

## VI. IEP Data Management Plan

YEAR: PEN: DATE UPDATED:

2019 012 2018-06-05

### STUDY TITLE:

Shrimp Abundance and Distribution Survey

**PRINCIPAL INVESTIGATOR:** Individual(s) responsible for the project. Include name, agency, e-mail, & phone.

Kathy Hieb  
Kathy.Hieb@wildlife.ca.gov, (209) 234-3484

**POINT OF CONTACT:** Individuals who data users should contact for access to the data or questions about the data. Include name, agency, e-mail, & phone number or write "same as above."

Same as above.

**DATA DESCRIPTION:** A very brief description of the information to be gathered; the nature and scale of the data that will be generated or collected.

The San Francisco Bay Study was established in 1980 to monitor the effects of freshwater outflow on the abundance and distribution of fish, brachyuran crabs, and caridean shrimp and in the San Francisco Estuary. Currently, the Bay Study samples 52 fixed stations monthly with trawl nets from south San Francisco Bay through the lower Sacramento and San Joaquin rivers. Data for the shrimp element includes species, sex, counts, lengths, and presence of the branchial cavity parasite *Argeia*. Historically, we also recorded stomach contents (presence-absence) and egg and ovary stage data.

Approximately 3 MB of shrimp data are added per year with the current data format. Older formats include additional data;

**RELATED DATA:** Optional. Existing datasets that you incorporate into analysis and reporting for this program element, existing data that are relevant to your study, or data that are collected simultaneously.

The station, tow, salinity and temperature, fish count and length and crab count sex and size data are collected as part of the "Estuarine and Marine Fishes and Crabs Abundance and Distribution Survey", work plan (#2019-011)

**METADATA:** A description of the metadata to be provided along with the generated data, including the metadata standards used. Provide the file name and information on how users can access the metadata (e.g., a link).

Metadata and methods are found with the catch matrices on our FTP site, available for the public to download.

<ftp://ftp.wildlife.ca.gov/BayStudy/CatchMatrices/>

Additional metadata is in the shrimp Access file Table and Field properties.

Other Metadata (project history, equipment descriptions, sampling protocol, analysis procedures, etc) can be found in "Bay Study SOP doc" and the shrimp sample processing protocol. The SOP is available from any of the Bay Study FS staff

**STORAGE & BACKUP:** A description of the short-term storage methods and backup procedures for the data, including the physical and electronic resources to be used for the short-term storage of the data.

The Bay Study's shared databases, such as the annual shrimp entry file, are stored on the local Stockton server, which is backed up to Sacramento daily. Serial backups of annual and "master" data files are also stored on an external hard drive that is stored in an off-site firebox. In addition, data is uploaded annually a Tier 3 server in Sacramento. Other working files are stored on staff computers, with the most important files backed up to the Stockton server and flash drives that are stored off site.

**ARCHIVING & PRESERVATION:** The procedures for long-term archiving and preservation of the data, including succession plans for the data should the expected archiving entity go out of existence.

Archived copies of Access databases, including serial backups, and relevant working data files are stored on the local server, staff computers, and an external drive that is stored in an off-site firebox. All historical datasheets are stored in the Stockton office, either in file cabinets in the office or file totes in the warehouse.

**ACCESS & SHARING:** A description of how data will be shared. Include (1) access procedures, (2) embargo periods, (3) technical mechanisms for dissemination (e.g., website addresses, listserv information), (3) whether access will be open or granted only to specific user groups, and (4) a timeframe for data sharing and publishing.

The shrimp catch matrix is available for download on the public FTP site, <ftp://ftp.wildlife.ca.gov/BayStudy/CatchMatrices/>. Files are updated and available within 2 to 3 months after a year's samples are processed and QC checked. The master shrimp Access file and the annual abundance indices are available to the public, upon request (contact information above).

**FORMAT:** Formats in which the data will be generated, maintained, and made available. Include BOTH general data type (e.g., spreadsheet, relational database) and file format (extension). Include approximate size (in MB) of the resulting data set.

Data is recorded directly into a MS Access database in the lab as the shrimp samples are processed. With Direct Entry, there is no hard copy data sheet with hand-recorded data. Once a year's samples are processed, QC checked, and the data edited, the length data is converted from one length per record to length frequency data and the sample, catch, and length-frequency data is transferred to a "master" shrimp data file with data from 1980 to current.

A shrimp catch matrix is produced annually in MS Excel. This file has a row for each valid tow and columns for station, tow, and catch data by species. Excel is also used for the final calculation of the annual abundance indices and data.

**QUALITY ASSURANCE:** Brief description of procedures for ensuring data quality. Provide links to Quality Assurance Project Plan and/or QA/QC Standard Operating Procedures.

Staff who process the shrimp samples undergo a rigorous training regime, with specific requirements for correct identification, sexing, and measurement before they "graduate" to processing samples without a check. The entry data base does not allow or produces a warning for certain types of errors, such as length out of range for the species.

For trained staff, 5-10% of their samples are QC checked, which involves identification and counts for the entire sample by another staff person. In some situations, the sex and size of the shrimp may also be checked.

Through a series of Access queries and Excel pivot tables, the data is checked for outliers - distribution, size by species, sex, and season and missing data, such as the plus count, and other errors.

**RIGHTS & REQUIREMENTS:** A link to or instructions to locate the agency's rights and requirements for data use

We ask of data users: "If you use any of these data in a paper, report, or presentation, please acknowledge CDFW's San Francisco Bay Study and the Interagency Ecological Program for the San Francisco Estuary." The Read Me file for the matrix includes the disclaimer "The California Department of Fish and Wildlife makes no warranty of the accuracy, completeness, or fitness of this data for any use. The Department assumes no liability for damages arising from errors, omissions, or the use of this information. ..." Note: There is more text in this disclaimer, but not enough space in this box.