

# CI and Data Delivery Update

Jeffrey Glatstein

May 13<sup>th</sup>, 2020





# Agenda

- PY II work in progress
  - Achievements To Date
  - Data Discovery Tool (Keryx project)
  - Asset Meta Data Management (Roundabout project)
  - Large Efforts
- PY III planned work highlights



# Achievements To Date

- **Redmine ticket statistics**
  - 443 tickets open at start of PY II
  - 354 tickets open as of 4/30 – approximately 20% reduction
  - 157 tickets were opened in PY II
  - 257 tickets were closed in PY II
- **Highlighted achievements**
  - Improved ADCP data with focused efforts across bugs and enhancements (started in PYI, 20+ tickets)
  - Enhanced ability to utilize data on co-located instruments
  - Restructured the Raw data server for cabled data
  - Transitioned Redmine server and management to Software Administration
  - Improved data management
    - Added ability to replay data with time bounds through existing ingestion process (multi-year effort)
    - Added ability to delete data by stream id
  - Communication of Meta Data clean-up to end users
    - Implemented widget that allows user to search for changes
    - Added banner to OOINET to announce changes
    - Mailchimp campaign, website articles and new pages created to support communications





# Data Discovery Tool (Keryx)

- User and MIO interviews completed – personas and features identified
- Roadmap of functionality and tasks reviewed and approved by PIs
- Science data staged and refreshed daily
- Testing of largest data sets into the new tool – bottom up approach
- Determine parameters by priority to be loaded into new tool
- Determine provenance parameters to move into NetCDF files
- Align styling with new website development by Trapeze
- Milestones
  - June 15<sup>th</sup> - Data and metadata ingested for prioritized parameters to support beta release
  - June 15<sup>th</sup> - Beta release for internal review
  - July 1<sup>st</sup> – Start focused sessions with selected end users and feedback prioritization for development into v1. User interaction and interviews will widen to include more audiences as time permits based upon feedback
  - July 30<sup>th</sup> - PY3 work plan
  - July - Sept- Data and metadata ingested for prioritized parameters to support version 1
  - Sept- Release version 1 internally
  - Oct 1 – Version 1 public release (Note: OOINET will continue to be available)



# Asset Meta Data Management (Roundabout)

- PY II deliverable will be an Asset metadata management interface that can be populated from the current GitHub repository
- Development team and cross MIO have gathered requirements and reviewed them with PIs
- Data delivery code analyzed for impact to current method of retrieving metadata
- Calibration management in Roundabout coding in progress
- Advanced search feature coding in progress





# Large Efforts

- **Quality**
  - Development of the Climatology QARTOD compliant test is well underway
  - Standardization of naming conventions and NetCDF file builds
  - Milestones
    - June 2<sup>nd</sup> – Migrate Gross Range test into production for CTD
    - September 2<sup>nd</sup> – Migration of Climatology test to production for CTD
- **Data Center Provider Evaluation – panel formed and RFP developed**
  - Timeline for evaluation
    - 5/1/20: Release RFP
    - 6/12/20: RFP response deadline
    - 7/13/20 – 8/14/20: Responder interviews
    - 9/20: Recommendation
- **Cassandra Architecture Review**
  - DataStax engaged to review DB health, current configuration and data model.
  - 134 page analysis with recommendations prioritized by immediate (9), near term (15) and long term (3).
  - Recommendations of note:
    - Upgrade Cassandra version to 2.2.16 then to latest version of 3.x
    - Significantly reduce partition size – requires reorganization of database
    - Increase number of nodes and replace disks with SSD drives



# PY III Planned Work Highlights

- Work with CI Systems sub-awardee to implement new system architecture in a new data center or cloud
- Upgrade of uFrame components
- QA/QC implementation of QARTOD code
- Asset management migrations – includes Roundabout development
- Keryx development
  - Implement login feature
  - Combine features of OOINET with Keryx for single user experience
  - Jupyter notebook access







