

# IEP Data Management Plan

## Project Element Number:

072

## Year:

2024

## Date Updated:

June 1, 2023

## Start Date:

January 1, 2024

## Study Title

Environmental Monitoring Program

## Principal Investigator

Theodore Flynn, Ph. D.

California Department of Water Resources

Email: [theodore.flynn@water.ca.gov](mailto:theodore.flynn@water.ca.gov)

Phone: 916-376-9715

## Point of Contact

Elizabeth Wells (Benthic Invertebrates)

California Department of Water Resources

[elizabeth.wells@water.ca.gov](mailto:elizabeth.wells@water.ca.gov)

916-202-4760

Scott Waller (Continuous Water Quality)

California Department of Water Resources

[scott.waller@water.ca.gov](mailto:scott.waller@water.ca.gov)

916-882-1245

Sarah Perry (Discrete Water Quality)

California Department of Water Resources

[sarah.perry@water.ca.gov](mailto:sarah.perry@water.ca.gov)

279-789-4574

Tiffany Brown (Phytoplankton)

California Department of Water Resources

[tiffany.brown@water.ca.gov](mailto:tiffany.brown@water.ca.gov)

916-202-4805

## Data Description

Since 1975, IEP's Environmental Monitoring Program (EMP) has collected water quality and biological monitoring data in the San Francisco Bay-Delta estuary ("the estuary" hereafter). A description of the data collected by each program element is as follows:

### *Discrete Water Quality*

EMP has collected samples of discrete water quality in the estuary since 1975. The number of stations monitored has varied over time in response to programmatic reviews and new water right decisions, the most recent being D-1641. In 2023, the EMP will monitor the concentration of nutrients, suspended solids, and chlorophyll a at 24 fixed sites across the estuary. Up to four additional floating "entrapment zone" (EZ) stations will be sampled depending upon specific conductance criteria at the bottom of the channel. Criteria for EZ stations are a measured bottom conductivity of either 2,000 or 6,000  $\mu\text{S}/\text{cm}$ . The approximate size of the compiled data collected yearly is 300 MB.

### *Continuous Water Quality*

EMP has monitored water quality in the estuary using continuous sensors since 1983. The number of stations monitored has varied over time in response to programmatic reviews and new water right decisions. The EMP monitors continuous real-time water quality constituents and chlorophyll at 15 fixed sites, 1 meter below the water's surface. Bottom specific conductance and water temperature values are also collected at four of those monitoring stations. Meteorological data is collected at ten stations in the existing continuous real-time monitoring network. Data is recorded every 15 minutes. The approximate size of the compiled data collected yearly is 14 GB.

### *Benthic Invertebrates*

EMP has collected benthic data within the estuary since 1975. The number of stations monitored has varied over time in response to programmatic reviews and new water right decisions. The EMP currently monitors the benthos at 10 fixed sites across the estuary. The approximate size of the compiled data collected yearly is 620 MB.

### *Phytoplankton*

EMP has collected phytoplankton data within the estuary since 1975. The number of stations monitored has varied over time in response to programmatic reviews and new water right decisions. The EMP currently monitors phytoplankton abundance and biovolume at the same sites as discrete water quality (described above). The approximate size of the compiled data collected yearly is 50 MB.

## Related Data

EMP collects data for discrete water quality, phytoplankton, and zooplankton (PEN# 077, managed by the California Department of Fish and Wildlife) simultaneously at the stations described above. Benthic sampling is conducted at separate (although often nearby) stations, with limited, sonde-based water quality measurements performed in

tandem with each sampling event. Continuous water quality is also recorded at separate, often nearby stations to discrete water quality.

## **Metadata**

### *Discrete Water Quality*

Data and metadata are updated annually on the Environmental Data initiative and are formatted in Ecological Metadata Language (DOI: 10/kdng). Metadata includes citation information, geographic scope, project history, data collection methods, laboratory methods, and parameters collected.

### *Continuous Water Quality*

Metadata was last updated in 2022. Some metadata elements are under revision are currently being updated to meet DWR QA Program requirements. EMP staff follow the same metadata format for each data type. The Continuous Real-time Water Quality Metadata is available from the point of contact listed above.

### *Benthic Invertebrates*

Benthic data and metadata (dates, locations, gear used, etc.) are updated annually on the Environmental Data Initiative and are formatted in Ecological Metadata Language (DOI: 10/j7cp). Metadata includes citation information, geographic scope, project history, data collection methods, laboratory methods, and parameters collected.

### *Phytoplankton*

Metadata is updated annually on the Environmental Data Initiative and are formatted in Ecological Metadata Language (doi: 10/j284). EMP staff follow the same metadata format for each data type. Metadata includes citation information, geographic scope, project history, data collection methods, laboratory methods, and parameters collected.

