluating identification and positional uncertainty. The Program has made considerable effort to assess the quality of data included in the National Database and to exclude data that are not up to standard. Additionally, numerous metadata fields are included to provide users with additional information that can be used to evaluate the data. Despite those efforts, data quality issues do still exist in the database and users should conduct their own quality checks prior to relying on the database for particular analyses and management applications.

Conclusion

The Program's goal is to provide broad public access to a comprehensive, consolidated, quality assured data resource for deep-sea corals and sponges in the U.S. and around the world. It is our intention that all data providers receive the credit they deserve for helping build this resource, which is designed to enhance deep-sea conservation and management decision making. By creating the National Database for Deep-Sea Corals and Sponges and the data portal at DeepSeaCoralData.NOAA.gov, the Program has provided a comprehensive and convenient one-stop source to view and access these data.

This report has summarized the current holdings and quality level of NOAA's National Database of Deep-Sea Corals and Sponges in order to enhance understanding, appropriate use, and to indicate areas for further improvement. The Program welcomes feedback from portal and data users. Please send us your comments and questions to *deepseacoraldata@noaa.gov*. This report, and any follow-on updates, will continue to track the ongoing evolution of this key resource for deep sea ecologists, biologists, managers, policy makers, and explorers.

Acknowledgements

The Program would like to thank all the individuals and institutions who have contributed their data to be published in the common framework of the National Database. We recognize the considerable time and effort that it takes to collaborate with us, and for that, we are grateful. Also, thanks to all the contributors to the conceptualization and design of the database and the deep sea coral data portal. This resource is only possible with all of the individual and institutional contributors from the deep-sea research community. We would also like to thank all of the users of the data and portal for providing the real-world application of these resources. A huge thanks to all of the reviewers for their insights and improvements to this document (in alphabetical order by last name): Rachel Bassett, Abigail Benson, Just Cebrian, Elizabeth Duncan, Meredith Everett, Enrique Salgado, and Curt Whitmire. Thanks to all who continue to help us evolve these information resources to enhance deep-sea conservation and management.

References

Amante, C. and B.W. Eakins (2009) ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA. *doi:10.7289/V5C8276M* [access date].] Available from *https://www.ngdc.noaa.gov/mgg/global/*.

Horton, T. et al. (2018) World Register of Marine Species. Available from *http://www.marinespecies.org* at VLIZ. Accessed 2019-12-10. *doi:10.14284/170*.

Hourigan, T.F., P.J. Etnoyer, R.P. McGuinn, C. Whitmire, D.S. Dorfman, L.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. 27 pp. Silver Spring, MD. Available from *https://deepseacoraldata.noaa.gov/internal-documents/program-guidance/science-team-guidance-for-data-management/database_intro*.

Hourigan TF, Etnoyer PJ, Cairns SD (2017) The State of Deep-Sea Coral and Sponge Ecosystems of the United States. NOAA Technical Memorandum NMFS-OHC-004 Silver Spring, MD 467 pp. *https://deepseacoraldata.noaa.gov/library/2017-state-of-deep-sea-corals-report*.

Howell KL, Davies JS, Allcock AL, Braga-Henriques A, Buhl-Mortensen P, Carreiro-Silva M, et al. (2019) A framework for the development of a global standardised marine taxon reference image database (SMarTaR-ID) to support image-based analyses. PLoS ONE 14(12): e0218904. https://doi.org/10.1371/journal.pone.0218904.

R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from *https://www.R-project.org/*. RStudio Team (2016) RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Available from *http://www.rstudio.com/*.

Scanlon, K.M., R.G. Waller, A.R. Sirotek, J.M. Knisel, J.J. O'Malley, and S. Alesandrini (2010) USGS cold-water coral geographic database-Gulf of Mexico and western North Atlantic Ocean, version 1.0: U.S. Geological Survey Open-File Report 2008-1351. Available from *https://pubs.usgs.gov/of/2008/1351/*.

Appendix 1: Data Dictionary for the National Database

National Database field names, definitions, and their valid values, organized by field name. This table includes only the public fields in the National Database. The fields are arranged alphabetically for ease of use. The current version of this table is also available as an Excel file on the DSCRTP web portal: *https://deepseacoraldata.noaa.gov/internal-documents/program-guidance/science-team-guidance-for-data-management/20170707.xlsx/view*.

Important Note on Data Dictionary Version: For the most current version available, access the link above. The version printed here may be out of date due to post-publication updates of the schema. Users are advised to use the latest Excel version of the table for preparing a new data for submission or for analytical guidance.

