2022

OOI Biogeochemical Sensor Data: Best Practices and User Guide. Version 1.0.0.

Hilary I. Palevsky

Sophie Clayton
Old Dominion University, sclayton@odu.edu

Dariia Atamanchuk

Roman Battisti

Jennifer Batryn

See next page for additional authors

Original Publication Citation

Palevsky, H., Clayton, S., Atamanchuk, D., Battisti, R., Batryn, J., Bourbonnais, A., Briggs, E. M., Carvalho, F., Chase, A. P., Eveleth, R., Fatland, R., Fogaren, K. E., Fram, J. P., Hartman, S. E., Le Bras, I., C.M. Manning, C., Needoba, J. A., Neely, M. B., Oliver, H., . . . Wingard, C. (2022). *OOI Biogeochemical Sensor Data: Best Practices and User Guide, Version 1.0.0*. Ocean Observatories Initiative, Biogeochemical Sensor Data Working Group, 134 pp. http://dx.doi.org/10.25607/OBP-1865

Authors

Hilary I. Palevsky, Sophie Clayton, Dariia Atamanchuk, Roman Battisti, Jennifer Batryn, Annie Bourbonnais, Ellen M. Briggs, Filipa Carvalho, Alison P. Chase, Rachel Eveleth, Rob Fatland, Kristen E. Fogaren, Jonathan Peter Fram, Susan E. Hartman, Isabela Le Bras, Cara C.M. Manning, Joseph A. Needoba, Merrie Beth Neely, Hilde Oliver, Andrew C. Reed, Jennie E. Rheuban, Christina Schallenberg, Michael F. Vardaro, Ian Walsh, and Christopher Wingard

OOI Biogeochemical Sensor Data: Best Practices & User Guide

Version Number: 1.0.0
December 2022

Hilary I. Palevsky[1]*, Sophie Clayton[2]*, Dariia Atamanchuk[3], Roman Battisti[4], Jennifer Batryn[5], Annie Bourbonnais[6], Ellen M. Briggs[7], Filipa Carvalho[8], Alison P. Chase[9], Rachel Eveleth[10], Rob Fatland[11], Kristen E. Fogaren[1], Jonathan Peter Fram[12], Susan E. Hartman[8], Isabela Le Bras[5], Cara C. M. Manning[13], Joseph A. Needoba[14], Merrie Beth Neely[15], Hilde Oliver[5], Andrew C. Reed[5], Jennie E. Rheuban[5], Christina Schallenberg[16], Michael F. Vardaro[17], Ian Walsh, Christopher Wingard[12]

[1] Boston College, Department of Earth and Environmental Sciences, [2] Old Dominion University, [3] Dalhousie University, [4] Cooperative Institute for Climate, Ocean and Ecosystem Studies/Pacific Marine Environmental Laboratory, [5] Woods Hole Oceanographic Institution, [6] University of South Carolina, School of the Earth, Ocean and Environment, [7]University of Hawaii at Manoa, [8] National Oceanography Centre, UK, [9] Applied Physics Laboratory, University of Washington, [10] Oberlin College, Department of Geosciences, [11] University of Washington, eScience Institute, [12]Oregon State University, [13] University of Connecticut, Department of Marine Sciences [14] Oregon Health & Science University, OHSU-PSU School of Public Health [15] Global Science and Technology, Inc., [16] Commonwealth Scientific and Industrial Research Organisation, Environment, Hobart, Australia [17] University of Washington, School of Oceanography

* co-lead, corresponding authors: Hilary I. Palevsky, palevsky@bc.edu; Sophie Clayton, sclayton@odu.edu

Author contributions: HIP and SC led all activities of the OOI Biogeochemical Sensor Data Working Group, which collaboratively envisioned and wrote this document. KF, JPF, SH, CCMM, and IL led Chapter 2 (Oxygen). JB, AB, JAN, HO, and IW led Chapter 3 (Nitrate). DA, RB, EMB, RE, ACR, and JER led Chapter 4 (Carbonate chemistry). FC, APC, RF, MBN, CS, and CW led Chapter 5 (Bio-optics). MV made significant contributions to information about the OOI. All authors after HIP and SC are listed in alphabetical order.

