Collecting Critical Climate Data from the NSF's Ocean Observatories Initiative Regional Cabled Array

Wendi Ruef, Mariela White, Joe Duprey, Mike Vardaro, Deb Kelley

School of Oceanography, University of Washington

wruef@uw.edu

The Array

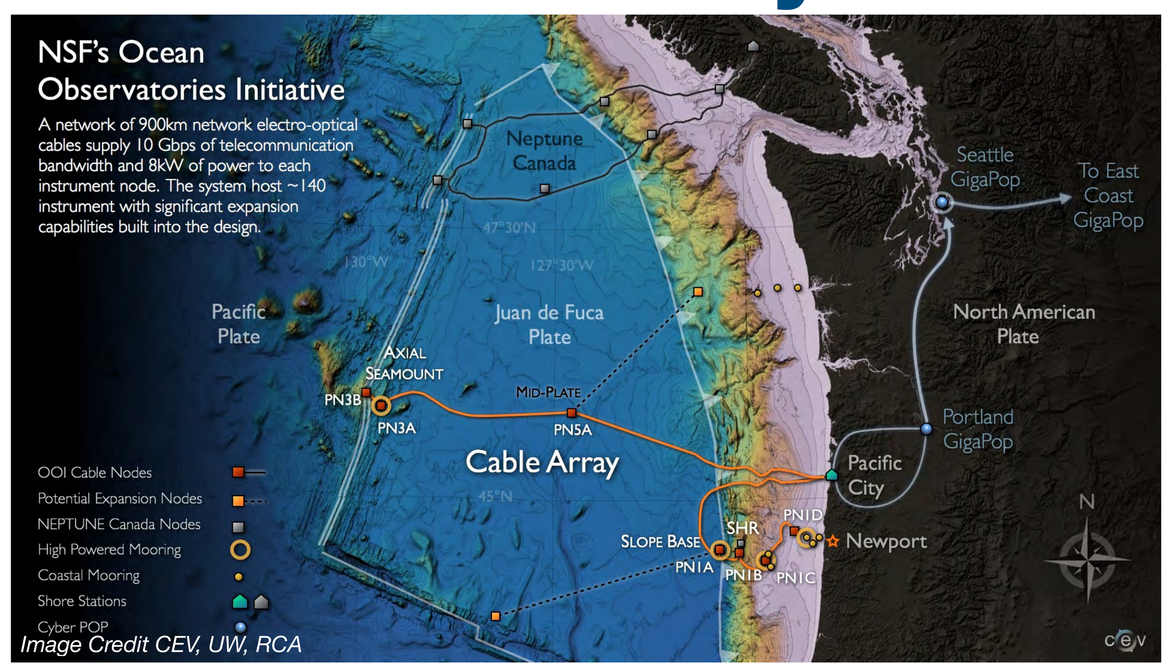


Image Credit CEV, UW, RCA

The NSF-funded Ocean Observatories Initiative Regional Cabled Array (OOI-RCA) is a state-of-the-art cabled submarine network delivering high-resolution physical, biological, chemical, and geological data from over 150 seafloor and water-column sensors since its deployment in the Northeast Pacific in 2014. Designed and maintained by the University of Washington, the OOI-RCA uses high-bandwidth optical fiber cables to provide two-way communication and real time in-situ data from a diverse set of ocean environments, including highly productive waters along the Cascadia Margin, methane seeps on the continental slope, and active hydrothermal vents at the Axial Seamount volcano 300 miles offshore, the largest and most active volcano off of the OR-WA coast.