Field Name	Field Definition	Valid Values
AphiaID	AphiaID of 'ScientificName' from World Register of Marine Species (WoRMS). Access these numeric codes are here: http://www.marinespecies.org/index.php.	[value]; -999
AssociatedSequences	Additional ID of samples where genetic sequences have been submitted to GenBank, for example. Multiple sequences may be separated by the pipe character with a leading a trailing space, '	'.	[value]; NA
AssociatedTaxa	Notable organisms that co-occur with the coral or sponge observation. A list of organisms may be separated by the pipe character with a leading a trailing space, '	'. Commercially fished species are of particular interest to regional fishery management councils. Methods for determining association (e.g., 'touching' or 'fish within one meter') should be noted in 'OccurrenceComments'.	[value]; NA
CatalogNumber	Unique record identifier assigned by the DSCRTP. It is persistent and the numbers are retired if records are deleted from the database.	[value]; -999
CategoricalAbundance	Abundance category of corals/sponges of the same ScientificName at the sample location. If data is presence only, write 'present' in CategoricalAbundance and set IndividualCount = '-999'.	present; 1; 2-10; 11-20; 21-50; 51-100; >100; NA
Citation	A list (concatenated and separated) of identifiers	[value]

Field Name	Field Definition	Valid Values
	(publication, bibliographic reference, global unique identifier, URI) of literature associated with the Occurrence. Equivalent to the term associatedReferences in Darwin Core. Multiple citations should be separated by the pipe character with a leading a trailing space, '	'. The preferred style guide is Chicago/Turabian. Michigan State University Libraries publishes a helpful guide on how to cite data at the following link: http://libguides.lib.msu.edu/citedata.	
Class	The Class in which the taxon is classified.	[value]; NA
CMECSBiotic	Biotic community present at the observation site. See the Coastal and Marine Ecological Classification Standard catalog for description https://www.cmecscatalog.org/cmecs/index.jsp .	see CMECS
CMECSGeoForm	Tectonic or geologic features, such as shelf, canyon, or seamount. See the Coastal and Marine Ecological Classification Standard catalog for description https://www.cmecscatalog.org/cmecs/index.jsp .	see CMECS
CMECSSubstrate	Geologic or biogenic substrate. See the Coastal and Marine Ecological Classification Standard catalog for description https://www.cmecscatalog.org/cmecs/index.jsp .	see CMECS
CombinedNameID	A simple concatenation of 'ScientificName' and 'Morphospecies'.	[value]; NA
Condition	Condition of the organism (coral or sponge) when collected or observed. Use 'Damaged' when greater than 20 % of the colony is damaged. If both live and dead organisms are present, use 'Live'.	Live; Damaged; Dead; NA
Country	Country in whose Exclusive Economic Zone the observation or sample was collected , based on Exclusive Economic Zones Boundaries (EEZ) [World EEZ v7 (2012-11-20)] from http://marineregions.org/downloads.php .	[value]; NA
Cover	Percent of sampled area covered by organism of interest. This field is generally used for	[positive integers from 0-100], -999

Field Name	Field Definition	Valid Values
	framework-forming stony corals for which individual counts are difficult.	
DatabaseVersion	Version of the entire database indicated as a date-based version in the format YYYYMMDD_<iteration>. Example: '20190226_0' for the version created on February 26, 2019. The zero on the end indicates the iteration number on that day. If another version of the database was created on that same day, it would be indicated as '20190226_1'.	[value]
DataContact	The name and e-mail of the individual who is the primary representative of the 'DataProvider'. The preferred format is 'Last Name, First Name; e-mail'. 'DataProvider' and 'DataContact' may be the same person, if 'DataProvider' is an individual and not an institution.	[value]
DataProvider	Identifies the originator of the data. This is the institution, publication, or individual, who ultimately deserves credit for acquiring or aggregating the data and making it available. A single institution or individual is preferred multiple. When an individual is listed, use 'Last Name, First Name' format. Publications should be abbreviated as such, 'Pante et al. 2012'.	[value]
DatasetID	Standardized identifier for each dataset within the larger database. This field is used as the primary grouping variable for associating records together and for creating the dataset dashboards	[value]
Density	Number of individuals (or colonies) per square meter for an observation.	[value], -999
DepthInMeters	Best single depth value for sample as a positive value in meters.	[value]
DepthMethod	Select method by which best singular depth in meters (DepthInMeters) was determined. Place method details in 'LocationComments'. Select 'averaged' when start and stop depths were averaged. Select 'assigned' when depth was derived from bathymetry at the location. Select	averaged; assigned; reported