Beta Testers: This group tested and reviewed an early draft of this document in May 2022 and participated in the June 2022 workshop, providing feedback which significantly improved the content and presentation of the document: Kohen Bauer, Baoshan Chen, Jose Cuevas, Susana Flecha, Micah Horwith, Melissa Meléndez, Tyler Menz, Al Plueddemann, Sara Rivero-Calle, Nick Roden, Tobias Steinhoff, Pablo Nicolás Trucco-Pignata, Mike Vardaro, and Meg Yoder.

Acknowledgments: The authors thank Heather Benway, Mai Mahegan, and Mary Zawoysky of the Ocean Carbon and Biogeochemistry Project Office for their invaluable support in organizing the OOI Biogeochemical Data Working Group, July 2021 online workshop, and June 2022 in person workshop that supported the development of this document. Funding for this effort was provided by the Ocean Observatories Initiative Facilities Board, the Ocean Observatories Initiative, and the National Science Foundation (Awards 2033988, 2034002, and 2033919).

Table of Contents:

Tables and Figures

Tables

Figures

Chapter 0: Quick Start Guide

The OOI Biogeochemical Sensor Data Best Practices and User Guide is intended to provide current and prospective users of data generated by biogeochemical sensors deployed on the Ocean Observatories Initiative (OOI) arrays with the information and guidance needed for them to ensure that the data is science-ready. This guide is aimed at researchers with an interest or some experience in ocean biogeochemical processes. We expect that users of this guide will have some background in oceanography, however we do not assume any prior experience working with biogeochemical sensors or their data.

Scope of the guide

While initially envisioned as a “cookbook” for end users seeking to work with OOI biogeochemical (BGC) sensor data, our Working Group and Beta Testers realized that the processing required to meet the specific needs of all end users across a wide range of potential scientific applications and combinations of OOI BGC data from different sensors and platforms couldn’t be synthesized into a single “recipe”. We therefore provide here the background information and principles needed for the end user to successfully identify and understand all the available “ingredients” (data), the types of “cooking” (end user processing) that are recommended to prepare them, and a few sample “recipes” (worked examples) to support end users in developing their own “recipes” consistent with the best practices presented here.

This is not intended to be an exhaustive guide to each of these sensors, but rather a synthesis of the key information to support OOI BGC sensor data users in preparing science-ready data products. In instances when more in-depth information might be helpful, references and links have been provided both within each chapter and in the Appendix.

Structure of the guide

This guide is split into 5 chapters, with an Introduction (Chapter 1) providing information relevant to all OOI BGC sensor data users and four subsequent chapters covering four groups of BGC variables and associated sensors:

- Chapter 2: Dissolved Oxygen
- Chapter 3: Nitrate
- Chapter 4: Carbonate chemistry
- Chapter 5: Bio-optical measurements

Chapter 1, intended as a starting point for all users of this guide, includes:

- An **overview of the OOI program**, its arrays, and the biogeochemical data they collect, including Table 1.1 summarizing the biogeochemical sensors deployed across all OOI arrays
- A summary of **how OOI processes data internally** before providing it to the end user, including human-in-the-loop (HITL) and automated quality assurance/quality control

(QA/QC) procedures, with Table 1.2 summarizing automated QA/QC tests currently either implemented or under review.
- An overview of **how to access OOI data**, emphasizing pointers to OOI-created resources
- An overview of **additional QA/QC procedures that we recommend be completed by the end user** of OOI BGC sensor data to prepare science-ready datasets for analysis, including a flowchart (Figure 1.1) highlighting key processing steps for all OOI BGC sensor types

The following four chapters each provide specific information on a subset of BGC sensors, with each chapter arranged in a parallel format that includes:

- An **introduction to the sensors** used by OOI to measure this BGC variable, including a table with key information about the sensors covered in the chapter and the OOI platforms on which they are deployed
- **OOI standard practices for deployment and calibration** of the BGC sensor(s) covered in the chapter
- A summary of the **internal-to-OOI data processing** for this BGC variable, with the processing from raw sensor output to end-user ready variables synthesized in a flowchart for each sensor type
- **Common data quality issues** that end users should be aware of when working with data from this type of BGC sensor(s)
- **Recommendations for end user data processing** to prepare analysis-ready data products from this type of BGC sensor(s), summarized in a sensor-specific end user data processing flowchart for each chapter
- A **worked example**, illustrating how to apply the recommended end user data processing described in the chapter & flowchart to an example OOI dataset. Each example includes **pseudo-code** to support users in developing their own data analysis pipeline suited to their specific application in the programming language of their choice. The code used by this guide's authors to prepare the examples is provided as a supplementary resource, but was not tested for application to other example datasets.