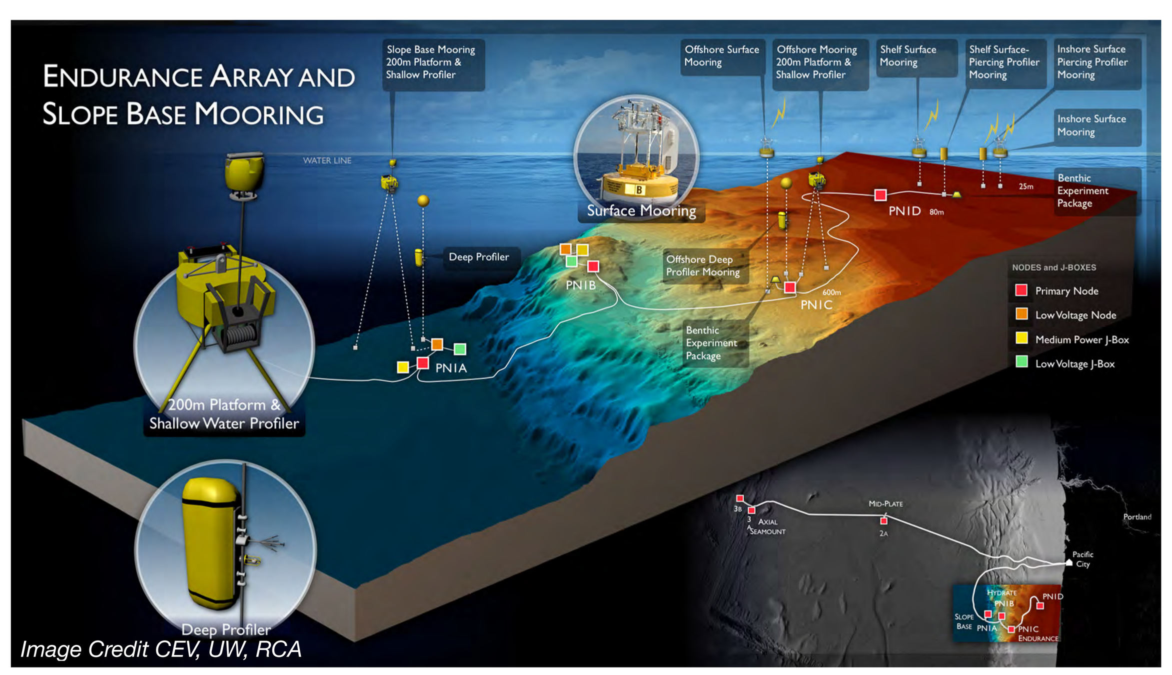


Image Credit CEV, UW, RCA

A representative swath of the Regional Cabled Array can be seen at the Endurance and Slope Base sites diagrammed above, which span from the inshore coastal shelf at 80 meters out to the continental slope at 2900 meters. These sites include a diverse set of seafloor sensors coupled with water column platforms and innovative profiling moorings that measure hourly a multitude of oceanographic parameters throughout the entire water column spanning 2900 m water depths to ~5 m beneath the surface. A small snapshot of data from the Slope Base shallow profiler are highlighted to the right, with profiles contoured and plotted on the left and center, and climatological anomalies plotted on the far right. Data are open-access, available in real-time, and viewable through the OOI Data Explorer portal.

