

Bodega Marine Lab (BML) IFCB Data Stream Plan



Figure 1. IFCB-163 is located in the seawater pump house at the UC Davis Bodega Marine Laboratory. The instrument is co-located with a CTD and Fluorometer and draws up water from the

First Draft: August 22, 2022.

Author: Patrick Daniel

Introduction

IFCB (S/N 163) is deployed in the seawater intake pump house at the University of California Davis' Bodega Marine Lab (BML), located on the Bodega Headland in Bodega Bay, California. Bodega Bay is located ~30 miles North of Pt. Reyes and ~95 miles south of Pt. Arena, a prominent upwelling center for the California Current Ecosystem.

Seawater is drawn up by a pump from Horseshoe cove into the pumphouse and used to fill a flow cell where the IFCB samples water from. Co-located are a CTD and a fluorometer part of the Bodega Marine Lab Seawater intake shore station. Intake waters are drawn from 1 to 3 meters below the surface, depending on the tidal cycle.

Intake lines are cleared once a year. The effect of filter-feeder growth within the intake line has not been explored with the IFCB.

Deployment Configuration

The IFCB is configured for full resolution. The Bleach, biocide, and Bead (BBB) routine runs every 25 samples. Data is processed and distributed hourly.

Telemetry: The IFCB is hardwired to the local network via ethernet cable.

Power: 120V AC is used to power the power supply.

Fouling/Corrosion: The instrument is protected in the pumphouse.

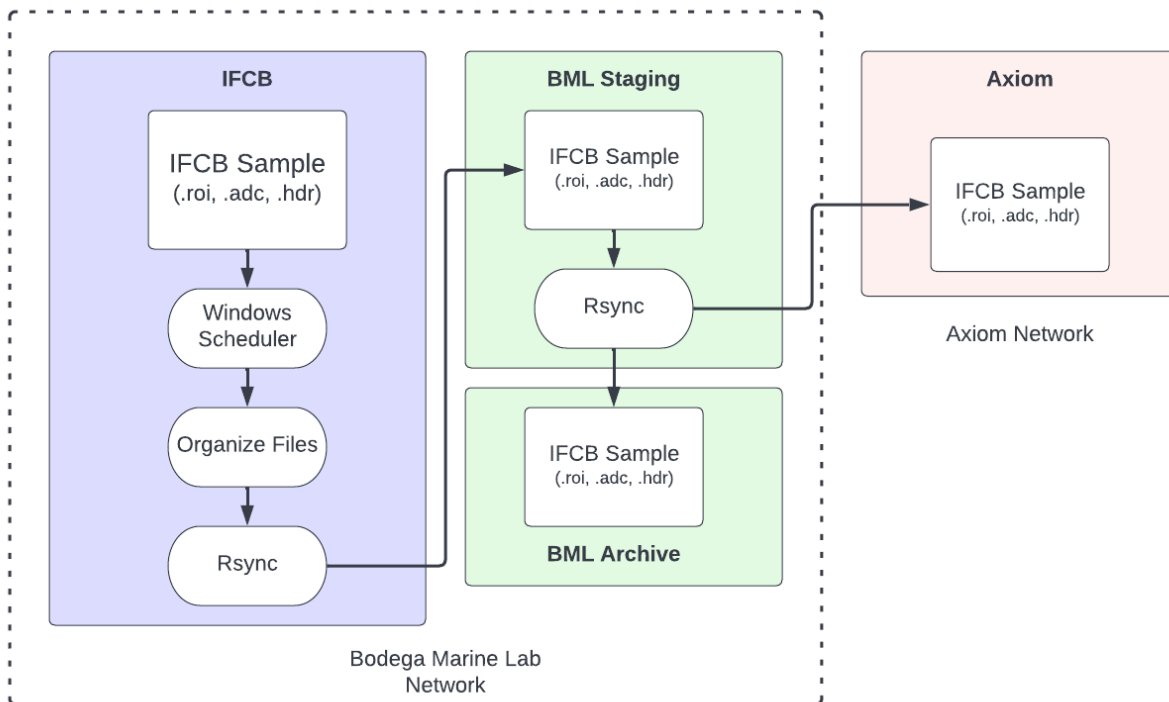
Data Flow

Newly generated files are copied onto a staging directory on the instrument computer and stored using the Axiom-standard convention, organized by **/year/day/sample**.

Example: Sample D20220823T172211_IFCB163

```
/2022/D20220823/D20220823T172211_IFCB163.roi  
/2022/D20220823/D20220823T172211_IFCB163.hdr  
/2022/D20220823/D20220823T172211_IFCB163.adc
```

New data is copied off the instrument using an RSYNC script pushed by a Windows Scheduling routine. Data are moved to a staging server on the BML network and then copied to an archival server at BML (BML Archive) and pushed to the Axiom FTP server. The local staging area is meant for temporary storage and is periodically wiped as storage becomes limited. BML Archive is meant for long-term storage and is configured with RAID for local data redundancy.



Data Storage and Redundancy

Local data redundancy is maintained from the BML Archive server configured for RAID. Axiom maintains offsite redundancy.

Data Access Points

Data is publicly accessible through the IFCB Dashboard deployed by Axiom (<https://ifcb.caloos.org/timeline?dataset=bodega-marine-lab>).

Post-processing

N/A

Are there ethical restrictions to data sharing?

No

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

University of California, Davis, Bodega Marine Laboratory (BML) 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