

IEP Data Management Plan

Project Element Number:

342

Year:

2024

Date Updated:

05-30-2023

Start Date:

2019-01-02

Study Title

Spatio-temporal Community Patterns for Early Life Stages of Fishes and their Associations with Zooplankton in the Upper San Francisco Estuary

Principal Investigator

Gonzalo Castillo, U.S. Fish and Wildlife Service

Point of Contact

Same as above if data access or questions specifically relate to data analyses. Otherwise contact the corresponding IEP Agency listed below maintaining the databases.

Data Description

The data to be used includes species relative abundance (fish, zooplankton, phytoplankton), other covariates (chlorophyll a, water quality, flows). Information will be obtained from various IEP agency websites:

[20 mm survey & Spring Kodiak Trawl survey](https://wildlife.ca.gov/Conservation/Delta/20mm-Survey/FTP%20Site): URL:

<https://wildlife.ca.gov/Conservation/Delta/20mm-Survey/FTP%20Site>

[DayFlow](https://data.ca.gov/dataset/dayflow): URL: <https://data.ca.gov/dataset/dayflow>

[Fall Midwater Trawl survey](https://wildlife.ca.gov/Conservation/Delta/Fall-Midwater-Trawl):

URL: <https://wildlife.ca.gov/Conservation/Delta/Fall-Midwater-Trawl>

[Spring Kodiak Trawl survey](https://wildlife.ca.gov/Conservation/Delta/Spring-Kodiak-Trawl):

URL: <https://wildlife.ca.gov/Conservation/Delta/Spring-Kodiak-Trawl>

Environmental Monitoring Program (EMP) at least since 1995 through 2017 will be requested to EMP staff (phytoplankton, chlorophyll a, zooplankton).

The entire IEP data set to be gathered is estimated to be 80 to 120 MB.

Related Data

No additional data sources, collected or to be collected, are anticipated.

Metadata

Detailed description of the data collected, compiled, and analyzed will be included in final manuscripts intended to be published as open access. No separate metadata plan or specific metadata standards are necessary beyond those already developed by IEP agencies (see data description).

Storage and Backup

Data will be backed up in the short- and long-term on the PIs networked shared drive at USFWS which has automatic back-up routines. The datasets to be generated will be stored in a folder named "20 mm Community Analyses 1995-2017". The queries used to derive the data used for this study and a backup of the resulting flat file will be stored on the CDFW Stockton office shared drive and backed up on an environmental scientist's hard drive.

Archiving and Preservation

The generated datasets will be archived and preserved on the USFWS and CDFW networked drives. The succession plan for the data is co-equally shared by the two agencies and is not expected to change in the event the principal investigator, Gonzalo Castillo, or another collaborator are not present in their current positions, the remaining collaborators will determine where the files will be stored.

Format

Data will be in a csv (.csv), Excel (.xlsx), or Access (.mdb) format. Data analyses will involve a variety of community and statistical analyses programs. Results will be available from the PI and collaborators upon request.

Quality Assurance

We will ensure data has been through QA/QC procedures that are routine for each individual dataset prior to data analyses. For example, in the case of the 20 mm survey, all collected CDFW field data is reviewed by a field crew leader for legibility, completeness, and accuracy. Data entry accuracy is checked by comparing listings of entered data against the data sheet entries twice immediately after their entry into the local database. Data entry is checked again after the end of the field season. The project ES performs reasonableness checks by running computer queries that flag outliers or erroneous entries. After these data checks are complete, a project lead (ES, environmental scientist) creates a flat file for this study to be independently reviewed for accuracy by another ES in the Stockton CDFW Office. Quality assurance procedures for each dataset used in this study are available upon request from each IEP Program. Preliminary analyses from compiled data sets generated in this study will be reviewed by the PI and collaborators based on Cox et al. 2001 ([Best practices for preparing ecological data sets to share and archive](https://www.jstor.org/stable/20168543)) [URL: <https://www.jstor.org/stable/20168543>].

Access and Sharing

Because this study will utilize data collected by existing long-term IEP projects, the availability of data to the public will depend on each IEP agency and specific program. All datasets used in our study are part of the IEP and are publicly available, either online or by request.

Rights and Requirements

Data management procedures are informed by:

- 1) [CDFW Department Bulletin Number 2017-02 "Department of Fish and Wildlife Scientific Integrity Policy"](#)

URL: <https://nrm.dfg.ca.gov/FileHandler.ashx?DocumentID=154775>

- 2) [Data use rights and requirements for federal government datasets:](#)

URL: <https://www.usa.gov/copyrighted-government-works>