Scripps Caselle Fish Survey Data

2022-08-17

Please **revise and fill in** as much information as possible. Fields highlighted in yellow must be completed.

The information in this document will be used to create the EML of the data package to be published on the [Environmental Data Initiative (EDI) repository](https://portal.edirepository.org/nis/home.jsp).

## 1. Data Package Title

Scripps Oceanographic Institute and Caselle Lab Fish Survey Data from 2008 - Present

## 2. Dataset Entities

*Include what, where, and when in the entity names*

| entity\_position | entityname | entitydescription | filename |
| --- | --- | --- | --- |
| 1 | Scripps Caselle Fish Survey Data 2008 - Present | Data from surveys done by the Caselle Lab and Scripps OI from 2008 to present | scripps\_caselle\_fish\_data.csv |
| 2 | Scripps Caselle Site List | List of the sites surveyed with geographical information for all surveys in the dataset | scripps\_caselle\_site\_list.csv |
| 3 | Scripps Caselle Taxonomy Table | Taxonomy Table for all of the fish species surveyed by both the Caselle Lab and the Scripps OI Team | scripps\_caselle\_taxonomy\_table.csv |

## 3. Abstract

*Please provide an abstract specifically of the dataset. This is different from the abstract of an associated publication. Include what, why, where, when, and how*

[See example here](https://docs.google.com/document/d/1KdHJObHl5Bxxr9t0LISTwX0cCSVvx7NPUjz5T_P9JEc/edit#)

## 4. Creators

*These are the people who will show up as authors in the dataset citation. These are the individuals who have provided intellectual or other significant contributions to the creation of this dataset.* ***Please add a row with the information of each individual that should be part of the authorship of this dataset.***

**We highly encourage you to create an ORCID if you don’t already have one.** [Find more information here.](https://info.orcid.org/what-is-orcid/)

| **First Name** | **Last Name** | **Organization** | **Email** | **ORCID ID** |
| --- | --- | --- | --- | --- |
| Peter | Carlson | University of California Santa Barbara | peter.carlson@ucsb.edu | 0000-0003-2893-2507 |
|  |  |  |  |  |
|  |  |  |  |  |

## 5. Other personnel names and roles

*Who should a data user contact with questions about this data?* *There must be at least one person or organization name to serve as the contact for this dataset.* *You may also list other personnel who participated in the project (such as field crew, lab tech, data entry, etc.) in this table with optional fields, email addresses, organization, and ORCID ID.*

**We highly encourage you to create an ORCID if you don’t already have one.** [Find more information here.](https://info.orcid.org/what-is-orcid/)

| First Name | Last Name | Organization | Email | ORCID ID | Role in Project |
| --- | --- | --- | --- | --- | --- |
| Peter | Carlson | University of California Santa Barbara | peter.carlson@ucsb.edu |  | Contact |
| Camila | Vargas Poulsen | PADL Data Manager, University of California Santa Barbara | camilavargas@ucsb.edu |  | associated party |
| Paloma | Cartwright | Environmental Data Initiative | palomacartwright@ucsb.edu |  | associated party |

## 6. License

***Please highlight*** *the license for the release of your data. Click on the license name for more information.*

[CC BY](https://creativecommons.org/licenses/by/4.0/)

Attribution is required. CC licenses require that those who reuse a work provide attribution to the work’s creator by retaining “identification of the creator(s) of the Licensed Material and any others designated to receive attribution, in any reasonable manner requested by the Licensor (including by pseudonym if designated).”

**Message for user:** Data user is required to cite it appropriately in any publication that results from its use. The Data User should realize that these data may be actively used by others for ongoing research and that coordination may be necessary to prevent duplicate publication. The Data User is urged to contact the authors of these data if any questions about methodology or results occur. Where appropriate, the Data User is encouraged to consider collaboration or co-authorship with the authors. The Data User should realize that misinterpretation of data may occur if used out of context of the original study. While substantial efforts are made to ensure the accuracy of data and associated documentation, complete accuracy of data sets cannot be guaranteed. All data are made available “as is.” The Data User should be aware, however, that data are updated periodically and it is the responsibility of the Data User to check for new versions of the data. The data authors and the repository where these data were obtained shall not be liable for damages resulting from any use or misinterpretation of the data.

## 7. Keywords

*Using keywords from a controlled vocabulary (CV) will improve your data’s future discovery and reuse. The LTER CV is a good source for keywords. Access the* [*LTER CV here*](https://vocab.lternet.edu/vocab/vocab/index.php)*. Also, please determine one or two keywords that best describe your lab, station, and/or project (e.g., Trout Lake Station, NTL LTER).). Add as many rows to this table as needed*

| key\_words |
| --- |
| Palmyra Atoll |
|  |

## 8. Funding of this work

*List only the main PI of a grant that supported this project, starting with the main grant first. Add rows to the table if several grants were involved.*

| PI First Name | PI Last Name | PI Email | PI ORCID ID | Title of Grant | Funding Agency | Funding ID Number |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

## 9. Timeframe

| Begin Date | End Date | Is data collection ongoing or completed? |
| --- | --- | --- |
| 2008 | - | Ongoing |