How to use this guide

Each of the sensor chapters are designed to build on the general information given in Chapter 1, but to be independent from each other. We recommend that all users familiarize themselves with the content in Chapter 1 before proceeding to the sensor-specific chapters that follow. For example, a user who is primarily only interested in working with bio-optical sensor data should read Chapter 1 to find the general information related to the configuration of the OOI arrays and data access, and then proceed to Chapter 5 to find the more specific information related to processing data from bio-optical sensors.

The flowcharts in each chapter summarizing both the internal-to-OOI data processing and recommended end user data processing are intended as both a starting point and as a reference to return to while the user completes their own data processing:

Internal-to-OOI Data Processing Flowcharts:

- Dissolved Oxygen
- Nitrate
- Carbonate chemistry
- Bio-optical measurements

Recommended End-User Data Processing Flowcharts:

- Overview for all OOI BGC sensors
- Dissolved Oxygen
- Nitrate
- Carbonate chemistry
- Bio-optical measurements

The worked examples and the accompanying pseudocode at the end of each chapter illustrate the key steps shown in the recommended end-user data processing flowcharts, but are not an exhaustive template for the user to follow in addressing all data QA/QC and calibration issues discussed in each chapter. All code and calculations used in processing the worked examples are provided as a supplementary resource (see Section A5), though these are intended solely as a reference. We emphasize that it is beyond the scope of this guide to provide full code pipelines for data processing. This is in part because we recognise that there is no single coding language of choice within the oceanographic community, and in part due to the magnitude of the task required to develop and maintain a codebase that would cover all cases of required end user data processing across all the OOI arrays, platforms, and sensor types discussed here. Users of this guide are recommended to develop their own data processing code suited to their specific application of OOI BGC sensor data, and to share their tools with other users through the OOI Discourse channels.

We recognize that working with OOI BGC data can, at first glance, be a daunting task. The number of different methods for accessing and downloading data, integrating multiple sources of data, the different sensors with their individual characteristics, and the sheer quantity of data can require a sizable investment of effort. The intention of this document is to lower the barrier to entry and support users in accessing and making use of the rich BGC datasets collected and provided by the OOI.

Chapter 1: Introduction

1.1 Scope and Goals

The Ocean Observatories Initiative (OOI) deploys sensors that measure key biogeochemical properties on both moored and mobile autonomous platforms across ocean observing arrays in the Atlantic, Pacific, and Southern Oceans (Smith et al., 2018; Trowbridge et al., 2019). These sensors provide great potential to support the oceanographic community in studying a wide range of important and interdisciplinary questions. However, OOI biogeochemical sensor data have thus far been underutilized by the oceanographic community. Studies making primary use of OOI's suite of biogeochemical sensors make up less than 15% of published studies to date known to have used OOI data. One reason for this underutilization of OOI biogeochemical sensor data is that research quantifying biogeochemical fluxes and addressing many questions of scientific interest (e.g., rates of air-sea CO_2 flux, productivity, and export; comparison across sites; monitoring of long-term changes) require effective calibration and validation, including post-deployment human-in-the-loop (HITL) processing outside of the scope of what OOI is charged to implement.

This OOI Biogeochemical (BGC) Sensor Data Best Practices and User Guide is the result of a grass-roots community effort to broaden the use of OOI biogeochemical sensor data and increase community capacity to produce analysis-ready data products. This effort brought together an international group of 39 ocean observing experts, across all career stages, from 19 institutions and 5 countries, each of whom brings expertise on biogeochemical sensors, data analysis and ocean observing infrastructure as well as research expertise in ocean biogeochemistry. The initial OOI Biogeochemical Sensor Data (OOI BGC) Working Group was formed in July 2021 through an open application process. A three-day virtual meeting in July 2021 launched the activities of the Working Group, with consensus-building activities to develop the scope and structure of the Best Practices and User Guide. From July 2021 to June 2022, the Working Group drafted a beta version of this Best Practices and User Guide that went through two rounds of internal review within the Working Group. A draft version of this document was Beta Tested by 14 current and prospective OOI BGC data users, who joined the Working Group members for a 3-day workshop in June 2022 to provide feedback that has since been used in revising and finalizing the document.

The best practices outlined here encompass four groups of biogeochemical variables and associated sensors which have been widely deployed across the OOI arrays and are also commonly used in other ocean observing and autonomous biogeochemical sensor programs.

