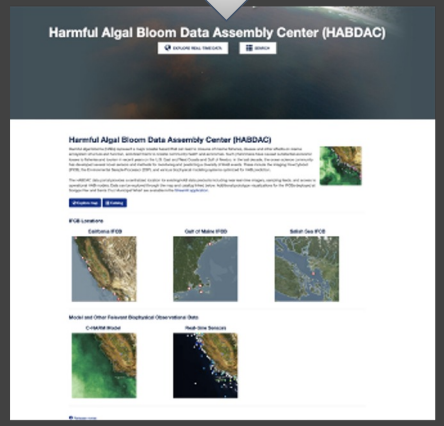


The Harmful Algal Bloom Data Assembly Center: A National Cyberinfrastructure Framework for Plankton Imagery

Clarissa Anderson (SIO)*, Rob Bochenek (Axiom Data Science), Heidi Sosik (WHOI), Stace Beaulieu (WHOI), Raphael Kudela (UCSC), and Henry Ruhl (MBARI)

***SCCOOS Director, CIMEAS Director**

Scripps Institution of Oceanography

6 May 2025

OOI FB-DSC Meeting

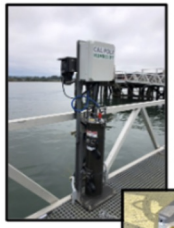
**Thank you to the NOAA NCCOS PCMHAB program*

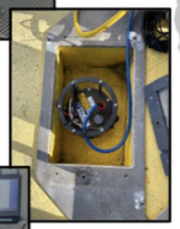
Ocean Observing and Modeling

Data Management and Cyberinfrastructure

Stakeholder Engagement and Co-Design

CA IFCB Network



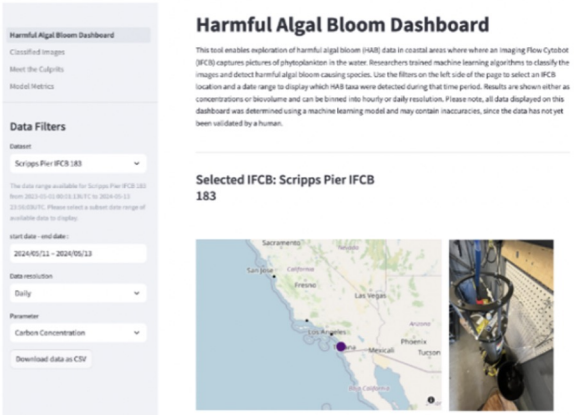


CalOOS IFCB Dashboard

ifcb.caloos.org

Data available on CalOOS IFCB dashboard (ifcb.caloos.org) **and** available for download through the UCSD Library Archive

HABDAC



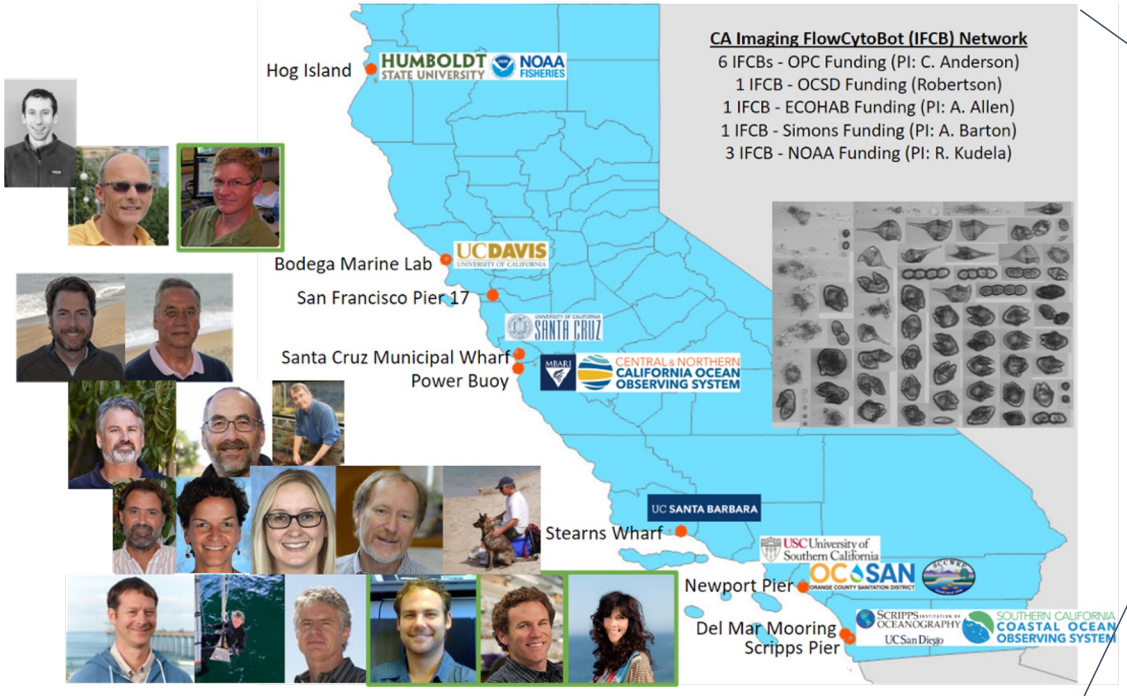
Data illustrated in a comprehensible graphical interface for effective HAB management