## **Storage and Backup**

### *Discrete Water Quality*

Data is collected from the EXO2 sondes using a custom software platform “Moped” written using Microsoft Web Forms with C# as the backend. Following each discrete sampling event, Moped generates electronic field data sheets in PDF and CSV format that are stored both locally on the computer’s hard drive and as well as EMP’s secured SharePoint site. Backup copies of electronic field data sheets are stored by on the Division of Integrated Science and Engineering (DISE) shared drive, which is backed up daily by the Division of Technology Services (DTS). Discrete water quality data is entered into FLIMS using Microsoft Access 2016 and then imported into DWR’s Oracle-based Water Data Library (WDL).

### *Continuous Water Quality*

Data is stored electronically in the EXO2 sondes and retrieved on station visit sheets. YSI electronic data files are transferred monthly from the sondes in binary format to a DISE server that is backed up daily by DTS. Since 2021, electronic station visit and field sheets are stored permanently in a DTS-managed Oracle database that is backed up yearly.

#### *Benthic Invertebrates*

Digital data, both sediment and biological, is stored in an MS Access database (.mdb format) on a dedicated server managed by DTS and backed up daily. Sediment reports from Bryte Laboratory (.pdf format) and field notes (.doc or .docx) are stored on a DISE server that is managed by DTS and backed up daily. Original paper copies of taxonomic data are stored in the DWR/DISE warehouse.

#### *Phytoplankton*

Beginning in 2018, the Microsoft Access 2003 database (.mdb format) is kept in the cloud on a Microsoft OneDrive server dedicated to EMP data, which is managed by the DTS and backed up daily. Once data is entered into Microsoft Access 2003, it is QA/QC'ed for errors, and backed up to an external hard drive that is stored away from the office in case the main database (or the server housing it) is damaged.

## **Archiving and Preservation**

#### *Discrete Water Quality*

Data from this project is archived both on paper and digitally. Electronic field data sheets are printed out and filed at the EMP offices in West Sacramento. Digital records are permanently stored on DWR's Water Data Library (WDL), an Oracle database managed by DTS. Annual updates are also posted to the cloud-hosted Environmental Data Initiative (EDI) data repository. DWR is a long-established agency with no anticipated end date of existence.

#### *Continuous Water Quality*

The continuous real-time water quality data is stored permanently in a DTS-managed Oracle database that is backed up once a year. Electronic visit sheets are stored on a server that is managed by DTS and is backed-up daily. Historic reports are also scanned and stored on a DISE server, while the hard copies are stored internally in the DWR/DISE warehouse. DWR is a long-established agency with no anticipated end date of existence.

#### *Benthic Invertebrates*

The benthic digital data (taxonomic and sediment) is stored permanently in a Microsoft Access database that is backed up daily by DTS. Original paper datasheets with taxonomic are stored internally in the DWR/DISE warehouse. Original sediment reports and field notes are stored on DTS servers. DWR is a long-established agency with no anticipated end date of existence.

## *Phytoplankton*

All electronic data, such as sample data, invoices, or COCs, is stored in the same OneDrive cloud server as the Microsoft Access 2003 Database. Physical samples are kept for 2 years after analysis before being discarded. Data prior to 2008 is stored in a Microsoft Access 2000 database (.mdb format). Prior to 2003, samples were analyzed by Bryte Laboratory using hard copy datasheets. Scanned versions of these datasheets are stored on the same OneDrive cloud server as the data. The original hard copy datasheets are stored in the DES office. DWR is a long-established agency with no anticipated end date of existence.

## **Format**

### *Discrete Water Quality*

Field data are documented in Microsoft Web Forms and exported into field sheets in .pdf format and as tables in comma-separated (.csv) format. Field data is entered into FLIMS, a Microsoft Access 2016 application, within 24 hours or as soon as possible. Field data is appended to lab analyses, and both are stored in Water Data Library. Data is published on EDI as .csv flat files.

### *Continuous Water Quality*

Quality assured continuous real-time water quality data can be accessed by specific DWR personnel (only) from the DWR Water Quality Portal (WQP) Oracle database. Data is populated into the WQP Oracle database once every 15 minutes. Data from the WQP Oracle database can be exported in MS Excel or text files, which can then be saved as different file formats including .xlsx, and .csv.

Preliminary (subject to revision) continuous real-time water quality data can be accessed from DWR's California Data Exchange Center (CDEC). Data is populated onto the CDEC data collection network once every 15 minutes. Data from CDEC can be exported in MS Excel or text files, which can then be saved as different file formats including html, .txt, .pdf, and others.