Field Name	Field Definition	Valid Values
	'reported' when depth was reported based on instrumentation or described in literature.	
DIC	Dissolved inorganic carbon, units should be micromoles per kilogram. Methods should be reported in 'OccurrenceComments'.	[value]; -999
EndLatitude	Ending latitude in decimal degrees for observations or samples collected over a distance (e.g., trawls, transects). Precision should be the same as that originally reported; ideally to five decimal places. Geographic coordinate system should be WGS1984. Use 'NavType' and 'LocationAccuracy' to report measurement method and accuracy, respectively. Use 'LocationComments' to make addtional notes about location determination.	[value], -999
EndLongitude	Ending longitude in decimal degrees for observations or samples collected over a distance (e.g., trawls, transects). Precision should be the same as that originally reported; ideally to five decimal places. Geographic coordinate system should be WGS1984. Use 'NavType' and 'LocationAccuracy' to report measurement method and accuracy, respectively. Use 'LocationComments' to make addtional notes about location determination.	[value], -999
EventID	ID of the survey event (e.g., dive number, transect, or trawl haul number) on which the sample was made.	[value]; NA
Family	The Family in which the taxon is classified.	[value]; NA
FishCouncilRegion	U.S. fishery management council (FMC) region where the occurrence is located. In cases of boundary overlap, two councils can be listed separated by the pipe character with a leading a trailing space, '	'. If the occurrence is in state waters, the adjacent FMC region will be listed. For a map, see: http://www.nmfs.noaa.gov/sfa/management/councils/ .	North Pacific; Pacific; Western Pacific; Gulf of Mexico; South Atlantic; Caribbean; Mid-Atlantic; New England; NA
footprintSRS	A Well-Known Text (WKT) representation of the	[value]; NA

Field Name	Field Definition	Valid Values
	Spatial Reference System (SRS) for the footprintWKT of the Location. Do not use this term to describe the SRS of the Latitude and Longitude, even if it is the same as for the footprintWKT. 'footprintSRS' must always be populated if 'footprintWKT' is given. Example: The WKT for the standard WGS84 SRS (EPSG:4326) is "GEOGCS["GCS_WGS_1984",DATUM["D_WGS_1984",SPHEROID["WGS_1984",6378137,298.257223563]],PRIMEM["Greenwich",0],UNIT["Degree",0.0174532925199433]]" without the enclosing quotes. For discussion see http://terms.tdwg.org/wiki/dwc:footprintSRS	
footprintWKT	A Well-Known Text (WKT) representation of the shape (footprint, geometry) that defines the sample area. As an example, for downward looking or oblique still images, this would be a polygon representing the footprint of the viewable area (an orthogonal projection using the sea floor as the projection plane). For a video transect, this would be a polygon that represents the footprint of the entire viewable area throughout the length of the video. To report the actual area of the sample, see 'SampleAreaInSquareMeters'. Also please report 'footprintSRS' if you report 'footprintWKT'. Example: the one-degree bounding box with opposite corners at (longitude=10, latitude=20) and (longitude=11, latitude=21) would be expressed in well-known text as POLYGON ((10 20, 11 20, 11 21, 10 21, 10 20)). For discussion see http://terms.tdwg.org/wiki/dwc:footprintWKT	[value]; NA
Genus	The Genus in which the taxon is classified.	[value]; NA
gisCRMDepth	GIS derived depth (bathymetry is Coastal Relief Model)	[value], -999
gisEtopoDepth	GIS derived depth (bathymetry is ETOPO1) .	[value], -999
gisGEBCODepth	GIS derived gisGEBCODepth. The GEBCO_2019 grid is used. It is a global terrain model for ocean and land at 15 arc-second intervals. https://www.gebco.net/data_and_products/gridded_bathymetry_data/	[value], -999

Field Name	Field Definition	Valid Values
gisMEOW	GIS derived Marine Ecoregions of the World, MEOW (Spalding et al., 2007) see: http://www.marineregions.org/sources.php#meow.	[value], NA
Habitat	A description of the geological and biotic environment in which the organism was sampled or observed, described in plain language, preferably in a manner consistent with Coastal and Marine Classification Standard (CMECS) or other regional habitat classification standards. Methods should be reported in 'OccurrenceComments'.	[value]; NA
HighlightImageURL	URL that points to the highlight image hosted on the DSC portal or appropriate archive that displays the specimen of record. A highlight image is one that is an especially good representative image of the taxa as determined by the 'DataProvider'. In the case of records submitted to the Program, this record is based on the image file indicated in the HighlightImageFilePath variable. This value is created by the Program. If this image is null, then it is assumed that the imagery associated with this record is not particularly notable. This image URL can be the same as that reported in 'ImageURL', but does not necessarily have to be the same image.	[value], NA
IdentificationComments	Comments and other info about the characteristics (e.g., color) or taxonomy of the of the sampled or observed organism. If the record is of a type specimen (e.g., holotype), this should be identified in this field along with the associated scientific name and citation.	[value]; NA
IdentificationDate	UTC date (YYYY-MM-DD). Date on which the sample was identified. If day or month is unknown, use YYYY-MM or YYYY.	YYYY-MM-DD; NA
IdentificationQualifier	Taxonomic identification method and level of expertise. Desired format would be along the following lines: 'genetic ID'; 'morphological ID from sample by taxonomic expert'; 'ID by expert from image'; 'ID by non-expert from video'; etc. An	[value]; NA