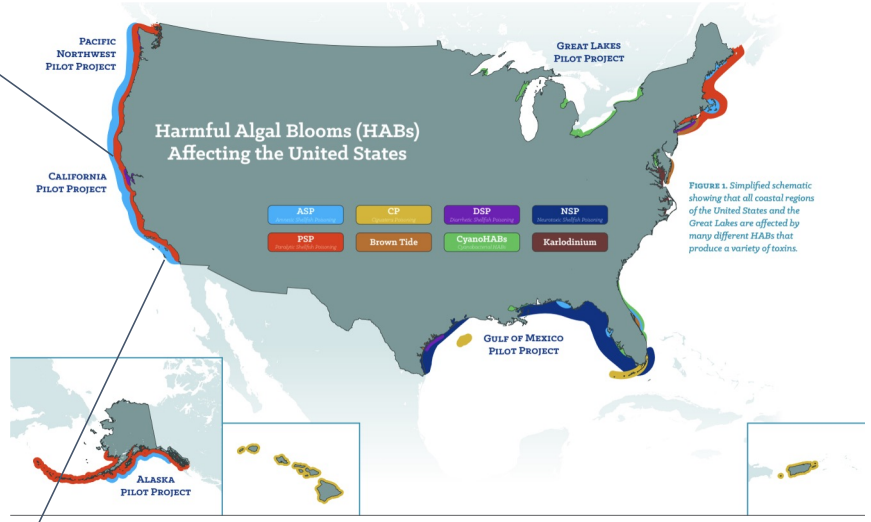
Customizable data product resolution (hourly, daily, weekly etc.) and Easy download

California Harmful Algal Bloom Early Warning System

Largest network of robotic microscopes in the world with 12 operational Imaging FlowCytobots

OCEAN PROTECTION COUNCIL · IOOS Integrated Ocean Observing System · SOUTHERN CALIFORNIA COASTAL OCEAN OBSERVING SYSTEM · CENTRAL & NORTHERN CALIFORNIA OCEAN OBSERVING SYSTEM · NCCOS · OCSAN ORANGE COUNTY SANITATION DISTRICT





NHABON Pilot Project at SCCOOS + CeNCOOS helps support **9** operational IFCBs at **7** piers and **2** at offshore moorings (12 total in CA Network)
https://ifcb.caloos.org/

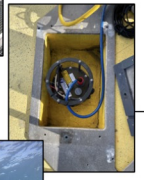
SCRIPPS INSTITUTION OF OCEANOGRAPHY UC San Diego | WOODS HOLE OCEANOGRAPHIC INSTITUTION | Axiom DATA SCIENCE A TETRA TECH COMPANY | UNIVERSITY OF CALIFORNIA SANTA CRUZ | CENTRAL & NORTHERN CALIFORNIA OCEAN OBSERVING SYSTEM

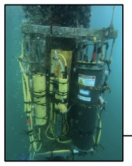
National HAB Observing Network

The NHABON CoP fosters scientific collaboration, information sharing, and other activities leading to the development, operation, and application of the NHABON at local, state, and regional levels

California IFCB Network

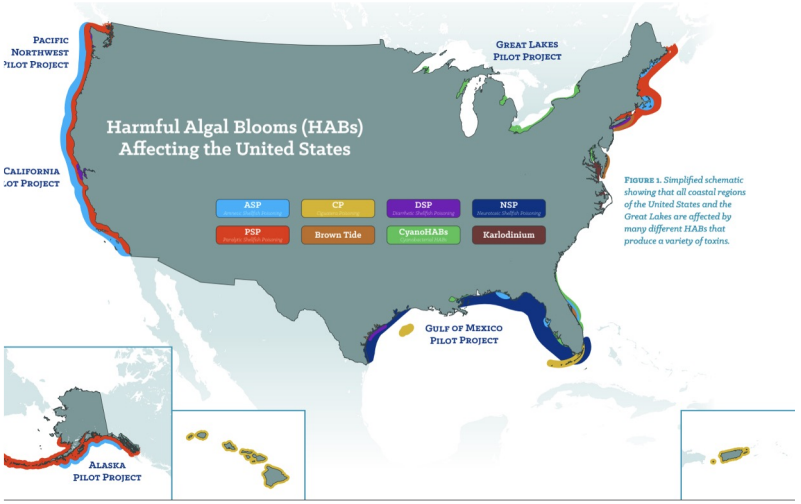
Technical support
- Monthly technical meetings
- Troubleshooting protocols in the CA IFCB Network's Standard Operating Procedures (SOP)
- Data quality control – focus on reliable concentration estimates





1 GB/day

https://library.ucsd.edu/dc



NHABON Pilot Project at SCCOOS + CeNCOOS elps support **9** operational IFCBs at **7** piers and **2** at offshore moorings (12 total in CA Network)
https://ifcb.caloos.org/