### *Benthic Invertebrates*

Data received in hard copies on paper from Hydrozoology (taxonomic contractor, analyzing biological data) and as PDFs from the Bryte Soils and Concrete Laboratory (sediment data) is entered into the benthic database using MS Access 2003 software (.mdb file format). Data produced by queries in MS Access can be saved as different file formats (.xlsx, .csv, .pdf and others). Field notes are written in MS Word documents.

## *Phytoplankton*

Phytoplankton samples are analyzed for abundance and biovolume by an outside contractor, who returns the data electronically as Excel spreadsheets (.xls or .xlsx formats). These spreadsheets are archived in their original format as they are received

from the contractor, and any modifications made for entry into the Microsoft Access 2003 database (.mdb format) are saved as separate files (.xls or .xlsx formats).

## **Quality Assurance**

### *Discrete Water Quality*

All field data is entered into FLIMS, which integrates the field and laboratory data. Bryte Laboratory manages the QA/QC for the laboratory constituents. When lab reports are received, EMP conducts a monthly check to ensure data completeness. EMP staff confirm that the data in WDL properly match what was written on the field sheets. If errors occur, they are corrected in WDL prior to making data public. EMP has developed a Quality Assurance Project Plan (QAPP) in collaboration with DWR's Quality Assurance Program. Standard Operating Procedures are available by request to the point of contact listed above.

### *Continuous Water Quality*

Continuous water quality data is populated into the WQP Oracle database every 15 minutes. EMP conducts a monthly check to ensure data completeness. EMP staff confirm that the data sheets properly match what was recorded on station visit sheets and investigates missing or incompleteness due to transcription errors. A graphical editing program is used to determine outliers of the continuous real-time data stream. All errors are corrected or flagged in WQP Oracle database prior to making data public. The approved Quality Assurance Project Plan (QAPP) was finalized in April 2022. EMP is working with DWR's QA program to continue improving and updating our QA/QC methods and documentation.

### *Benthic Invertebrates*

DWR Soils Laboratory employs sieve-testing (as per ASTM D-422) to assure consistency in grain size testing, and the lab is accredited through the American Association of State Highway and Transportation officials (AASHTO), which does quality audits and inspects the facilities, worksheets, and reporting documents are accurate and results are repeatable. Hydrozoology manages the QA for the taxonomic identification by employing the recognized taxonomic expert in this region, who consults with subject matter experts when necessary. When reports are received from either the Soils Lab or Hydrozoology, EMP conducts a monthly check to ensure data completeness. After an EMP staff member enters data into the benthic database, another staff member double-checks the data entry as QC; data entry and data check are initialed. Any discrepancies are resolved using original data sheets, in consultation with either the Soils Lab or Hydrozoology if necessary. A further QA/QC process is in the progress with the development of a Quality Assurance Project Plan. The EMP is working with DWR's QA Program to improve and build upon QA methods and documentation.

### *Phytoplankton*

Phytoplankton data is entered as it is received, usually monthly. The format of the database prevents typos and spelling errors in taxon names, so the QA/QC process is used to check for numerical or label typos (e.g., incorrect station number or incorrect organism count). A phytoplankton QAPP was completed in 2017. The QAPP is currently not available online, but a copy can be requested by emailing the point of contact listed above.

## **Access and Sharing**

Data from 1975 – 2022 is publicly available on the Environmental Data Initiative (EDI). Reports and figures are publicly available online through the EMP's Github site (<https://emp-des.github.io/emp-reports/>). Data and reports are updated annually but data requests can be sent to points of contact listed above at any time. QA/QC'd data are typically available by request 2 months after collection.

### *Continuous Water Quality*

Preliminary (subject to revision) continuous real-time water quality data is publicly available on CDEC (<http://cdec.water.ca.gov>). Data are populated onto the CDEC data collection network once every 15 minutes and is readily available online. QA/QC'd real-time water quality data can be accessed by specific DWR personnel (only) from DWR's Water Quality Portal (WQP) Oracle database. Data requests can be sent to the point of contact listed above.

### *Benthic Invertebrates*

Data from 1975-2022 is available on EDI (DOI listed above). Newer data can be made available by request to the point of contact listed above. There is always at least an approximately two-month delay between sampling date and any data availability, to allow time for samples to be analyzed by the taxonomist and the data to be entered and undergo QA/QC. All mandated data is publicly available, and special datasets that would be tedious to create using the flat files (e.g., all instances of a single species across the entire period of record) are performed upon request to the point of contact listed above.

### *Phytoplankton*

Data from 2008 – 2021 is available on on EDI (DOI listed above). Older data is available by contacting the point of contact listed above. Data requests for more recent data before the annual update can be made to the point of contact listed above.

## **Rights and Requirements**

Data from this project is publicly and may be used freely with attribution. Cite the version used and, if using data from EDI, the DOI from the version of the data used and the appropriate citation listed on the EDI website.