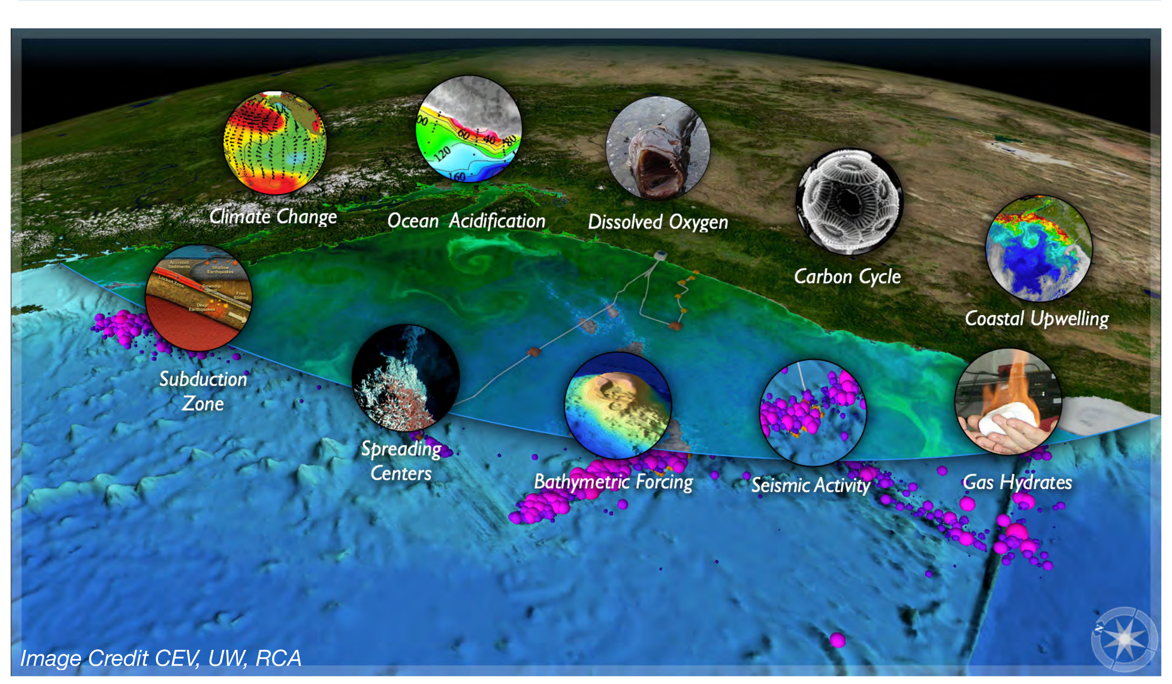
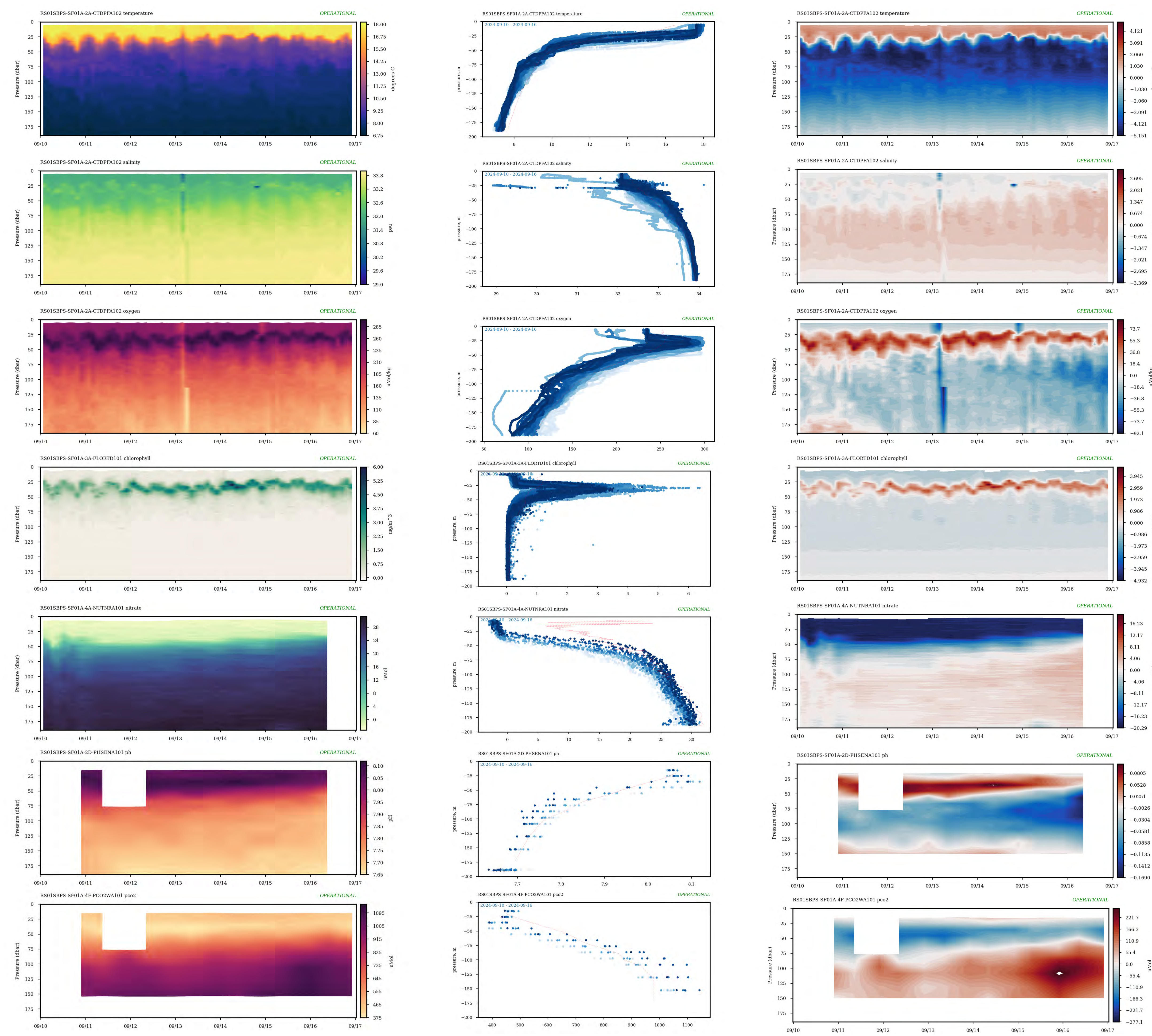


