

IEP Data Management Plan

Basic Information

Year: 2020; PEN:011; Date Updated 2019-06-14, Start Date:

Study Title

Estuarine and Marine Fishes and Crabs Abundance and Distribution Survey (San Francisco Bay Study)

Principal Investigator

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

Kathy Hieb, CDFW

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. Include approximate size (in MB) of the resulting data set.

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 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 collected for the fish and crab portion of this study includes station, tow, salinity and temperature, fish species, counts and lengths, and crab species, counts, sex and size. From this data, we calculate annual abundance indices and regional CPUE for 40+ fish and crabs to track annual and seasonal abundance trends and distributional patterns.

- Station, tow, fish catch and length - 90 MB (1980-2017 data)
- Station, tow, crab catch and crab size - 19 MB (1980-2017)
- EC-Salinity-Temperature profile data - 53 MB (1980-2017)

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 Bay Study shrimp samples are collected by the same trawls as the fish and crabs, but processed in the laboratory after collection (PEN 2020-012). The shrimp data is managed in a separate MS Access database.

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 data on our FTP site. The zipped Methods file includes a User's Guide called "Bay Study Access Database Use Guide_Public.doc". In the MS Access file, metadata is also in the Table and Field properties. For the flat files (matrices), there are metadata worksheets and various methods files, such as Species Codes.

Links to [Bay Study Access Database](#) and [Catch Matrices](#) public files can be found on the CDFW ftp site online.

Other metadata (project history, equipment descriptions, sampling protocol, analysis procedures, etc) can be found in Bay Study SOP.

Storage and 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 entry files, are stored on a local Stockton server, which is backed up to Sacramento daily. Serial backups of annual and "master" data files are also copied to an external hard drive that is stored in an off-site firebox. In addition, data is uploaded quarterly to 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. Data sheets for the sample year are kept in an on site firebox until the data is proofed, entered, and QC checked and then moved to file cabinets or file boxes stored at the Stockton DFW office.

The Tier 3 server is maintained by DFW and Resources Agency IT staff. It is our understanding that the Tier 3 server is backup several times a day to a secure location, possibly out of the state. We use this as a back-up, not an archive, as we have only one version of each file stored on this server.

Archiving and 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 MS 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 data sheets are stored in the Stockton office, either in file cabinets in the office or file totes in the warehouse.

Access and 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 [fish and crab catch matrices](#) are available for download on the public FTP site,. The [station, tow, fish catch and length data](#) in a MS Access file that can be downloaded from Other data, including annual abundance indices and the crab catch and size data, are available to the public upon request, contact information above. Files are updated and available to the public within 3 to 4 months after the end of each sampling year.

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).

Data is initially recorded on data sheets and later entered into MS Access databases. MS Excel is also used to generate flat files for more specific analyses and needs, such as catch matrices and abundance indices and catch per unit effort (CPUE) for individual species.

Water quality data is downloaded from a Conductivity-Temperature-Depth (CTD) profiling instrument, converted to text files, and stored in MS Access. Excel spreadsheets are also used to create data summaries.

As of June 2019, most MS Access files are still Access 2000 or 2002 (.MDB), but we have been able to open and work with them in Access 2016 without conversion. Excel files are .XLS or .XLSX, although we are working in Excel 2016.

Quality Assurance

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

All data is collected according to protocol, and a field QA/QC program is in place to ensure accurate sample processing, with 5% of the tows subject to a field QC check for identification and counts.

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. After data is entered into MS Access, it undergoes a line-by-line review to check for accuracy. The station and tow data is double entered and all fields compared. Through a series of Access queries and Excel pivot tables, the station, tow, fish and crab count and size data is checked for outliers and missing data.

Rights and 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 files for the matrices 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. Users of these data are advised to be aware of the locational accuracy, data collection dates, compilation methods, and cartographic format applicable to these data. Users are advised to use these data appropriately. This disclaimer shall apply to any authorized or unauthorized use and transfer of all or parts of these data."