Field Name	Field Definition	Valid Values
	additional qualifier may be included (e.g., 'ID Uncertain').	
IdentificationVerificationStatus	IdentificationVerificationStatus - Score of the quality of the identification. 1 = identified from image only, 2 = identified from image and physical specimens sampled from the same region, 3 = identified from image and that specific physical specimen.	1;2;3;-999
IdentifiedBy	Name of individual or institution that assigned the 'ScientificName'. Desired format is 'LastName, FirstName' or a list of names delimited by the pipe character with a leading a trailing space, '	'.	[value]; NA
ImageURL	URL that points to the image on the DSC portal or appropriate archive that displays the specimen of record. In the case of records submitted to the Program, this record is based on the image file indicated in the ImageFilePath variable. This value is created by the Program.	[value], NA
IndividualCount	Number of individuals (e.g., sponges or non-colonial corals), colonies (e.g., octocorals, black corals) or patches (e.g., framework forming corals where individual colonies cannot be distinguished) as a positive integer. If categorical, use 'CategoricalAbundance'. Write -999 if no count is available for presence-only data. Add 'present' to the 'CategoricalAbundance' field.	[positive integers]; -999
LargeMarineEcosystem	Large Marine Ecosystem using designations set forth at http://www.lme.noaa.gov/. 'NA' = outside of LME boundary. This field is calculated by the Program based on occurrence location.	[list of LME's]; NA
Latitude	Latitude in decimal degrees where the sample or observation was collected. Precision should be the same as that originally reported. Geographic coordinate system should be WGS1984. Use 'NavType' and 'LocationAccuracy' to report measurement method and accuracy, respectively. Use 'LocationComments' to make addtional notes about location determination.	[value]
LifeScienceIdentifier	Life Science Identifiers are a way to name and	[value]; NA

Field Name	Field Definition	Valid Values
	locate pieces of information on the web. Essentially, an LSID is a unique identifier for some data, and the LSID protocol specifies a standard way to locate the data (as well as a standard way of describing that data). They are a little like DOIs used by many publishers. An LSID is represented as a uniform resource name (URN) with the following format: 'urn:lsid:Authority:Namespace:ObjectID[:Version]'. An example using WoRMS where genus = 'Thesea': 'urn:lsid:marinespecies.org:taxname:125315'. When searched on common web search services, this LSID provides a clear link to this resource. http://www.marinespecies.org/aphia.php?p=taxdetails&id=125315. This value is assigned by DSCRTP and uses the WoRMS database as the primary authority.	
Locality	A specific named place or named feature of origin for the specimen or observation (e.g., Dixon Entrance, Diaphus Bank, or Sur Ridge). Multiple locality names can be separated by the pipe symbol,' | ', arranged in a list from largest to smallest area (e.g., Gulf of Mexico; West Florida Shelf, Pulley Ridge).	[value]; NA
LocationAccuracy	The horizontal distance (in meters) from the given Latitude and Longitude that describes the range of error in the position. Leave the value empty if the uncertainty is unknown or cannot be estimated. Zero is not a valid value for this term.	20m; 50m; 100m; 500m; 1000m; >1000m
LocationComments	Comments and other info about how the location (including depth) was determined. Also comment on any uncertainty regarding the geographic or depth coordinates.	[value]; NA
Longitude	Longitude in decimal degrees where the sample or observation was collected. Precision should be the same as that originally reported. Geographic coordinate system should be WGS1984. Use 'NavType' and 'LocationAccuracy' to report measurement method and accuracy, respectively. Use 'LocationComments' to make addtional notes about location determination.	[value]

Field Name	Field Definition	Valid Values
MaximumDepthInMeters	Positive integer used to express the maximum depth of the observation. If 'DataProvider' gives a single depth, the 'MinimumDepthInMeters' and 'MaximumDepthInMeters' will be equal. If no depth information was provided, both the min and max terms will be -999.	[value], -999
MaximumSize	Maximum size of the taxon observed. Methods for determining size (e.g., height or maximum diameter) should be noted in 'OccurrenceComments'.	[value]; -999
MinimumDepthInMeters	Positive integer used to express the minimum depth of the observation. If 'DataProvider' gives a single depth, the 'MinimumDepthInMeters' and 'MaximumDepthInMeters' will be equal. If no depth information was provided, both the min and max terms will be -999.	[value], -999
MinimumSize	Minimum size of the taxon observed. Methods for determining size (e.g., height or maximum diameter) should be noted in 'OccurrenceComments'.	[value]; -999
Modified	The most recent date the DataProvider updated or verified the record. Format: YYYY-MM-DD. If the DataProvider does not record or provide this information, the date the record became available to NOAA will be used.	YYYY-MM-DD
Morphospecies	A taxonomic species characteristic determined based wholly on observed morphological differences from related species. This field allows the 'DataProvider' to record the the extra detail needed to distinguish among different looking organisms of the same 'ScientificName'. For example, it is common within a given survey that a coral or sponge is clearly identifiable to the 'TaxonRank' of 'genus'. By visual inspection the analyst may also notice morphological differences that are evident, but not diagnostic enough to assigned to 'TaxonRank' of 'species. These morphological notations, e.g., 'msp1, msp2, msp3, or in the case of sponges encrusting, vase, fig, sponge, massive globose etc., may be placed in 'Morphospecies'. This field does not currently have a recommended set of valid values and can	[value]; NA

