

# Santa Cruz Wharf IFCB Data Stream Plan

Author: Patrick Daniel, UCSC  
Version 1: June 13, 2023



Santa Cruz Wharf (c. City of Santa Cruz)

## **Introduction**

The Santa Cruz Wharf IFCB is deployed on a platform underneath the Santa Cruz Wharf (SCW), California, USA. The Santa Cruz Wharf is a municipal pier located in the Northern region of Monterey Bay. The platform is roughly 100 meters offshore of the coastline and generally outside the surf zone. The Santa Cruz wharf was chosen for access and precedence. The SCW is owned and maintained by the city of Santa Cruz, which provides convenient access to a secured space to deploy the IFCB under the wharf. Additionally, the SCW has been the site of weekly water quality sampling (nutrients, toxins) and phytoplankton net tows dating back to October 2006.

## **Deployment Description**

**Coordinates:** 36.960 N, 122.020 W

**Ocean Depth:** ~10 meters

**Deployment Length at Full Resolution:** indefinite

**Sampling Depth:** 0.5 - 1 meter from surface

**Auxiliary Measurements:** Co-located multiparameter sonde (YSI Exo Sonde), weekly sentinel mussel, SPATT , weekly wharf whole water sampling (nutrient, toxins, Chl-a), phytoplankton net tows.

## **Power**

110v AC power for the IFCB instrument, peristaltic pump, and telemetry equipment are provided by the Santa Cruz Wharf. A 24v DC power supply run through an uninterruptible power supply (UPS) powers the IFCB. Telemetry equipment and peristaltic pumps are powered from 120V AC.

## **Telemetry**

Data is telemetered through a cellular mobile hotspot (4G Jetpack, Verizon Wireless). In this configuration, it is not possible to assign a static IP address. The instrument is accessed remotely using Teamviewer. Remote power cycling is done using the iBoot G2.

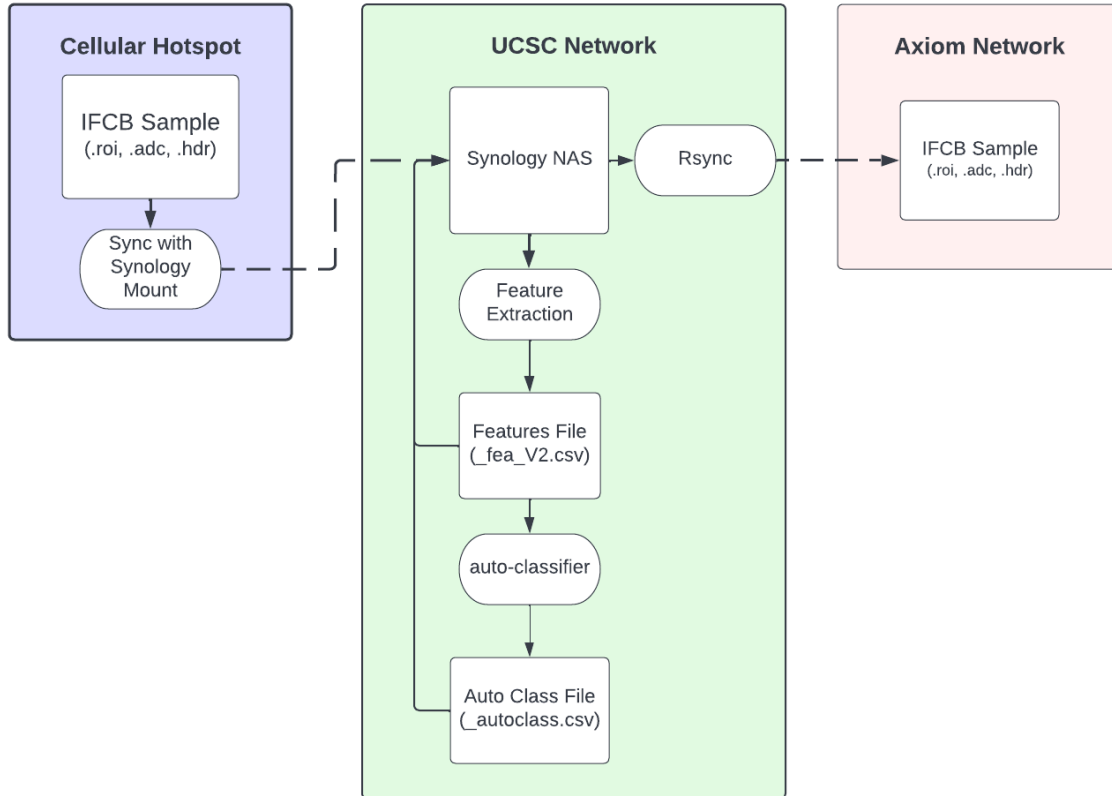


**Figure 2.** Water is drawn up from the surface through a peristaltic pump. IFCB samples are drawn from a tee fitting set inline with the pump intake.