1) ***Dissolved oxygen***, measured by two sensor types: Aanderaa (now a subsidiary of XYLEM) optodes, used on most platforms, and Sea-Bird electrochemical sensors, used on some profiler moorings

2) ***Nitrate***, measured by two similar instruments using the same optical principles: the In Situ Ultraviolet Spectrophotometer (ISUS) and the Satlantic (now part of Sea-Bird) Submersible Ultraviolet Nitrate Analyzer (SUNA),
3) ***Carbonate chemistry (pH and pCO_2)***, measured by Sunburst SAMI sensors using similar colorimetric reagent methods for both variables, and a Pro-Oceanus sensor deployed on surface buoys to measure air and surface seawater pCO_2 using a nondispersive infrared detection method,
4) ***Chlorophyll fluorescence, optical backscatter and fluorescent dissolved organic matter (FDOM)***, measured optically by Sea-Bird's WETLab fluorometer series.

This document is intended to support current and potential users of OOI biogeochemical data from across the oceanographic research community by synthesizing existing knowledge about best practices for calibration and validation of the types of biogeochemical sensor data collected by OOI and context specific to the OOI program. We do not assume any prior knowledge of OOI systems or procedures, nor does the user need detailed knowledge of the operation of the various sensors covered in the following chapters in order to implement the procedures described therein. We have endeavored to ensure that all information is up to date at the time of writing (completed December 2022) and to point to internal-to-OOI resources that will continue to be updated to reflect future developments in the OOI program.

In the remainder of Chapter 1, we provide information intended as a starting point and overview for end users working with any type of OOI BGC data. The following 4 chapters provide information specific to each of the 4 groups of variables covered here (Dissolved oxygen, Nitrate, Carbonate chemistry, and Bio-optical measurements). For a summary of the overall Best Practices & User Guide document structure and links to key resources, see the Chapter 0: Quick Start Guide.

1.2 Overview of OOI program and OOI BGC data

1.2.1 Summary of OOI Arrays

The Ocean Observatories Initiative (OOI) is an NSF-funded major research facility that includes five arrays currently in operation as well as two arrays in the Southern Hemisphere that were decommissioned in 2017 (with the Southern Ocean Array collecting data until January 2020, with support from the UK's Natural Environment Research Council). All data, including that collected from the now-decommissioned arrays, is publicly available through the OOI cyberinfrastructure (see Data Access section below). Detailed descriptions of the OOI program's development, motivation, and design are available elsewhere (see especially Smith et al., 2018; Trowbridge et al., 2019, and the current OOI Science Plan: OOIFB, 2021). Here we provide a summary of the OOI arrays, with a focus on platforms and instruments collecting the biogeochemical data types addressed in this document.

The OOI includes three array types, each with its own configuration of platforms and instruments:

- **Global Arrays**, including two Northern Hemisphere arrays with ongoing measurements, in the subarctic Northeast Pacific (Global Station Papa Array) and the subpolar North Atlantic (Global Irminger Sea Array), as well as two now-decommissioned Southern Hemisphere arrays (the Global Southern Ocean Array and the Global Argentine Basin Array).
- **Coastal Arrays**, including the Coastal Endurance Array, permanently deployed off the coast of Washington and Oregon, and the Pioneer Array, deployed since OOI's inception at the New England shelf-break front (Pioneer - NES) but designed to be relocatable and currently slated to be redesigned and relocated to the southern Mid-Atlantic Bight (Pioneer - MAB) in 2024.
- **Regional Cabled Array (RCA)**, instrumenting the Juan de Fuca tectonic plate in the Northeast Pacific, uses two electro-optical cables to stream real-time data from water column, seafloor, and sub-seafloor sensors. The coastal component of the RCA (including Slope Base, Southern Hydrate Ridge, and the Oregon Offshore and Shelf sites connected to the Endurance Array) includes profiling and seafloor platforms equipped with biogeochemical sensors within the scope discussed here. The network also includes seafloor sensors at Axial Seamount and Hydrate Ridge focused on understanding connections between volcanism, seismicity, biogeochemical fluxes, and life in extreme environments which are outside the scope of this document.

Schematic drawings of each of the OOI arrays, with accompanying captions identifying the locations within each array of all biogeochemical sensors covered in this Best Practices & User Guide are provided in the Appendix (Section A7, Figures A.1-A.6).