Field Name	Field Definition	Valid Values
	be used differently by separate research groups. See related variables, 'Morphospecies' and 'CombinedNameID'.	
NavType	Navigation type used to determine coordinates (e.g., 'USBL' or 'ship GPS'). Further detail on NavType can be captured in 'LocationComments'.	ship GPS; USBL; LORAN; other
ObservationDate	Date formatted as YYYY-MM-DD. If month or day is unknown use YYYY-MM or YYYY. Please convert from local date to Universal Time Code (UTC). In some rare cases, the 'ObservationDate' is unknown, as for some museum records. Every effort should be made to recover 'ObservationDate' if missing.	YYYY-MM-DD; YYYY-MM; YYYY; NA
ObservationTime	Time as hh:mm:ss when the sample/observation occurred. Use UTC.	hh:mm:ss; NA
ObservationYear	Year the observation was made in YYYY format (or beginning year of a range, if exact year is unknown).	YYYY; NA
OccurrenceComments	Biological or other detail about the observation event that is not addressed elsewhere (e.g., fishing debris). May be remarks from a corresponding column in original data, or 'IndividualCount' values reported as text 'few scattered', or 'pieces' when integers are expected.	[value], NA
Ocean	Ocean basin where the observation or sample was collected.	Arctic; North Atlantic; South Atlantic; North Pacific; South Pacific; Indian; Southern
OperationalTaxonomicUnit	An operational taxonomic unit (OTU) is an operational identifier used to classify groups of closely related individuals. OTU is a pragmatic identifier to group individuals by similarity, equivalent to, but not necessarily in line with classical Linnaean taxonomy or modern evolutionary taxonomy. This field does not currently have a recommended set of valid values and can be used differently by separate research groups. See related variables, 'Morphospecies' and 'CombinedNameID'.	[value]; -999

Field Name	Field Definition	Valid Values
Order	The Order in which the taxon is classified.	[value]; NA
OtherData	Other data types collected on the same event where this sample was taken. List of data types separated by a comma (e.g., 'CTD, push cores').	[value]; NA
Oxygen	Dissolved oxygen in mL/L for EventID. Use best value or averaged value. Methods should be reported in 'OccurrenceComments'.	[value]; -999
pCO2	Partial pressure of CO2 in seawater, units = microatmospheres. Methods should be reported in 'OccurrenceComments'.	[value]; -999
pH	pH value in units of total pH. Methods should be reported in 'OccurrenceComments'.	[value]; -999
pHscale	Scale used to measure pH. Methods should be reported in 'OccurrenceComments'.	total; NBS; seawater; free; NA
Phylum	The Phylum in which the taxon is classified.	[value]; NA
PI	Principal Investigator or Chief Scientist of the cruise during which the sample was collected, or individual who collected the sample. Desired format is 'LastName, FirstName'.	[value]; NA
PIAffiliation	Affiliation of the Principal Investigator. Well established acronyms or abbreviations are acceptable (e.g., MBARI or WHOI).	[value]; NA
Purpose	Purpose of the cruise during which the sample or observation was collected (e.g., 'deep-sea coral survey', 'exploration', or 'groundfish survey').	[value]; NA
RecordType	Denotes the origin and type of record. The options are: published literature ('literature'); a collected specimen, including those from literature ('specimen'); observation from a still image ('still image'); observation from video ('video observation'); cumulative observation from a video transect ('video transect'); notation without a specimen or image ('notation'); or observation	literature; specimen; still image; video observation; still image transect; video transect; catch record

Field Name	Field Definition	Valid Values
	from trawl surveys, longline surveys, and/or observer records ('catch record').	
Repository	Location where the physical sample, or image is stored. This may be a museum, if samples are physical (e.g., 'Smithsonian NMNH') or it may be an institution (e.g., 'NOAA Central Library') that maintains master copies of images or video.	[value]; NA
Salinity	Salinity in PSU for 'EventID'. Use best value or averaged value. Methods should be reported in 'OccurrenceComments'.	[value]; -999
SampleAreaInSquareMeters	The area in square meters, if projected vertically to a plane representing the seafloor surface, of the sample space in which the observation was made. Use 'footprintWKT' field to specify the exact shape of the sampling space.	[value]; -999
SampleID	The primary identification label or code of the specimen or catch record exactly as reported by the 'DataProvider'. This field, along with 'TrackingID', is used to establish record provenance at the 'DataProvider' institution. Image-based records should also have a SampleID that is a filename or other identifier chosen by the 'DataProvider'. SampleID should be the primary link back to the original database owned by the 'DataProvider'. An individual SampleID can be repeated. For example, there may be multiple taxa in an single image that would require separate records, so that the same SampleID (ID of the image) may be listed multiple times.	[value]
SamplingEquipment	Method of data collection.	ROV; AUV; submersible; drop camera; towed camera; trawl; net; dredge; longline; pot; hook and line; grab; corer; SCUBA; other; epibenthic sled
ScientificName	Taxonomic identification of the sample as a Latin binomial (e.g., Primnoa pacifica), or lowest practical level (e.g., Primnoidae).	[value]
ScientificNameAuthorship	Author who originally described the species or	[value]; NA