OCEANDESERVATORIES.ORG





# Supporting Details







# Quality Control

- Implementation of QARTOD compliant Gross Range test in production for CTD scheduled for June release.
- Development of the Climatology QARTOD compliant test
- Researching Timing Test
- Standardization of naming conventions and NetCDF file builds
  - Pre-load database review currently in process by working group
  - NetCDF redmine tickets being worked
- Milestones
  - June 2<sup>nd</sup> – Migrate Gross Range test into production for CTD
  - September 2<sup>nd</sup> – Migration of Climatology test to production for CTD





# Data Center Provider Evaluation

- Rutgers University is currently responsible CI Systems scope and will be moving that scope to a new sub-awardee during PY III
- Process to select new sub-awardee kicked off 5/1/20
- Panel has been formed to author an RFP, evaluation matrix and evaluate responses
- Timeline for evaluation
  - 5/1/20: Release RFP
  - 6/12/20: RFP response deadline
  - 7/13/20 – 8/14/20: Respondent interviews
  - 9/20: Recommendation





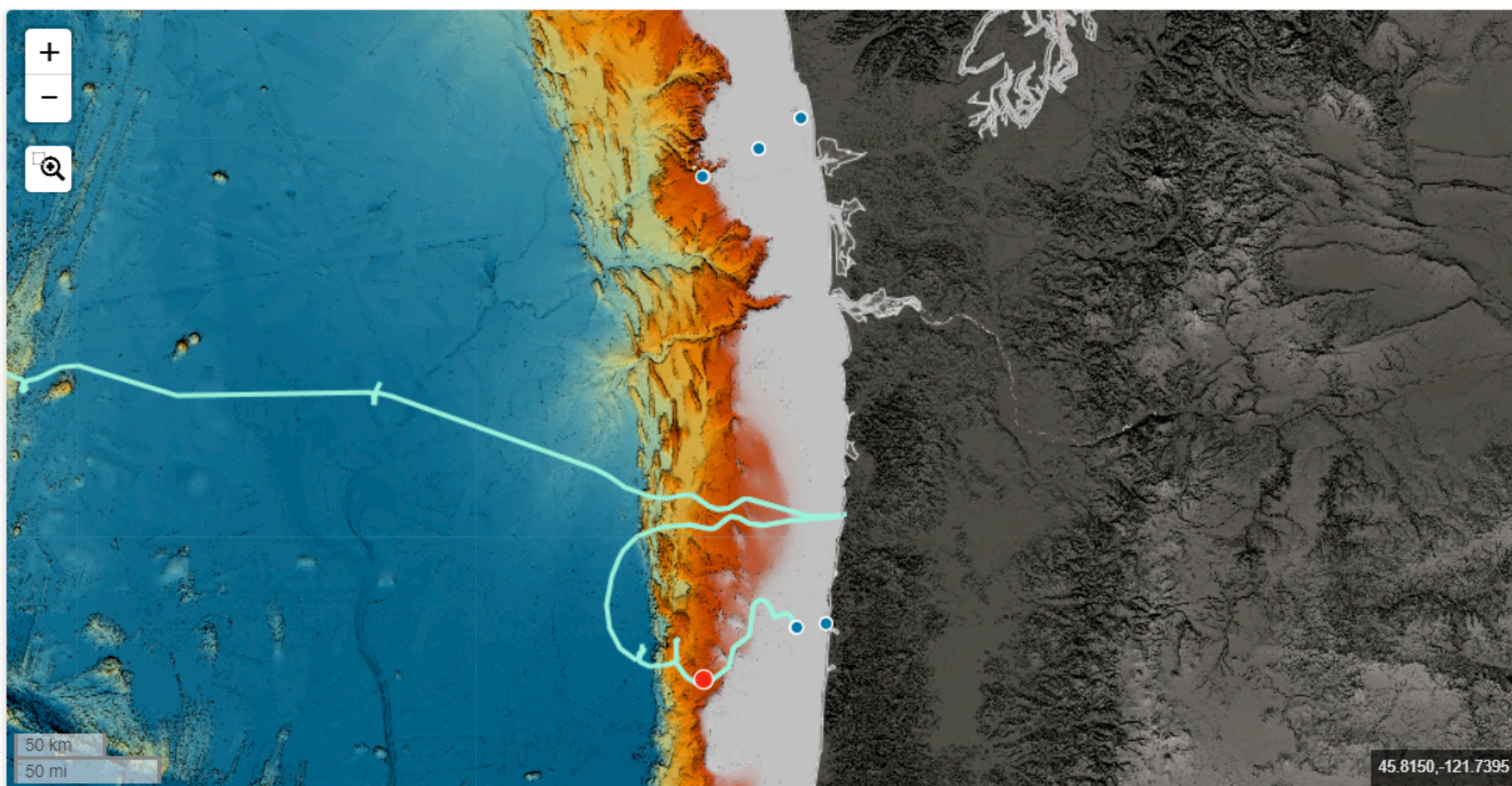
# Cassandra Architecture Review

- DataStax was engaged to do a health check of the current Cassandra implementation including review of current configuration and data model.
- Delivered 134 page highly detailed analysis with recommendations prioritized by immediate (9), near term (15) and long term (3).
- Recommendations of note:
  - Upgrade Java 8 version
  - Upgrade Cassandra version to 2.2.16 then to latest version of 3.x
  - Significantly reduce partition size – requires reorganization of database
  - Significantly reduce amount of data managed per node (currently 21)
  - Implement SSD disk technology
  - Normalize configuration across all nodes
  - Reduce number of tables
  - Adjust commit failure policy to *die*
  - Research and resolve long delays in garbage collection