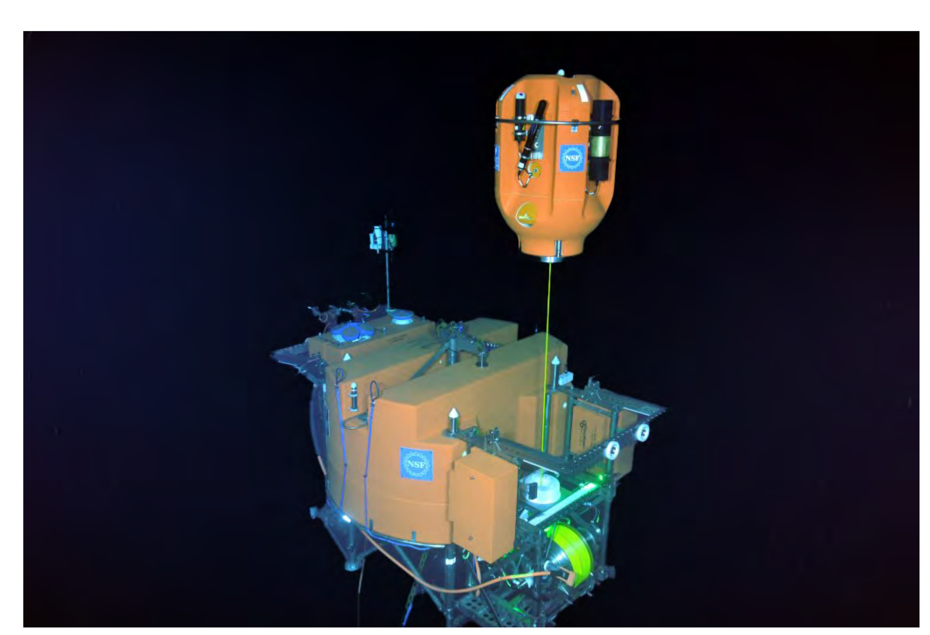
Image Credit CEV, UW, RCA

Most RCA datasets are available in real-time through satellite telemetry and the high-bandwidth Regional Cabled network in the Northeast Pacific, which allows for interactive, rapid adaptive sampling changes in response to events such as storms, volcanic eruptions, and earthquakes. With over 150 individual sensors deployed, these data comprise a rich set of opportunities to gain insight into the complexities of ecosystem variability and myriad ocean processes and environmental challenges that are difficult to address using traditional ship-based sampling, such as water column dynamics, thin layer formation, ecosystem response to storms, internal waves, coastal upwelling, ocean acidification, hypoxia events, tsunami waves, volcanic activity, and seismicity over the greater Cascadia Subduction Zone.