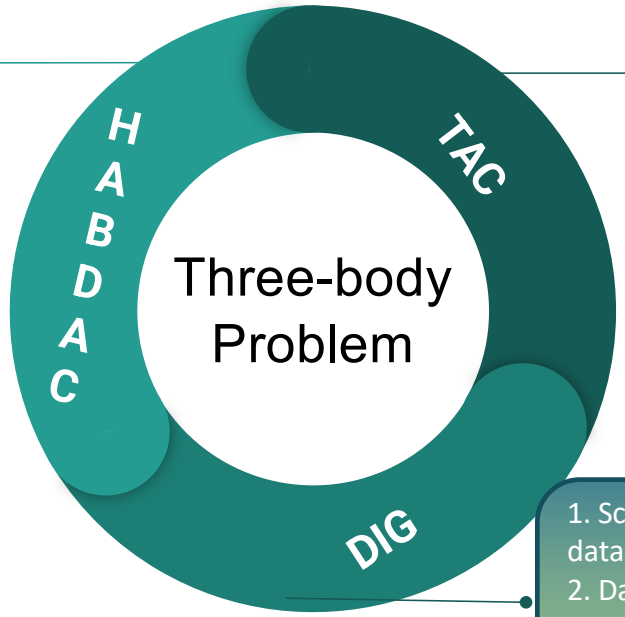
Scripps Institution of Oceanography, UC San Diego · Woods Hole Oceanographic Institution · Axiom Data Science, A Tetra Tech Company · University of California Santa Cruz · Central & Northern California Ocean Observing System

California IFCB Network

Diverse institutional involvement and deployment configurations across the network requires redundant data archival and a centralized cyberinfrastructure

HABDAC Team

Clarissa Anderson, SIO
Rob Bochenek, ADS
Heidi Sosik, WHOI
Stace Beaulieu, WHOI
Raphe Kudela, UCSC
Henry Ruhl, MBARI
Patrick Daniel, UCSC
Joe Futrelle, WHOI
Karina Khazmutdinova, ADS
Kasia Kenitz, SIO
Andrew Barton, SIO



1. Stand up the TAC
2. Central CI for CA Network
3. Ingest new data/models - new regions
4. Centralize ML models and data sets
5. Transition Plan for tech transfer – SOP/Best Practices

Transition Advisory Committee

1. Downstream products
2. Broaden conversation
3. Regional stakeholders
4. Synergy w/ programs
5. End-user requirements
6. Broaden participation
7. Help sustain HABDAC

Barb Kirkpatrick, GCOOS
Jan Newton, NANOOS
Tom Shyka, NERACOOS
Lisa Campbell, Texas A&M
Andrew Barton, SIO
Abby Benson, USGS/OBIS
Noah Ben-Aderet, OPC
Mike Brosnahan, WHOI
Stephanie Moore, NWFSC
Tenaya Norris, TMMC
Jules Jaffe, SIO
Felix Martinez, NCCOS

Data Integration Group

1. Scope of HAB data & data types to be ingested
2. Data levels & process standardization
3. Metadata format & content

Kendra Negrey, UCSC
Patrick Daniel, UCSC
Eric Orenstein, MBARI
Abby Benson, USGS/OBIS
Mike Brosnahan, WHOI
Kate Hubbard, FWC-FWRI
Marco Corrales, GFDL

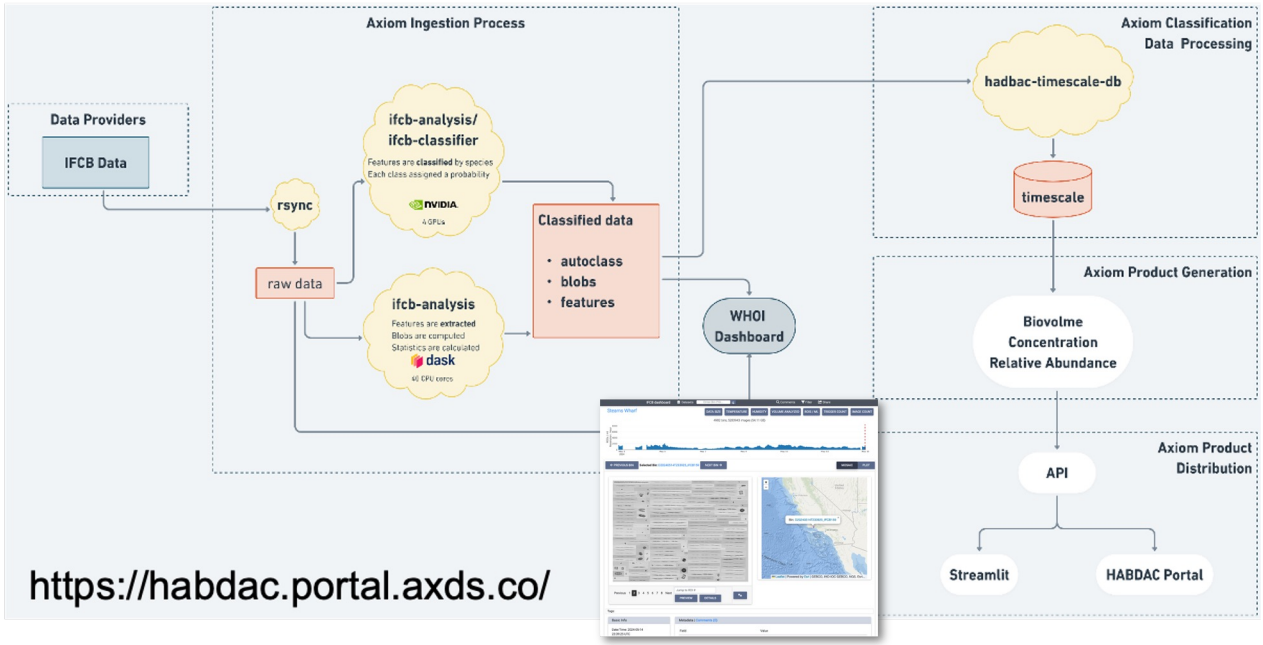
Scripps Institution of Oceanography, UC San Diego · Woods Hole Oceanographic Institution · Axiom Data Science, A Tetra Tech Company · University of California Santa Cruz · Central & Northern California Ocean Observing System