### **Sampling**

Water is drawn up from the ocean surface using a peristaltic pump (figure 2). IFCB samples are drawn from a tee fitting located inline to the pump intake and taken before passing through the pump to avoid damage to organisms. Peristaltic pump exhaust is returned to the ocean through an output line that extends to below the floor. Samples are drawn up through a 150  $\mu$ m mesh filter set in the intake line. Sampling is set to the full resolution, 5ml per sample and ~60 per day.

Bead cycles and bleach and biocide routine is run every 40th sample to prevent fouling and provide qualitative metrics.



**Figure 3.** Wire diagram of the SCW IFCB data flow.

### Data Processing

Sample data are remotely synced with a NAS (Synology) on campus at UCSC via a cellular hotspot modem using the Synology Cloud Sync and files are synced upon creation. Once per day or as is feasible, data quality is reviewed by technical staff and the processing queue is initiated. A Matlab routine is run to generate blob and feature files. Currently, version 2 (V2) blob and feature files are being generated in operational use. Version 4 (V4) blob and feature files currently being generated for testing and validation, with the intention of replacing the V2 files in the future. All blob and feature files are stored on a NAS at UCSC. Currently, images are classified from V2 feature files using a random forest model (Model Name, Fischer et al 2020). For each syringe, a plain text “Autoclass” file is generated containing the probability outputs for each class and image of the random forest classifier. Model output is saved on a NAS located at UCSC. A convolution neural network is currently being evaluated to replace the operational random forest classifier and will be implemented upon completion of testing and validation.



Once per day, raw data (.roi, .hdr, .adc) and derived data products (V2 feature, V2 blob, and autotclass files) are pushed to the Axiom FTP server using an rsync script that is run on a CRON job.

### **Data Storage**

All IFCB data (raw and generated products) are stored locally at the UCSC on the Synology NAS configured with RAID. Data is also sent to Axiom, which acts as a secondary, offsite data redundancy. To date around 462 GB of data have been generated by the SC Wharf IFCB.

### **Data Access**

Instrument generated files (.roi, .adc, .hdr) and derived product files (V2 feature, V2 blob, and autotclass files) are publically accessible through the UCSC Ocean Data Center IFCB dashboard instance (<http://akashiwo.oceandatacenter.ucsc.edu:8000/timeline?dataset=SCW>) and the Axiom maintained Caloos IFCB Dashboard (<https://ifcb.caloos.org/timeline?dataset=santa-cruz-municipal-wharf>). Local data access is available through a Synology NAS, but is not available to the public.