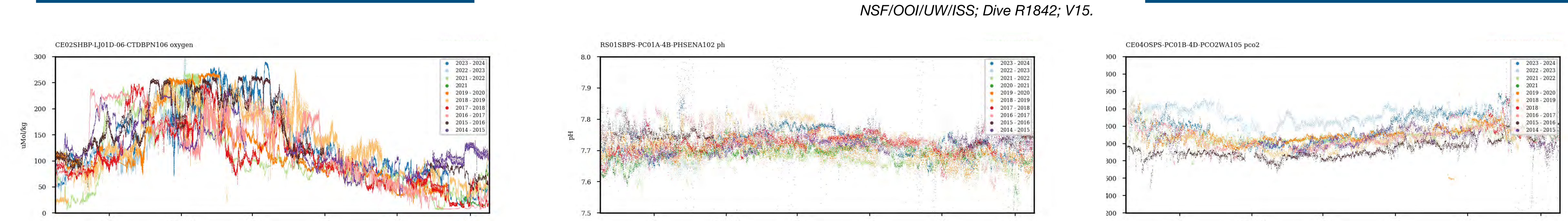
Data



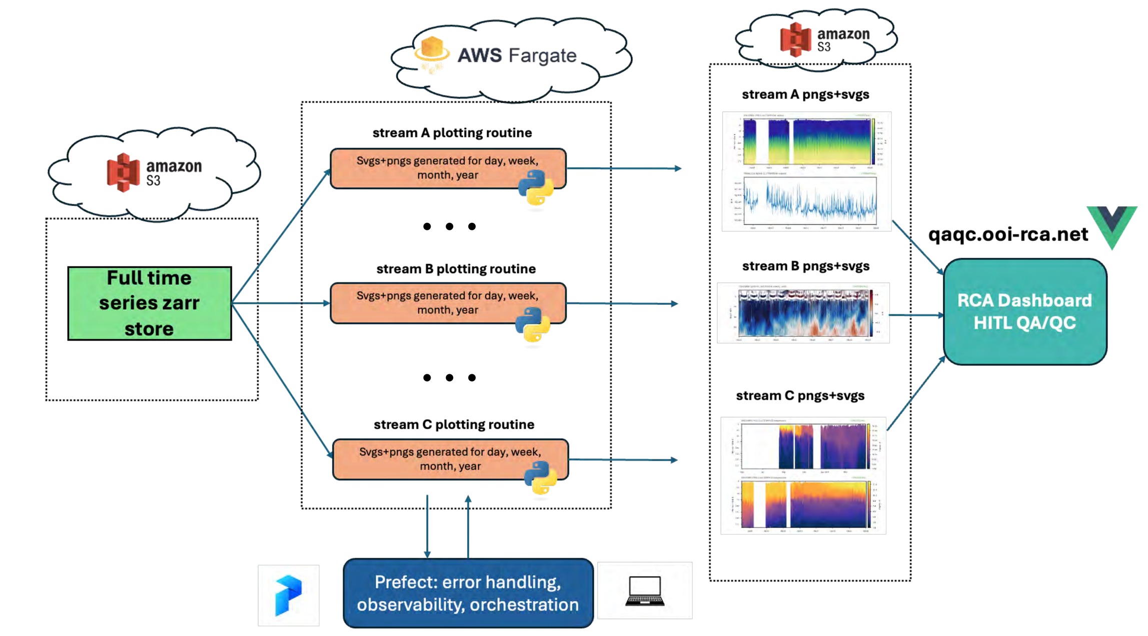
Carbon and oxygen are among some of the most critical climate parameters that are also some of the most sparsely sampled offshore. Long-term high-frequency datasets are imperative to understanding current and future environmental challenges in a rapidly changing ocean. Below are three exemplar time-series for oxygen, pH, and pCO2 at select sites across the RCA that include hourly samples from 2014 though present. These datasets, along with co-located instruments measuring water temperature, salinity, oxygen, nutrients, chlorophyll concentrations, spectral irradiance, and current velocities and coupled with discrete sample validation, provide an unparalleled opportunity for in-depth climate research.


NSF/OOI/UW/ISS; Dive R1842; V15.