## 10. Geographic location

*Use decimal degrees to define a point or a bounding box.* Every point you add will be displayed on a map once the data package is published. [See example here](https://portal.edirepository.org/nis/mapbrowse?scope=knb-lter-sbc&identifier=108)

**Verbal description:** Palmyra Atoll National Wildlife Refuge

| Directions | Coordinate |
| --- | --- |
| Northbound | 5.88333 |
| Southbound | 5.87100 |
| Eastbound | -162.04300 |
| Westbound | -162.08300 |

## 11. Methods

*Be specific about the study design and field and lab methods for collecting and processing the data. Include instrument descriptions and protocol citations.*

Find an example [in this link](https://docs.google.com/document/d/1KdHJObHl5Bxxr9t0LISTwX0cCSVvx7NPUjz5T_P9JEc/edit#)

## 12. Data Provenance

*Is this data derived from other data? If so, you will want to document this information, so users know where this data came from. Please specify the source datasets used in the below provenance table, preferably with their DOI or URL.* [*Here is an example of a dataset derived from several others*](https://portal.edirepository.org/nis/mapbrowse?packageid=edi.101.3)*.*

| Dataset Title | Dataset DOI or URL | Creator Name | Contact Email |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

## 13. Data Table

Each row in the below table describes one column in your data table. Complete each row as follows:

* **Description:** Please give a specific definition of the column name. This can be lengthy.
* **Unit:** Identify units for all numeric variables. All rows where there is an \* under the unit column must be filled in with a unit.
* **Date format:** Please tell us exactly how the date and time are formatted: e.g., mm/dd/yyyy hh:mm:ss, plus the time zone and whether daylight savings were observed. ISO date format of YYYY-MM-DD or YYYY-MM-DD hh:mm:ss is preferred.
* **Missing value code:** If a code for ‘no data’ is used, please specify: e.g., -99999, NA
* **Missing value code Explanation:** Why are these values missing? e.g.: value not available, value not recorded.

| attributeName | attributeDefinition | class | unit | dateTimeFormatString | missingValueCode | missingValueCodeExplanation |
| --- | --- | --- | --- | --- | --- | --- |
| project |  | categorical |  |  |  |  |
| reef |  | categorical |  |  |  |  |
| year |  | Date |  | !Add datetime specifier here! |  |  |
| date |  | Date |  | !Add datetime specifier here! |  |  |
| station\_clean |  | character |  |  |  |  |
| diver |  | character |  |  |  |  |
| transect |  | categorical |  |  |  |  |
| depth\_zn |  | numeric | \* |  |  |  |
| species |  | character |  |  |  |  |
| species\_name |  | character |  |  |  |  |
| number |  | numeric | \* |  |  |  |
| size |  | numeric | \* |  |  |  |
| datatyp |  | categorical |  |  |  |  |
| tx\_area\_m2 |  | numeric | \* |  |  |  |
| depth\_actual\_m |  | numeric | \* |  |  |  |

| attributeName | attributeDefinition | class | unit | dateTimeFormatString | missingValueCode | missingValueCodeExplanation |
| --- | --- | --- | --- | --- | --- | --- |
| station\_clean |  | character |  |  |  |  |
| reef |  | categorical |  |  |  |  |
| site |  | character |  |  |  |  |
| station |  | character |  |  |  |  |
| habitat\_zone |  | character |  |  |  |  |
| depth\_zone\_m |  | numeric | \* |  |  |  |
| lat\_dd |  | numeric | \* |  |  |  |
| long\_dd |  | numeric | \* |  |  |  |
| activity |  | categorical |  |  |  |  |
| notes |  | character |  |  |  |  |

## 14. Attributes code

*If you use codes in your column, please define each code in the following table.* ***Fill in only if necessary***

| attributeName | code | definition |
| --- | --- | --- |
| datatyp | QUAN |  |
| project | SIO |  |
| project | TNC\_PARC\_Monitoring |  |
| project | UCSB - CASELLE |  |
| reef | PALMYRA |  |
| transect | A |  |
| transect | B |  |
| transect | C |  |
| transect | D |  |
| transect | E |  |
| transect | F |  |

| attributeName | code | definition |
| --- | --- | --- |
| activity | BENTHIC |  |
| activity | FISH |  |
| reef | PALMYRA |  |

## 15. Scripts/code (software)

*List any software scripts/code you would like to archive along with your data. These may include processing scripts you wrote to create, clean, or analyze the data.*

| Filename | Description | Scripting Language |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |

## 16. Other objects (misc.)

*List any other objects (e.g., .zip, .pdf, etc.) you would like to archive.*

**Note:** At the moment, we are not including any photographic data.

| Filename | Description | Data Type |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |

## 17. Articles

*List articles citing this dataset. Add as many rows as necessary*

|  |  |  |
| --- | --- | --- |
| **Article DOI or URL (DOI is preferred)** | Article title | Journal title |
|  |  |  |
|  |  |  |
|  |  |  |

## Notes, Comments, and Questions

*Please let us know if you have any additional comments or questions about your data or the EML information. THANK YOU!*