PCMHAB 2020: Harmful Algal Bloom Community Technology Accelerator

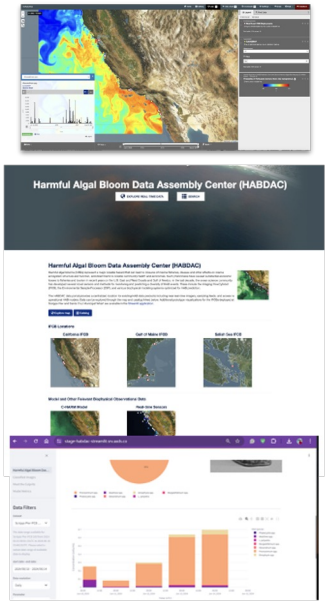
NOAA NCCOS funding awarded in 2020 to establish a national HAB Data Assembly Center that uses the CA & GoM IFCB Networks as prototypes

x12

HAB DAC development



CalOOS Catalog



SCRIPPS INSTITUTION OF OCEANOGRAPHY UC San Diego

WOODS HOLE OCEANOGRAPHIC INSTITUTION

Axiom DATA SCIENCE A TETRA TECH COMPANY

UNIVERSITY OF CALIFORNIA SANTA CRUZ

CENTRAL & NORTHERN CALIFORNIA OCEAN OBSERVING SYSTEM

PCMHAB 2020: Harmful Algal Bloom Community Technology Accelerator

VISION: Automated services to all IFCB users to immediately stand up an IFCB Dashboard and generate downstream products for science AND management

Value-added products

Reduce barriers, promote efficiencies, and positively engage varied community participants

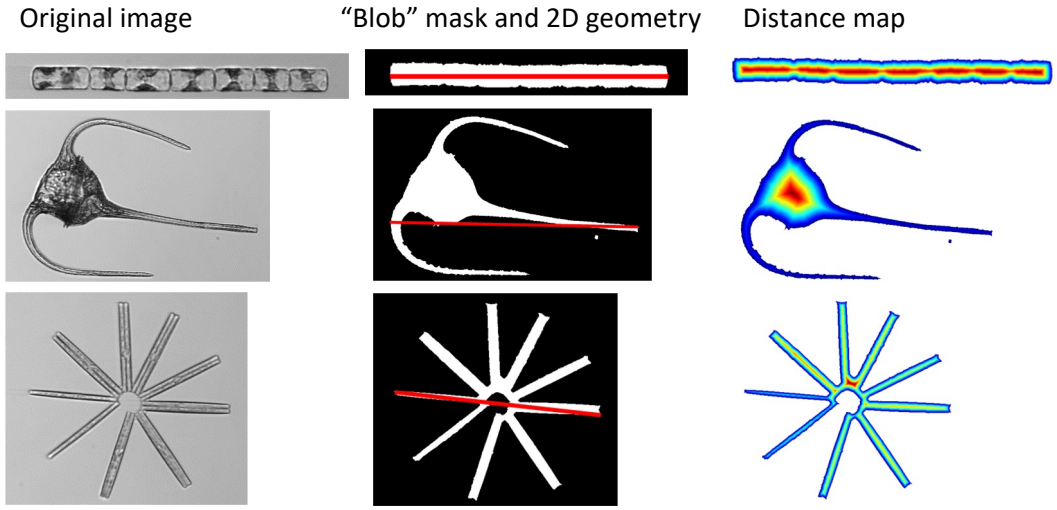
- IFCB users / data producers
- End users: research scientists, resource managers, ecosystem modelers, etc.