Coastal Endurance > Oregon Offshore Cabled Shallow Profiler Mooring ▾

Overview Parameters Instruments Nodes





Nodes 18

Gliders 175

Arrays 7

Variable Types 25

Data Layers 7

## Coastal Endurance ▾

Overview

Parameters

Instruments

Nodes

Gliders

## Shallow Profiler (SF01B)(SF01B)

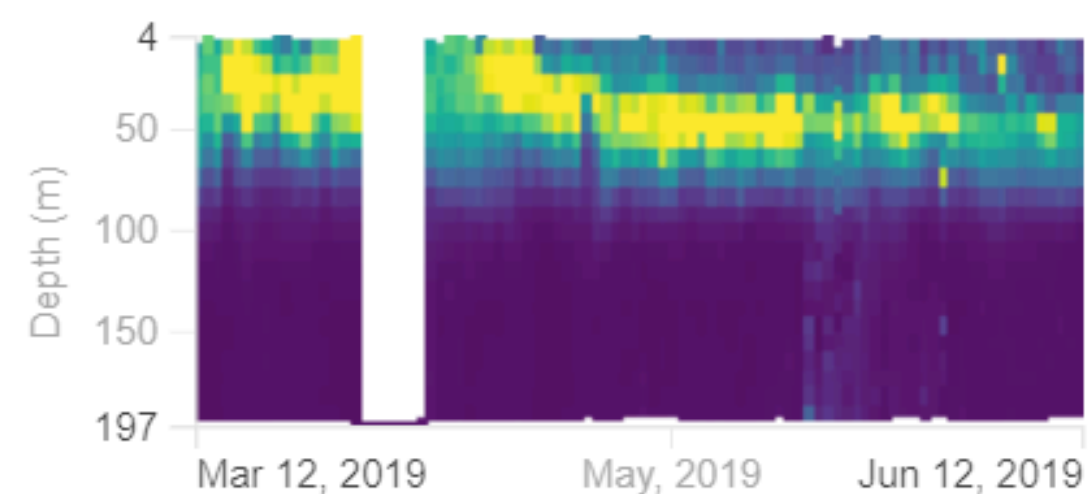
Data

Annotations

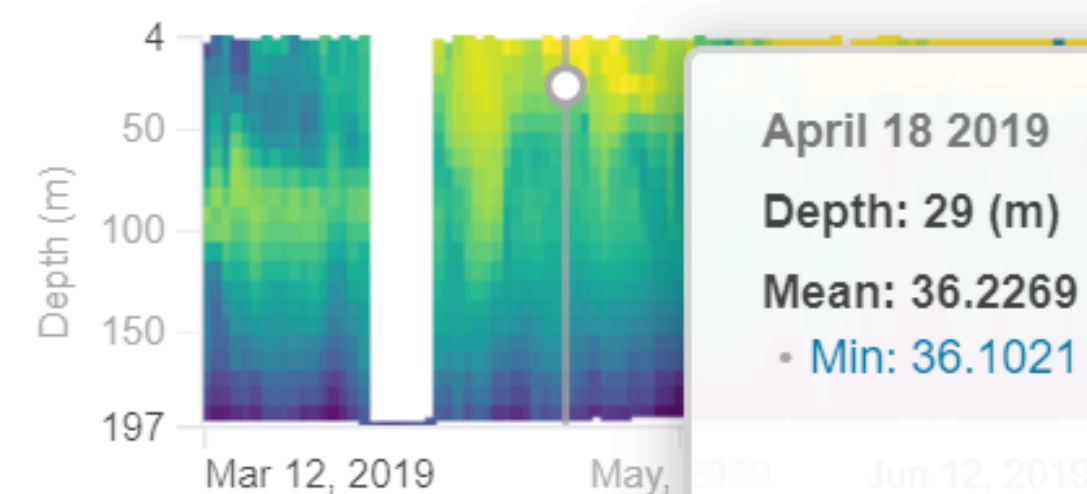
Deployments

Inventory

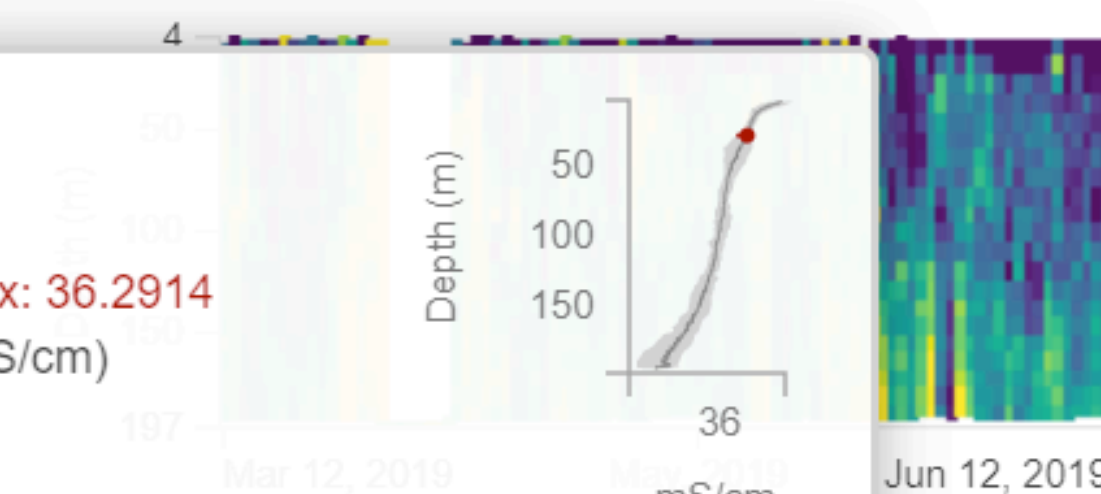
Chlorophyll a Mass Concentration