Field Name	Field Definition	Valid Values
	subspecies. The author should be consistent with AphiaID and TaxonRank.	
ShallowFlag	A value of 1 denotes a coral or sponge record where DepthInMeters is correct but it is less than 50 meters. A value of 0 denotes a coral that has a correct DepthInMeters greater than or equal to 50 meters.	1; 0; -999
Species	The specific epithet in which the taxon is classified.	[value]; NA
StartLatitude	Starting latitude in decimal degrees for observations or samples collected over a distance (e.g., trawls, transects). Precision should be the same as that originally reported; ideally to five decimal places. Geographic coordinate system should be WGS1984. Use 'NavType' and 'LocationAccuracy' to report measurement method and accuracy, respectively. Use 'LocationComments' to make addtional notes about location determination.	[value], -999
StartLongitude	Starting longitude in decimal degrees for observations or samples collected over a distance (e.g., trawls, transects). Precision should be the same as that originally reported; ideally to five decimal places. Geographic coordinate system should be WGS1984. Use 'NavType' and 'LocationAccuracy' to report measurement method and accuracy, respectively. Use 'LocationComments' to make addtional notes about location determination.	[value], -999
Station	Reference ID of the station where the sample or observation was collected. For example, '78G9-III'.	[value]; NA
Subclass	The Subclass in which the taxon is classified.	[value]; NA
Subfamily	The Subfamily in which the taxon is classified.	[value]; NA
Subgenus	The Subgenus in which the taxon is classified.	[value]; NA
Suborder	The Suborder in which the taxon is classified.	[value]; NA

Field Name	Field Definition	Valid Values
Subspecies	The Subspecies in which the taxon is classified.	[value]; NA
Substrate	Attachment or contact substrate (e.g., dropstone, outcrop, or boulder) of the organism observed. The purpose is to distinguish habitat (like a sandy plain) from attachment point (for instance, a dropstone in a sandy plain). Methods or classification system should be reported in 'OccurrenceComments'.	[value]; NA
SurveyComments	Comments and other information about the cruise/survey or programs sponsoring the cruise/survey.	[value]; NA
SurveyID	Name or ID of the cruise or project in which the sample or observation was collected (e.g, 'EX1202L2' or 'Lophelia II: Reefs, Rigs, and Wrecks'). Many events ('EventID') can be associated with one survey ('SurveyID'). Every effort should be made to use a 'SurveyID' which is standard for the vessel or program. The 'SurveyID' will often be referred to as 'cruise ID' in common communication.	[value]
Synonyms	Other scientific names used for the specimen or observation, including outdated or synonymized names for the taxon, alternate spellings, or previous identifications.	[value]; NA
TA	Total alkalinity, units should be micromoles per kilogram. Methods should be reported in 'OccurrenceComments'.	[value]; -999
TaxonRank	The TaxonRank term is a companion to the ScientificName term. TaxonRank identifies the level in the taxonomic hierarchy of the ScientificName term.	phylum; class; subclass; order; suborder; family; subfamily; genus; subgenus; species; subspecies; forma; variety
Temperature	Temperature in degrees Celsius for 'EventID'. Use best value or averaged value. Methods should be reported in 'OccurrenceComments'.	[value]; -999
TrackingID	Additional ID of the sample or observation provided by the 'DataProvider' (e.g., an additional ID of the occurrence used during field activities or	[value]; NA