⇒ **Centralized configuration and management of system components for data access and product workflows**

Automated data accession
Automated image product generation
Automated image product access

Classifier training and evaluation tools
On-demand classifier application
Automated level 2 and 3 product generation
including summary time series and maps

Automated image processing for trait estimation



MATLAB development environment → Python operational environment

SCRIPPS INSTITUTION OF OCEANOGRAPHY UC San Diego
WOODS HOLE OCEANOGRAPHIC INSTITUTION
Axiom DATA SCIENCE A TETRA TECH COMPANY
UNIVERSITY OF CALIFORNIA SANTA CRUZ
CENTRAL & NORTHERN CALIFORNIA OCEAN OBSERVING SYSTEM

PCMHAB 2020: Harmful Algal Bloom Community Technology Accelerator

Data Archive and Access: Images (ROIs)Metadata >>> Features >>> Image Classification >>> Automated Level 1 to Level 3 Product Generation

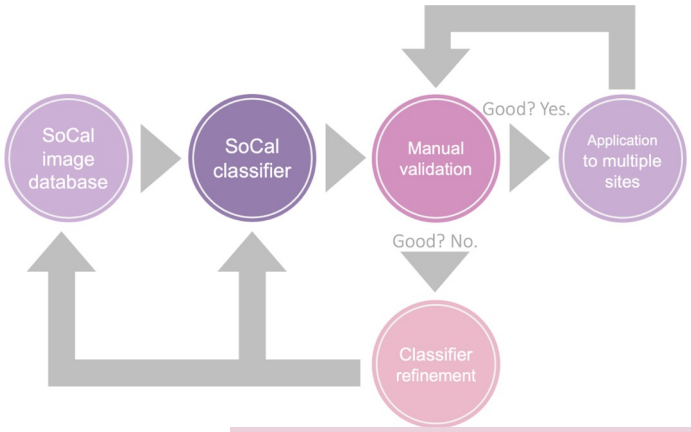
Raw Data:
Images (ROIs)
Metadata

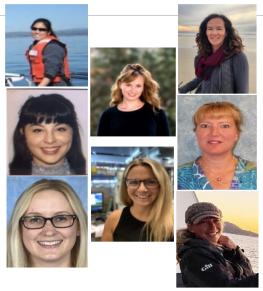
Features

Image classification

Sample-level concentrations:
- Abundance
- Biovolume
- Carbon

Time series data





- Annotation effort led by a group of expert taxonomists
- Standardized annotation and classification workflow across all sites
- Image database available for sites without taxonomic expertise
- Centralized automated classification of plankton images for the network available in real-time

SOCAL

- Convolutional Neural Network
- Inception_v3 architecture, pretrained with ImageNet
- 131 classes, including 33 diatom, 38 dinoflagellate and 20 ciliate categories
- Average F1 score = 0.87

Akashiwo: 0.89
Dinophysis: 0.96
Lingulodinium: 0.92
Phaeocystis: 0.98
Prorocentrum: 0.96
Pseudo-nitzschia: 0.93

NORCAL

- Convolutional Neural Network
- Xception architecture, pretrained with ImageNet
- 51 classes, including 22 diatom, 16 dinoflagellate and 4 ciliate categories
- Average F1 score = 0.58

Akashiwo: 0.95
Dinophysis: 0.92
Lingulodinium: 0.56
Phaeocystis: 0.81
Prorocentrum: 0.92
Pseudo-nitzschia: 0.62