Conductivity



Eastward Sea Water Velocity

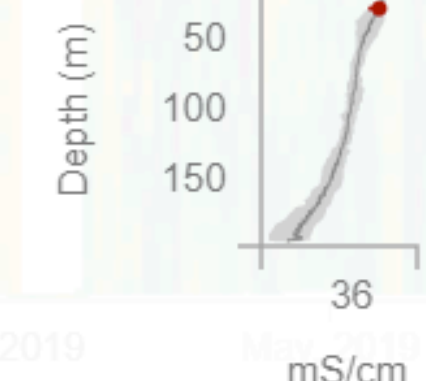


April 18 2019

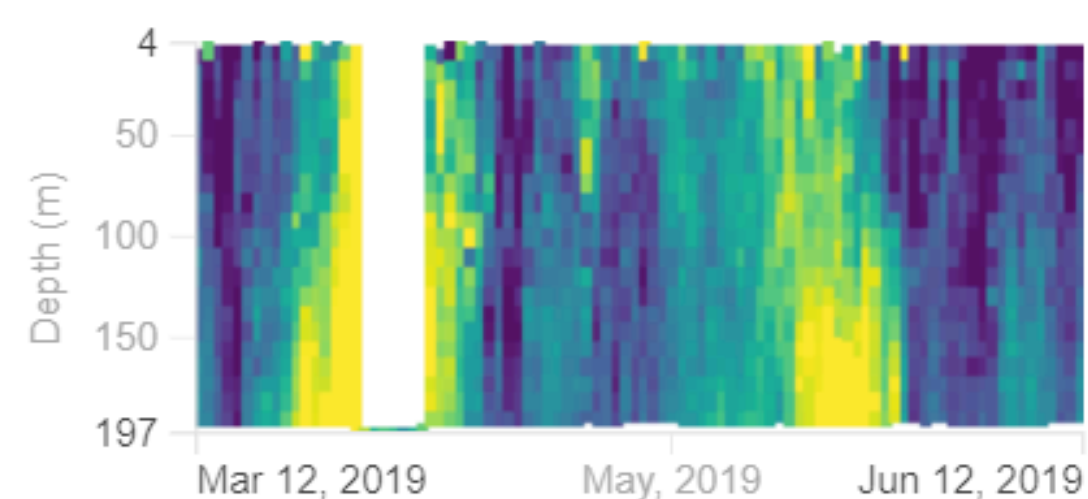
Depth: 29 (m)

Mean: 36.2269 • Max: 36.2914

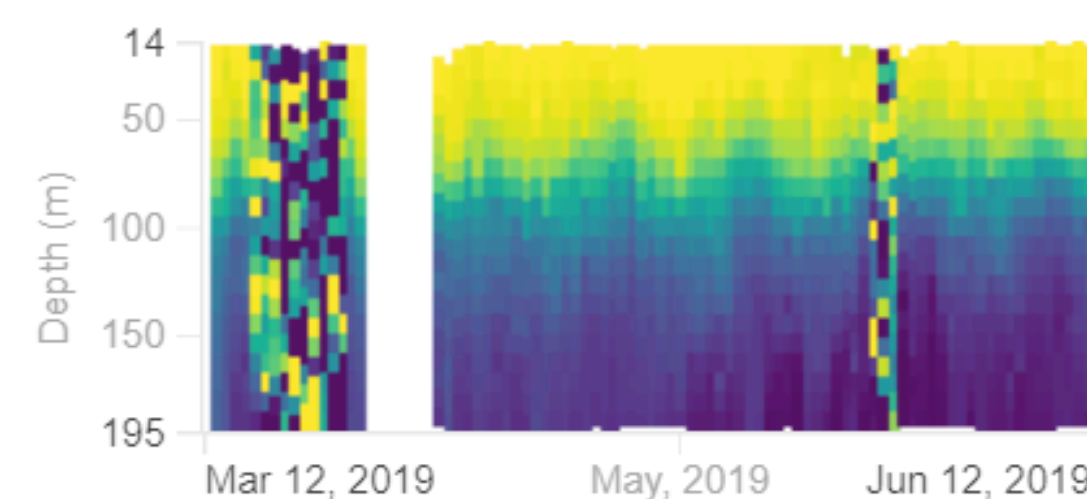
• Min: 36.1021 • (mS/cm)