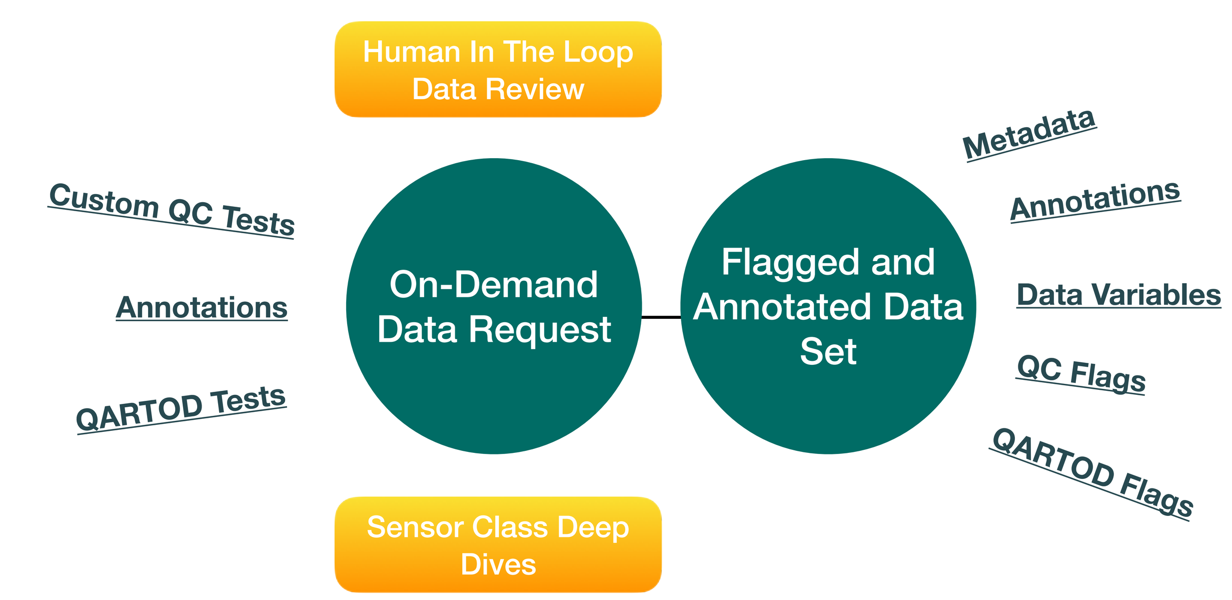
Above we highlight a snapshot of key climate datasets available through the OOI data portal, which includes a wealth of unparalleled high-frequency long-term time-series from sensors deployed on the seafloor, profiling vehicles and 200-meter platforms spanning coastal to blue water environments in the NE Pacific. The shallow profiler and platform sensors provide co-located multi-parameter datasets from up to 18 science instruments for investigation of biogeochemical cycling on timescales of days to weeks to years. The RCA shallow profilers have completed a combined total of over 50,000 profiles since 2015 across three sites, with comparable instrument suites to allow for cross-basin investigations.



QA/QC



Critical to the utility of the OOI-RCA datasets are the underlying quality assurance and quality control (QA/QC) protocols that allow users to evaluate the condition of the data and provide necessary context for robust analysis and interpretation. The increasingly large and complex datasets collected through the RCA, including over 30 sensor types and novel technologies unique to the RCA, require new QA/QC protocols to view and fully evaluate. RCA has developed a harvest and visualization pipeline utilizing cloud-friendly Zarr files and leveraging multiple resources within the AWS cloud to reduce processing and memory needs and allow use of existing open-source cloud orchestration software. Scaleable resources and parallel processing allow efficient human-in-the-loop review and data annotations.



The OOI-RCA QA/QC workflow consists of multiple review stages to the data. Automated algorithms in CI produce data flags for each data stream on-demand. These algorithms identify data spikes, stuck values, and data that fall outside of global ranges, as well as values that fall outside of QARTOD gross and local range and climatology limits. In addition to the automated filters a regular human-in-the-loop (HITL) data review is performed, along with sensor-class deep dives that include custom data filters to identify suspect and failed data through raw data ratios, signal ranges, and vendor calibration ranges. All HITL results are stored as data annotations and include a separate flag to exclude failed data, delivered with on-demand data requests and viewable through the online OOI Data Explorer.

Future RCA Quality Control

Upcoming Automated and HITL Dashboard Enhancements

- Discrete sample validation overlays
- Nearest Neighbors overlays
- Property-property analysis
- Model comparison overlays
- Event detection alerts
- Full implementation of all relevant QARTOD flags
- Custom sensor-specific QC flags

Machine Learning and AI

- Informed outlier detection
- Nearest neighbor comparisons
- Profile shape analysis
- Enhanced event detection and prediction
- Image and acoustic processing and annotation
- Large scale correlation analyses

The overarching principle of quality control for RCA is the same as that of OOI as a whole: to ensure that data being served are research-ready, meaning users have easy access to quality descriptors for each observation and have the context to determine which data can be incorporated into their analysis. With instrument technology advances allowing higher frequency sampling and more complex sensor data collection along with increasing compute power, the QA/QC workflow is a continuously evolving process with new opportunities to refine and improve the automated and HITL workflows to not only flag suspect data, but also allow event detection and a deeper understanding of key climate phenomena on timescales ranging from minutes/hours to days/weeks to seasonal/inter-annual variability.