Building a CoP around community annotation

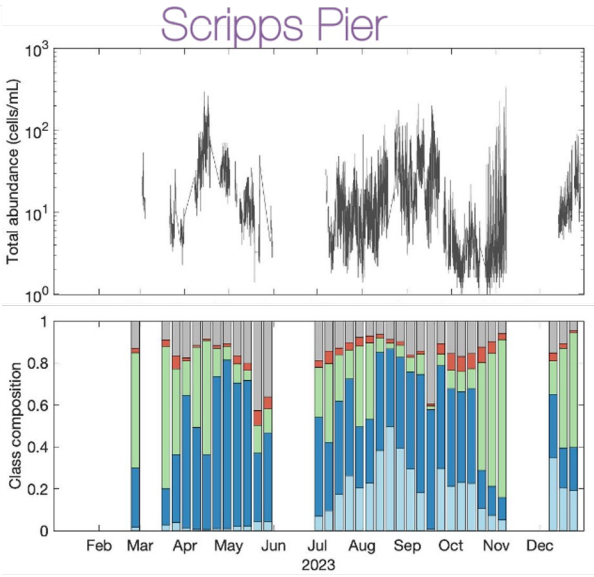
Slide courtesy of Kasia Kenitz

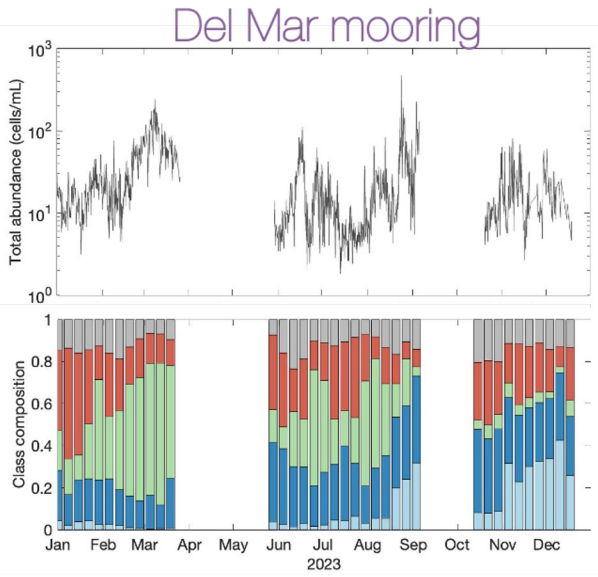
SCRIPPS INSTITUTION OF OCEANOGRAPHY UC San Diego | WOODS HOLE OCEANOGRAPHIC INSTITUTION | Axiom DATA SCIENCE A TETRA TECH COMPANY | UNIVERSITY OF CALIFORNIA SANTA CRUZ | CENTRAL & NORTHERN CALIFORNIA OCEAN OBSERVING SYSTEM

PCMHAB 2020: Harmful Algal Bloom Community Technology Accelerator

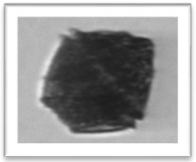
Building and sharing annotated image libraries and regional classifiers is essential to progress

NEARSHORE





OFFSHORE



L. polyedra Dinoflagellates Diatoms Ciliates Other

SCRIPPS INSTITUTION OF OCEANOGRAPHY UC San Diego

WOODS HOLE OCEANOGRAPHIC INSTITUTION

Axiom DATA SCIENCE A TETRA TECH COMPANY

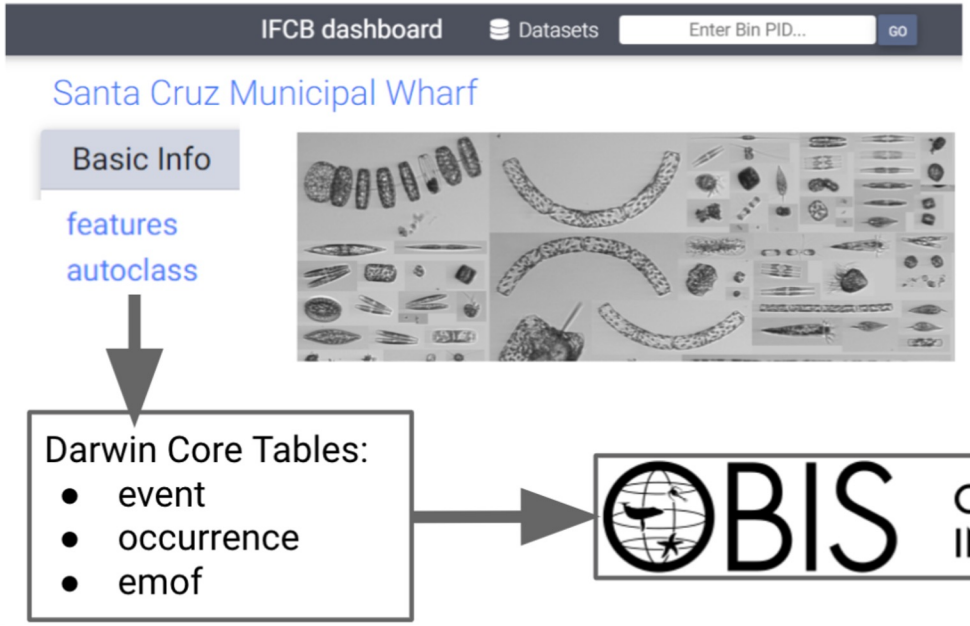
UNIVERSITY OF CALIFORNIA SANTA CRUZ

CENTRAL & NORTHERN CALIFORNIA OCEAN OBSERVING SYSTEM

PCMHAB 2020: Harmful Algal Bloom Community Technology Accelerator

Regional CNN classifiers appear necessary for accurate HAB assessment but can be repurposed and retrained

Downstream product development: OBIS use case



Building FAIR data products from IFCB autoclassification and size data

Success!

Link to our notebook

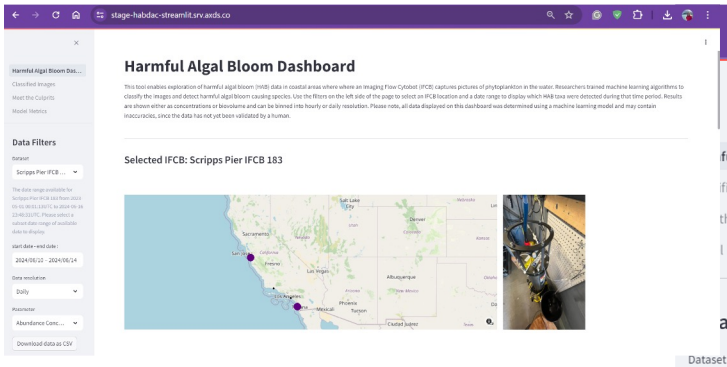
Implemented EU Horizon 2020 "Best practices and recommendations for plankton imagery data management"