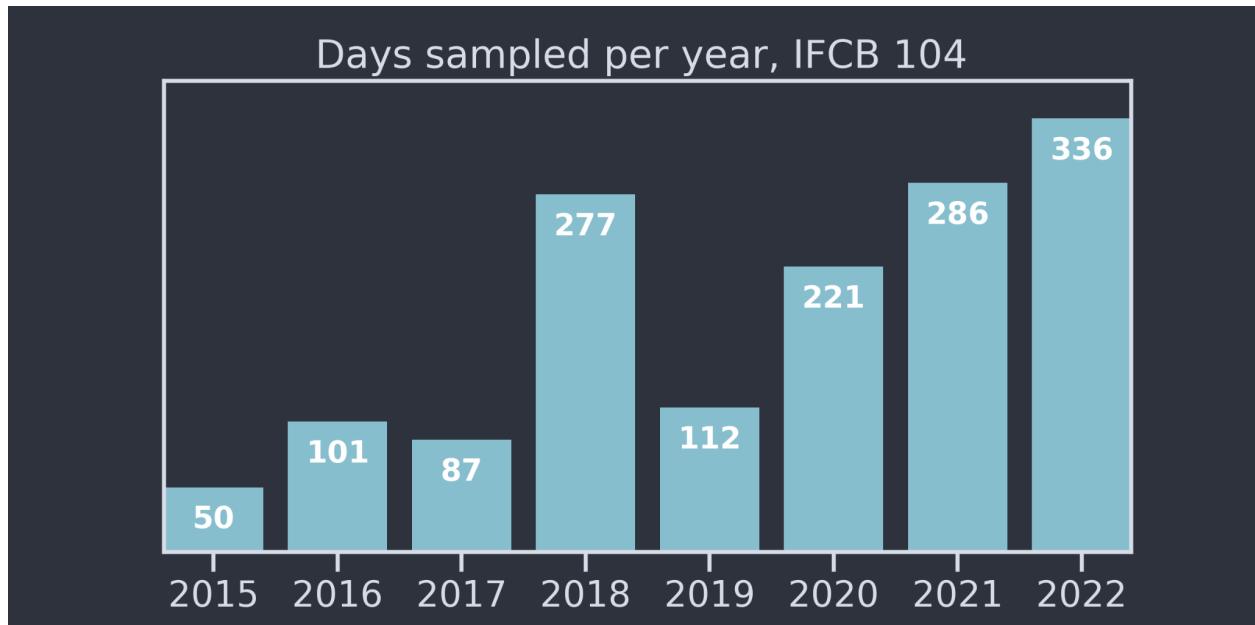
An experimental web-based data visualization tool, “HAB Tracker” (<http://eucampia.ucsc.edu:3000/>) displays near real-time classified cell concentrations from the SC Wharf IFCB. The site was built as a proof of concept visualization tool for exploring phytoplankton classes of interest (i.e. harmful algal bloom-forming species).

### **Maintenance/Issues**

The instrument is stored in a locked room located underneath a busy municipal pier and is subject to vibrations from traffic overhead, nesting pigeons, and corrosive seaspray. Occasionally, when high tides, storm surge, and large waves occur concurrently, sea level exceeds the platform base and the room floods. Electronics are stored in water proofed plastic boxes secured 4 feet off of the ground. The IFCB is secured to the wall

IFCB generated data is checked daily for quality assurance and checked weekly in person during weekly sampling at SCW. Service and repairs are generally made on site. Occasionally, the instrument is collected and returned to the lab for deep cleaning or servicing.

IFCB S/N 104 (aka “Tina”) is the instrument primarily deployed at the SCW. Tina has been deployed at the wharf with intermittency since 2015.



**Figure 4.** Number of days with at least one sample per year at the Santa Cruz Wharf.

**Are there ethical restrictions to data sharing?**

No

**Who holds intellectual property rights (IPR) to the data?**

University of California, Santa Cruz and CeNCOOS

**Describe any effect of IPR on data access.**

None

**Datas Source and Quality Control**

**Indicate the data source type (i.e. Federal, Non-Federal, University, Stage Agency, Local Municipality, Military Establishment (branch), private industry, NGO, non-Profit, Citizen Science, Private individual)**

University

**Indicate the data reporting type (e.g. real-time, historical)**

Real-time

**If real-time, list the QARTOD procedures that are currently applied.**

Data currently falls outside of the scope of QARTOD. The data are a set of photos and are inspected by the IFCB operators on an ad hoc basis. The ML model derived identifications are addressed in the Model Card reporting.

**What is the status of the reported data? (e.g. raw, some QC, incomplete, delayed mode processed but not QC'd)**

Both raw data and ML modeled identifications are presented. QC of the model output is discussed in the model card.

**Describe the data control procedures that were applied by the originator.**

The originator monitors the output images to make sure that the images are clear and that the instrument is operating properly and behaving in the specified manner.

**Describe the data control procedures that were applied by CeNCOOS.**

Axiom, the CeNCOOS data partner, applies the ML algorithm to the data and generates the output subject to all the caveats described in the model card.

**Provide a line to any documented procedures**

<https://www.cencoos.org/data/ifcb-doc-proto/index.html>

**List the procedures taken for data that could not be QC'd as directed.**

N/A

**Stewardship and Preservation Policies**

**Who is responsible for long-term data archiving?**

Long-term archiving is currently the responsibility of the operating institution with Axiom Data Science as a backup. Future plans are to use Ocean Vision AI with NCEI for archive of images.

**Which long-term data storage facility will be used for preservation?**

Near real-time images are planned to be stored through Ocean Vision AI through the NCEI connection.

**Describe any transformations necessary for data preservation.**

Data will be bundled into tar files for transfer.

**List the metadata or other documentation that will be archived with the data.**

N/A