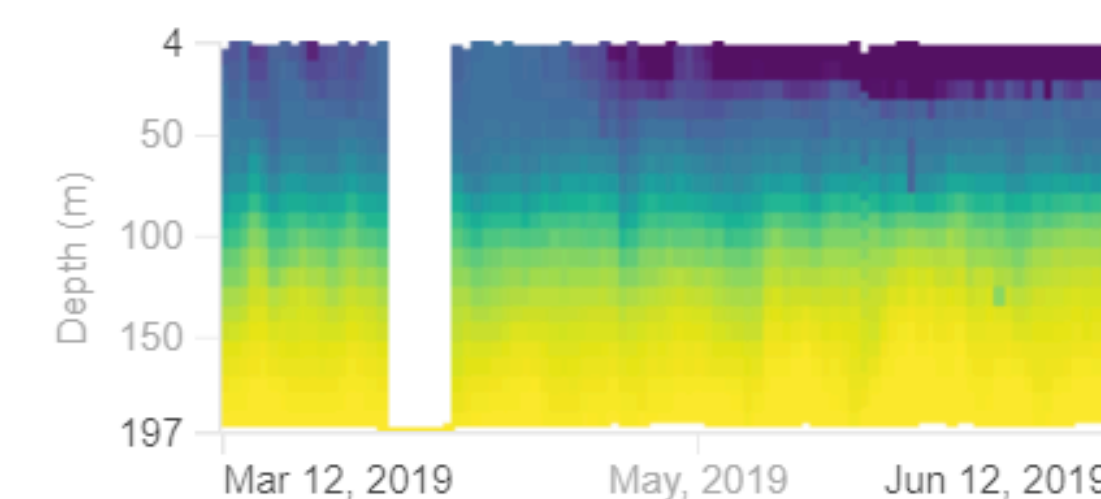
Northward Sea Water Velocity



pH



Salinity

Surface Piercing  
Profiler(SP002)

8

Surface Buoy(SBD17):  
CE.06ISSM.SBD17

9

Near Surface  
Instrument Frame(RID16):  
CE.06ISSM.RID16

11

Coastal Glider  
311(GL311)

11

Near Surface  
Instrument Frame(RID26):  
CE.07SHSM.RID26

6

Surface Buoy(SBD11):  
CE.04OSSM.SBD11

15

Coastal Glider  
381(GL381)

11



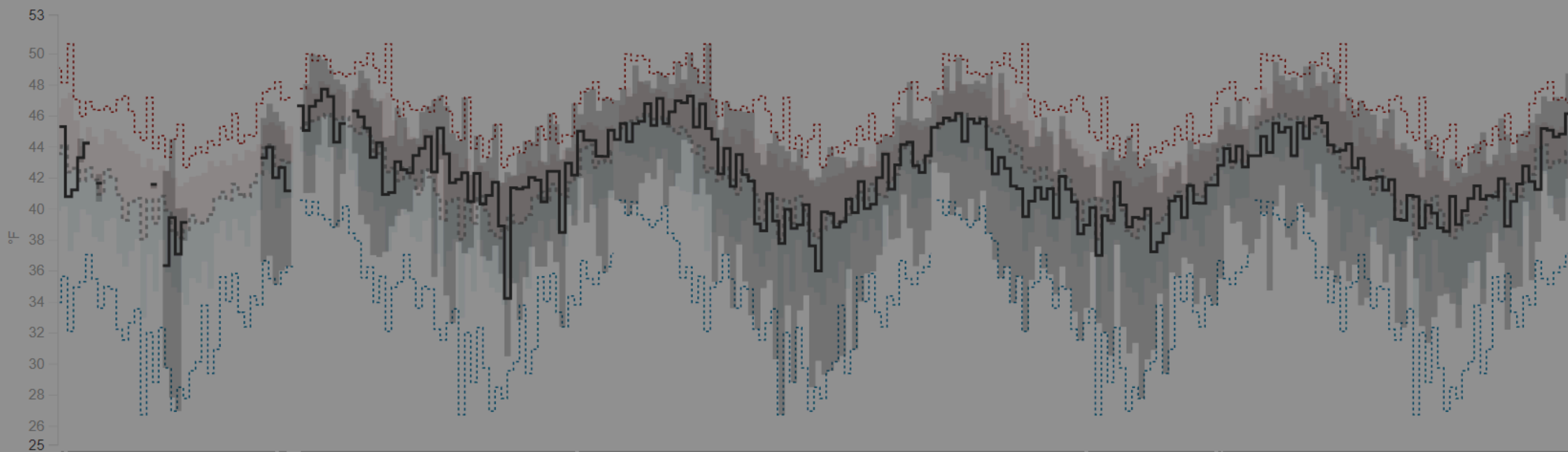
Nodes 18

Gliders 175

Global Southern Ocean &gt; Ape

## Air Temperature

Chart Time series ▾

☒ Autoscale

## Note



🕒 Nov 30, 2019 15:00 (GMT-0900) - Dec 2, 2019 03:00 (GMT-0900)

Deployment 4: Data gap for part of the day on 12/1/2019 through 12/2/2019. Data will be available for instruments that collect internally. [GS01SUMO-SBD11]

[Attributes ▾](#)