Field Name	Field Definition	Valid Values
	other sample tracking). May be used to indicate transects within an 'EventID'.	
TypeStatus	The type status (e.g., holotype, paratype, etc.), if any, of the specimen. Additional information (e.g., typified scientific name, publication) should be included in 'IdentificationComments'.	Holotype; Paratype; Syntype; Lectotype; Paralectotype; Neotype; Allotype; Hypotype; NA
VehicleName	Name or ID of the AUV, ROV, submersible, etc. used to collect sample (e.g., 'Jason II'). Provide the full name of the vehicle, in plain text with no prefix. Use the field 'SamplingEquipment' to indicate vehicle type. A list of valid vehicle names is maintained by DSCRTP to verify and standardize data entries in this field.	[value]; NA
VerbatimScientificName	The 'ScientificName' exactly as reported by the 'DataProvider'.	[value], NA
VerbatimSize	This field is used to store the originally reported Size value if they were not in our required format. If the data provider is available to provide values in our required format, then that is preferable.	[value]; NA
VernacularName	Common name(s) of the observed species.	[value]; NA
VernacularNameCategory	Common (vernacular) name category of the organism.	stony coral (cup coral); stony coral (branching); stony coral (unspecified); black coral; gold coral; gorgonian coral; soft coral; alcyonacean (unspecified); sea pen; lace coral; stoloniferan coral; lithotelestid coral; longhorn hydrozoan coral; other coral-like hydrozoan; sponge (unspecified); glass sponge; demosponge; calcareous sponge; homoscleromorph sponge
Vessel	Name of the vessel used for collecting the sample or observation. Provide the full name of the vessel, in plain text with no prefix, e.g; 'Ronald H. Brown'. A list of valid vessel names is maintained by DSCRTP to standardize the vessel naming convention.	[value]; NA

Field Name	Field Definition	Valid Values
WebSite	World Wide Web link (URL) to an online resource for the 'SurveyEvent' (e.g., a web page devoted to this dataset or expedition). Multiple URLs can be listed and they should be separated the pipe character with a leading a trailing space, '	'.	valid URL; NA
WeightInKg	The recorded weight in kilograms of the individuals in the observation. Use 'OccurrenceComments' for weight methods (e.g., dry weight), calculation or estimation procedures, or other details applied in determining weight.	[value], -999

Appendix 2: Required and Desired Fields for Submission of Data to the National Database

Each of the data sets in the database has unique 'DatasetID'. There are currently 743,758 records in the database composed of 238 individual data sets from 73 different data providers. You will find an an interactive dashboard for each data set at the following location: LINK (not live yet, see reviewer note below).

The following section is revised from Hourigan et al. (2015). Even though the database has a diverse array of fields that can accommodate many types of information, not all are required when submitting data to the Program. For example, the minimum number of required fields for submission to the database is eight. This number is for historical point-based records. All incoming data are evaluated according to the data type. A two-factor classification is used to characterize incoming data according to type. Three categories, 'Point', 'Transect', and 'Trawl', are used to differentiate the first factor and are defined below.

- **Point:** Discrete observations from a single geographic point. An example is an observation of a single taxon from an individual still image or alternatively an actual single point collection of a physical sample.

- **Transect**: Cumulative observations from a pre-determined linear transect. An example would be observations from multiple still images (within a transect) that are grouped together to form a single record in the National Database. Even though these records are currently represented as a single point on the DSCRTP map portal interface, they are an inherently linear sampling frame, and additional required fields help capture the start and end points of the transect line.

- **Trawl:** Specimens from trawl hauls. These data represent observations of multiple collected individuals over a single trawl-haul describing a single taxa, and represented by a single record in the National Database. These records also have an inherently linear sampling frame, and additional required fields help to capture the start and end points of the trawl line if available. Three categories, 'Historical', 'New', and 'Program' are used to differentiate the second factor and are defined below.

- **Historical:** 'Historical' data are already collected and the Program has no ability to influence the format or observational methods. An example is a data set already collected under the funding of another program and that might already be in the literature or another database format, such as records from Monterey Bay Aquarium Research Institute and the Smithsonian Institution at the National Museum of Natural History.

- **New:** 'New' data have not yet been collected, but are intended to be submitted to the Program. An example is data from an upcoming expedition that is not funded by the Program and therefore has fewer submission requirements. The Program would like to encourage these efforts to submit their observations to the National Database, and can provide direct assistance when researchers contact the Program prior to data collection.

- **Program:** Data collected under exploration funded by the Program have the most required fields for submission. In the case of Program funded data collection, it is expected that the principal investigators coordinate directly with the Program to process their data according to Program standards outlined here. Examples include data from upcoming planned cruises with Program funding and 'data rescue' projects designed to create new annotations from existing unanalyzed video or still photos.

Used in combination, these factors (each with three categories), lead to nine different types of incoming data with different numbers of required fields. The Data Working Group uses notation to refer to these combinations: Point Historical, Point New, Point Program, Transect Historical, Transect New, Transect Program, Trawl Historical, Trawl New, Trawl Program. *Table A-C* below summarize the required fields for the nine possible incoming data types.

Table A: Required and desired fields for different types of incoming point-based data. 'R' means the field is required for data submission and 'D' means that it is desired by the Program. 'P' means that the field is supplied by the Program.