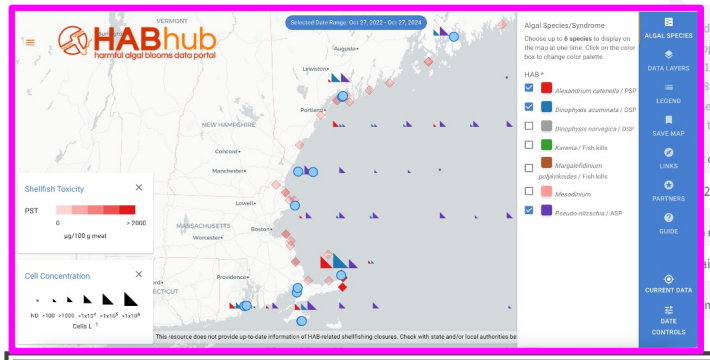
Slide courtesy of Stace Beaulieu & Ian Brunjes

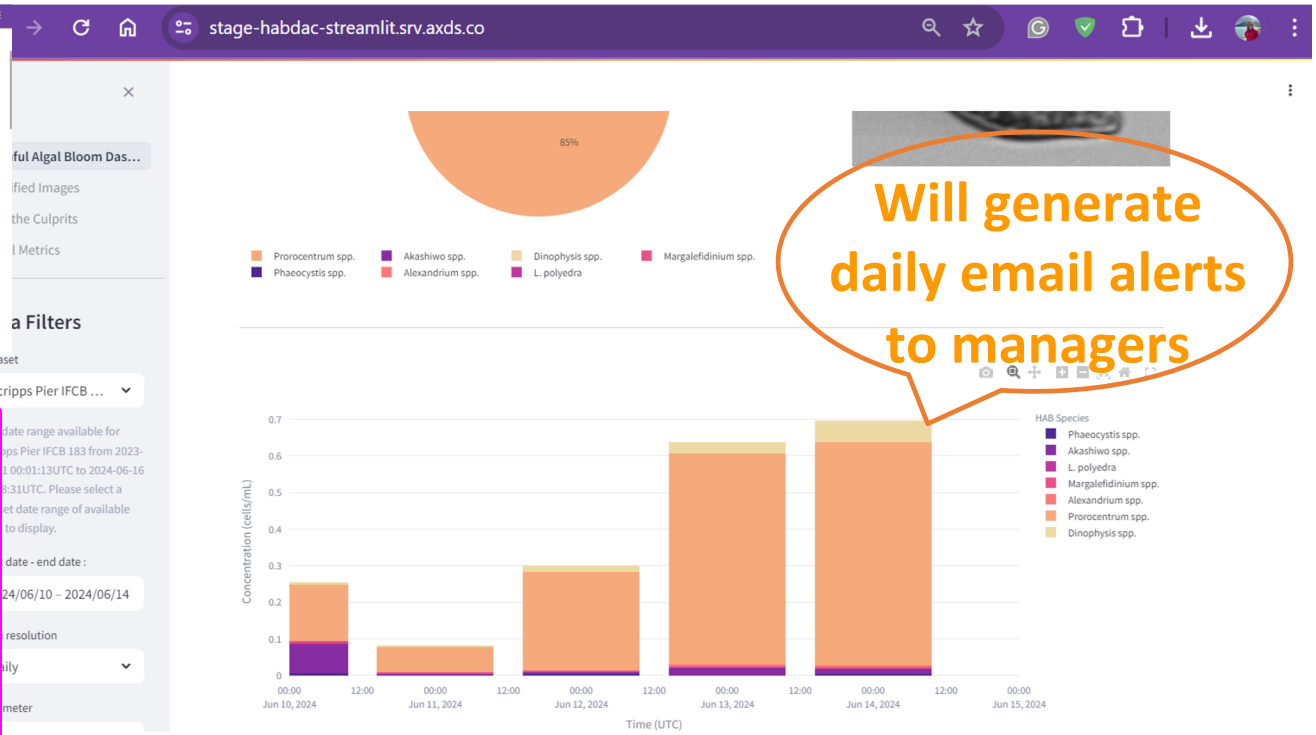
SCRIPPS INSTITUTION OF OCEANOGRAPHY UC San Diego · WOODS HOLE OCEANOGRAPHIC INSTITUTION · Axiom DATA SCIENCE A TETRA TECH COMPANY · UNIVERSITY OF CALIFORNIA SANTA CRUZ · CENTRAL & NORTHERN CALIFORNIA OCEAN OBSERVING SYSTEM

PCMHAB 2020: Harmful Algal Bloom Community Technology Accelerator

- *Building a code repository for community archival of complex image and classifier data*
- *NASA PACE project (PI Anderson) generating new routines for pushing IFCB data to SeaBASS*







SCRIPPS INSTITUTION OF OCEANOGRAPHY
UC San Diego
WOODS HOLE OCEANOGRAPHIC INSTITUTION
Axiom DATA SCIENCE
A TETRA TECH COMPANY
UNIVERSITY OF CALIFORNIA SANTA CRUZ
CENTRAL & NORTHERN CALIFORNIA OCEAN OBSERVING SYSTEM

