IEP Data Management Plan

Project Element Number:
2026-011

Year:
2026

Date Updated:
2025-05-30

Start Date:
2026-01-01

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

Principal Investigator
Kathy Hieb, CDFW; Kathy.Hieb@wildlife.ca.gov, work cell (209) 640-4642.

Point of Contact

- Kathy Hieb, CDFW; Kathy.Hieb@wildlife.ca.gov
- Alternate contact: Jillian Burns, CDFW, Jillian.Burns@wildlife.ca.gov
- Alternate contact: Kenji Soto, CDFW, Kenji.Soto@wildlife.ca.gov

Data Description
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 data - 102 MB (1980-2024 data)
- Station, tow, crab catch and crab size data- 22 MB (1980-2024)
- EC-Salinity-Temperature profile data - 54 MB (1980-2024)

Related Data
The Bay Study shrimp samples are collected by the same trawls as the fish and crabs but processed in the laboratory after collection (the shrimp study is Program Element # 2026-012). The shrimp data is managed in a separate MS Access database from the fish and crab data.

Metadata

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 public files shown below:

- https://filelib.wildlife.ca.gov/Public/BayStudy/AccessDatabase/
- https://filelib.wildlife.ca.gov/Public/BayStudy/CatchMatrices/

Other metadata (project history, equipment descriptions, sampling protocol, analytical procedures, etc.) can be found in the Bay Study SOP, available from the Point of Contact.

Storage and Backup

The Bay Study's “working” databases, such as the current year entry files, are stored on the 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 safe. In addition, data is uploaded annually to a Tier 3 CDFW server. Other working files are stored on staff computers with the most important files backed up to the Stockton server or OneDrive 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 in the Stockton DFW office or file totes stored in the Stockton DFW warehouse.

The Tier 3 server is maintained by CDFW and Resources Agency IT staff. It is our understanding that the Tier 3 server is backup several times a day to a secure location. 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

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 safe. All historical data sheets are stored in the Stockton office, either in file cabinets in the office or file totes in the warehouse. None of the Bay Study field sheets have been digitized.

Format

Currently, data is recorded on hard copy data sheets in the field and entered into MS Access databases after QAQC processes and edits. MS Excel is used to generate flat files for more specific analyses and summaries, such as catch matrices and abundance indices and catch per unit effort (CPUE) for individual species.

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

Current Access files are in Access 2016 (.ACCDB), while many archived MS Access files are Access 2000 or 2002 (.MDB). Note that we have been able to open and work with these older files using the current version of Microsoft 365 without conversion. Excel files are mostly .XLSX.

Quality Assurance

All data is collected according to protocol, and a field QAQC 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 for certain types of entry errors, such as an incorrect species code or lengths out of range for the species, or produces a warning. After data is entered and second entered into MS Access, the data from both entries are compared using R scripts and corrected. After this first round of corrections, a series of Access queries, Excel pivot tables, and data visualization in R, the station, tow, fish and crab count and size data are checked annually for outliers and missing data.

Access and Sharing

The fish and crab catch matrices are available for download on the public FTP site, https://filelib.wildlife.ca.gov/Public/BayStudy/CatchMatrices/. The station, tow, fish catch and length data in a MS Access file that can be downloaded from the link shown below https://filelib.wildlife.ca.gov/Public/BayStudy/AccessDatabase/.

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 annually and generally available to the public within 6 months after the end of each sampling year. Note: with digital field entry of Bay Study data, we anticipate that data will be publicly available soon after the first of the year.

We started the process to publish the Bay Study station, tow, fish catch, fish length, crab catch, and crab size data to EDI and expect to publish soon.

Rights and Requirements

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."