# Inventory

Filter by Part Type ▼

- 📍 Land
  - 📷 Hydrothermal Vent Fluid In-situ Chemistry - 1-20001
  - 📷 Spectrophotometer - 2-20001
- 📍 Sea
- 📍 Retired
- 📍 Trash Bin

Add Inventory

## Add Calibration

Calibration Date\*

09/27/2014



Approved\*

Draft ▼

### Calibrations

Calibration Name	Calibration Coefficient	Coefficient Notation Format	
CC_arr_hgo ▼	4.38978E-10	Scientific ▼	🗑 Remove
CC_arr_tac ▼	-1.88519E-7	Scientific ▼	🗑 Remove
CC_arr_hgo ▼	-0.000188232	Standard ▼	🗑 Remove
CC_arr_tac ▼	-2.80979E-9	Scientific ▼	🗑 Remove
CC_arr_hgo ▼	2.21477E-6	Scientific ▼	🗑 Remove
CC_arr_tac ▼	-0.000553586	Standard ▼	🗑 Remove
----- ▼		Standard ▼	🗑 Remove

+ Add Coefficient

Add Calibration



Inventory Search

←

→

↺

🏠

https://ooi-cgrdb-staging.who.net/search/inventory?f=.0.serial\_number&l=.0.contains&q=.0.optaa&f=.1.location\_\_name&l=.1.contains&q=.1.instrument&f=2.0.s

🔽

🔒

🔍

🌐

☰

🌐 WHOI Internal Web Site

✉️ [webmail](#)

📁 [Electrical](#)

📁 [Tech\\_Staff\\_Training](#)

📁 [OMS](#)

📁 [Python](#)

🖥️ [Welcome to WHOI En...](#)

📁 [weather](#)

📁 [Sustainability](#)

📁 [Real Estate](#)


📁 [TWR](#)

📁 [GliderOps](#)

📁 [3DAT Drives](#)

📁 [Shopping for Meg](#)

📁 [To Read](#)



[Inventory](#) [Builds](#) [Templates](#) [Users](#) [My Profile](#) [Sign Out](#) [⚙️](#)

Search Serial and Name

[🔍 Search](#)

[Inventory](#) [▼](#)

Search: [Inventory](#) ▼

Search Block

AND

⋮

Serial Number

▼

Contains

▼

optaa

—

AND

⋮

Location

▼

Contains

▼

instrument

—

+

ROW

Search Block 2

AND

⋮

Serial Number

▼

Contains

▼

dosta

—

NOT

⋮

UDF Value

▼

Contains

▼

4.

—

+

ROW

+

New Card

RESET

[🔍 Submit](#)

Search Results

10 items match your search!

[Download CSV](#)

🔍 Visible columns ▼

Serial Number	Name	Location	Date Modified	Unit Cost	Manufacturer	Firmware Version
<a href="#">CGINS-DOSTAD-00126</a>	DOSTA-D	Instrument Lab	03/19/2020 4:42 p.m.	0.00	Aanderaa	5.0.4
<a href="#">CGINS-DOSTAD-00128</a>	DOSTA-D	Instrument Lab	03/19/2020 4:42 p.m.	0.00	Aanderaa	
<a href="#">CGINS-DOSTAD-00495</a>	DOSTA-D	Retired	03/19/2020 4:42 p.m.	0.00	Aanderaa	
<a href="#">CGINS-DOSTAM-00400</a>	DOSTA-D	Retired	03/19/2020 4:42 p.m.	0.00	Aanderaa	
<a href="#">CGINS-OPTAAD-00123</a>	OPTAA-D	Instrument Lab	03/20/2020 11:31 a.m.	0.00	WET Labs	1.11
<a href="#">CGINS-OPTAAD-00130</a>	OPTAA-D	Instrument Lab	03/20/2020 11:31 a.m.	0.00	WET Labs	1.11
<a href="#">CGINS-OPTAAD-00151</a>	OPTAA-D	Instrument Lab	03/20/2020 11:31 a.m.	0.00	WET Labs	1.11
<a href="#">CGINS-OPTAAD-00206</a>	OPTAA-D	Instrument Lab	03/20/2020 11:31 a.m.	0.00	WET Labs	1.11