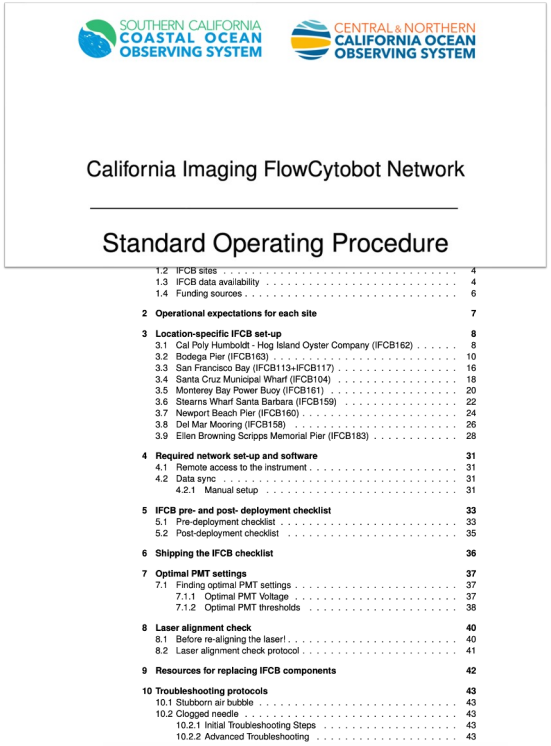
PCMHAB 2020: Harmful Algal Bloom Community Technology Accelerator

What do managers need and want? How do we develop best practices for generating alerts?

HABDAC Streamlit app being vetted in concert with apps like WHOI HABHub (M. Brosnahan)

Where are we headed with IRA funding over next two years?

1) Plug and play – user can bring an IFCB online almost immediately, i.e. ability to rapidly view sample images on dashboard
2) Classifier Menus – existing classifiers (CNNS, RF, etc) from different regions available for testing and/or running side-by-side
3) Annotated Image Libraries – libraries searchable by region; easily rearranged for training set generation
4) Automated Product Development – L1-L3 and management-ready products, merging all available HAB data (toxins, cell counts, models…)
5) Script Libraries for Data Archiving - customized routines to push data to OBIS, NASA SeaBASS, etc. in agency-compliant formats
6) Regional Diversity – **We have CA, (some) Gulf of Maine, Salish Sea!!!** What about Florida?, Texas?, Pacific Northwest? OOI IFCBs!!!
7) Management Products People Want! - feedback from diverse end-users


SOUTHERN CALIFORNIA COASTAL OCEAN OBSERVING SYSTEM
CENTRAL & NORTHERN CALIFORNIA OCEAN OBSERVING SYSTEM

California Imaging FlowCytobot Network

Standard Operating Procedure

SCRIPPS INSTITUTION OF OCEANOGRAPHY UC San Diego
WOODS HOLE OCEANOGRAPHIC INSTITUTION
Axiom DATA SCIENCE A TETRA TECH COMPANY
UNIVERSITY OF CALIFORNIA SANTA CRUZ
CENTRAL & NORTHERN CALIFORNIA OCEAN OBSERVING SYSTEM

PCMHAB 2020: Harmful Algal Bloom Community Technology Accelerator

END GOAL: Create a truly useful community of practice and community experience that stands up to computational and instrumentation advancements

Harmful Algal Bloom Data Assembly Center (HABDAC)

EXPLORE REAL-TIME DATA | SEARCH

PLEASE JOIN OUR COMMUNITY OF PRACTICE

clrander@ucsd.edu

Harmful Algal Bloom Data Assembly Center (HABDAC)

Harmful algal blooms (HABs) represent a major coastal hazard that can lead to closures of marine fisheries, disease and other effects on marine ecosystem structure and function, and direct harm to coastal community health and economies. Such phenomena have caused substantial economic losses to fisheries and tourism in recent years on the U.S. East and West Coasts and Gulf of Mexico. In the last decade, the ocean science community has developed several novel sensors and methods for monitoring and predicting a diversity of HAB events. These include the Imaging FlowCytobot (IFCB), the Environmental Sample Processor (ESP), and various biophysical modeling systems optimized for HAB prediction.

The HABDAC data portal provides a centralized location for existing HAB data products including near real-time imagery, sampling feeds, and access to operational HAB models. Data can be explored through the map and catalog linked below. Additional prototype visualizations for the IFCBs deployed at Scripps Pier and Santa Cruz Municipal Wharf are available in the Streamlit application.

Explore map | Catalog

IFCB Locations

California IFCB | Gulf of Maine IFCB | Salish Sea IFCB

Model and Other Relevant Biophysical Observational Data

C-HARM Model | Real-time Sensors

Release notes

https://habdac.portal.axds.co/

https://ifcb.caloos.org/

Data Disclaimer

All products published on this website are prototype products and are not intended to be used for navigational or operational purposes. This website provides an aggregation and display service for distributed data and modeling products that come from a wide array of contributors. We are grateful to those contributors and make every attempt to give them credit for their products. We welcome suggestions for clarification or correction, as appropriate.