DSCRTP Category	Field Name	PointHist	PointNew	PointProgram
Survey Data	PI	D	D	R
Survey Data	PIAffiliation	D	D	R
Survey Data	Purpose	D	D	R
Survey Data	SurveyID	D	D	R
Survey Data	Vessel	D	D	R
Event Data	DepthMethod	D	R	R
Event Data	EventID	D	R	R
Event Data	Locality	D	R	R
Event Data	LocationAccuracy	P	P	R
Event Data	NavType	D	D	R
Event Data	SamplingEquipment	D	D	R
Event Data	VehicleName	D	D	R
Observation Data	DepthInMeters	D	R	R
Observation Data	Latitude	R	R	R
Observation Data	Longitude	R	R	R
Observation Data	ObservationDate	D	R	R
Observation Data	RecordType	P	R	R
Observation Data	SampleID	R	R	R
TaxonomicID	IdentificationDate	D	D	R
TaxonomicID	IdentificationQualifier	D	P	R
TaxonomicID	IdentifiedBy	D	R	R
TaxonomicID	ScientificName	R	R	R
Occurrence Detail	IndividualCount	D	R	R
Metadata	Citation	R	R	R
Metadata	DataContact	D	R	R
Metadata	DataProvider	R	R	R
Metadata	Modified	R	R	R
Metadata	Repository	D	D	R

Table B: Required and desired fields for different types of incoming transect-based data. 'R' means the field is required for data submission and 'D' means that it is desired by the Program. 'P' means that the field is supplied by the Program.

DSCRTP Category	Field Name	TransHist	TransNew	TransProgram
Survey Data	PI	D	D	R
Survey Data	PIAffiliation	D	D	R
Survey Data	Purpose	D	D	R
Survey Data	SurveyID	D	D	R
Survey Data	Vessel	D	D	R
Event Data	DepthMethod	D	R	R
Event Data	EndLatitude	D	D	R
Event Data	EndLongitude	D	D	R
Event Data	EventID	D	R	R
Event Data	Locality	D	R	R
Event Data	LocationAccuracy	P	P	R
Event Data	MaximumDepthInMeters	D	D	R
Event Data	MinimumDepthInMeters	D	D	R
Event Data	NavType	D	D	R
Event Data	SamplingEquipment	D	D	R
Event Data	StartLatitude	D	D	R
Event Data	StartLongitude	D	D	R
Event Data	VehicleName	D	D	R
Observation Data	DepthInMeters	D	R	R
Observation Data	ObservationDate	D	R	R
Observation Data	RecordType	P	R	R
Observation Data	SampleID	R	R	R
TaxonomicID	IdentificationDate	D	D	R
TaxonomicID	IdentificationQualifier	D	R	R
TaxonomicID	IdentifiedBy	D	R	R
TaxonomicID	ScientificName	R	R	R
Metadata	Citation	R	R	R
Metadata	DataContact	R	R	R
Metadata	DataProvider	R	R	R
Metadata	Modified	R	R	R
Metadata	Repository	D	D	R

Table C: Required and desired fields for different types of incoming trawl-based data. 'R' means the field is required for data submission and 'D' means that it is desired by the Program. 'P' means that the field is supplied by the Program.

DSCRTP Category	Field Name	TrawlHist	TrawlNew	TrawlProgram
Survey Data	PI	D	D	R
Survey Data	PIAffiliation	D	D	R
Survey Data	Purpose	D	D	R
Survey Data	SurveyID	D	D	R
Survey Data	Vessel	D	D	R
Event Data	DepthMethod	D	R	R
Event Data	EndLatitude	D	D	R
Event Data	EndLongitude	D	D	R
Event Data	EventID	D	R	R
Event Data	Locality	D	R	R
Event Data	LocationAccuracy	P	P	R
Event Data	MaximumDepthInMeters	D	D	R
Event Data	MinimumDepthInMeters	D	D	R
Event Data	NavType	D	D	R
Event Data	SamplingEquipment	D	D	R
Event Data	StartLatitude	D	D	R
Event Data	StartLongitude	D	D	R
Event Data	VehicleName	D	D	R
Observation Data	DepthInMeters	R	R	R
Observation Data	ObservationDate	D	R	R
Observation Data	RecordType	P	R	R
Observation Data	SampleID	R	R	R
TaxonomicID	IdentificationDate	D	D	R
TaxonomicID	IdentificationQualifier	D	P	R
TaxonomicID	IdentifiedBy	D	R	R
TaxonomicID	ScientificName	R	R	R
Occurrence Detail	WeightInKg	D	D	R
Metadata	Citation	R	R	R
Metadata	DataContact	R	R	R
Metadata	DataProvider	R	R	R
Metadata	Modified	R	R	R
Metadata	Repository	D	D	R

Appendix 3: Data set Dashboards

Each of the data sets in the database has a unique 'DatasetID'. There are currently 743758 records in the database composed of 238 individual data sets from 73 different data providers. You will find an interactive dashboard for each data set at the following location: *https://deepseacoraldata.noaa.gov/DatasetID_Table/DatasetID_Table.html*. These dashboards provide important information about each constituent data set that makes up the National Database and to download each set individually with ISO compliant metadata.