The OOI program encompasses operations coordinated across multiple institutions to deploy and maintain these arrays and to provide access to the collected data. As of October 2018, Woods Hole Oceanographic Institution (WHOI) hosts the OOI Program Management Office, which coordinates operations across the entire program. The Global Arrays and the Pioneer Array are operated by the WHOI; the Endurance Array is operated by Oregon State University (OSU); the Regional Cabled Array is operated by the University of Washington (UW); and OSU is responsible for the Cyberinfrastructure. While the full operational structure and history of OOI are beyond the scope of this document and are not intended to be information needed by the end data user, it can be helpful to be aware of differences in institutional operators especially when comparing data across multiple arrays operated by different institutions.

The OOI arrays include three main types of platforms equipped with biogeochemical sensors:

- **Moorings with sensors at fixed depths** - these include both surface moorings, which have instrumented surface buoys, and sub-surface moorings; in both cases sensors are deployed at multiple fixed depths through the water column
- **Profiler moorings** - these moorings collect regular vertical profiles, using a combination of profiler technologies customized for the OOI program: sub-surface Wire-Following Profilers based on the McLane Moored Profiler deployed on the Coastal and Global Arrays and as Cabled Deep Profilers at the RCA; winched Cabled Shallow Profiler

Moorings at the RCA; and Coastal Surface-Piercing Profilers on the Endurance Array that sample across the air-sea interface as well as through the water column

- **Gliders** - Slocum gliders from Teledyne Webb Research transit the arrays, collecting profile data through the water column. Glider missions include collecting data in portions of the water column not sampled with the moorings (i.e., collecting profiles that include the upper water column above sub-surface profiler moorings) and array transits that pass by the moorings in a regular pattern.

Table 1.1 summarizes the biogeochemical sensors deployed across each of these platform types across each of the OOI arrays. Each sensor type is discussed in detail in the chapters that follow in the remainder of this document. Additional details about the configuration of the platforms within each array can be found through the OOI website and the current OOI Science Plan (OOIFB, 2021) and detailed information on the observation and sampling approaches for all data collected by mobile and moored sensors are provided here.

Biogeochemical sensor distribution within each of the array platforms is aimed to optimize for measurements of maximal scientific interest; for instance, fixed depth biogeochemical sensors on Global Array moorings are all found within the top 130 m of the water column, and most densely concentrated on the surface buoy (1 m) and near-surface instrument frame (12 m). The four Global arrays have similar configurations but differ in the exact number and configuration of sensors: for instance, the number of profiler sensors is halved at the Irminger Sea site because separate deep and shallow profilers are not needed due to the shallower water depth, and the number of fixed sensors is significantly reduced at Station Papa because OOI does not deploy a surface mooring due to the presence of a NOAA surface mooring at that site.

During deployments, instruments are set to both internally record and to transmit data. OOI provides data in near-real time, streamed from the Regional Cabled Array, and telemetered from surface moorings and gliders on the Global and Coastal Arrays. Telemetered data provided in near real time is generally decimated, due to the limitations of transmitting large amounts of data in challenging conditions and with limited battery power, particularly for glider and sub-surface mooring data being transmitted by the gliders. Recovered data, added after deployments have been completed and all assets have been recovered, are more complete, but at times may be missing data that was previously telemetered (for instance, if a sensor or platform experienced a major failure or was lost part way through the deployment).

Table 1.1. Summary of biogeochemical sensors included in the scope of this document and numbers deployed across all OOI arrays. Deployed sensors of each type are summarized among all fixed-depth mooring deployments, profiler moorings, and gliders at each site. Since the number of active gliders has varied across deployments, here we assume 2 gliders deployed at a time per array. This summary excludes autonomous underwater vehicles deployed at Pioneer during turn-around cruises. Benthic experiment packages at the Endurance Array are included in the "fixed" platform count.

Array	Platform	**Oxygen** (chapter 2)		**Nitrate** (chapter 3)	**Carbonate chemistry** (chapter 4)			**Bio-optics** (chapter 5)
		Aanderaa	*Sea-Bird*	*SUNA/ISUS*	*pH: SAMI*	*pCO_2: SAMI*	*pCO_2:Pro-Oceanus*	*WetLabs fluorometers*
Global Irminger	Fixed	7	0	2	4	4	1	7
	Profiler	1	0	0	0	0	0	1
	Gliders	2	0	1	0	0	0	2
Global Papa	Fixed	2	0	0	2	0	0	2
	Profiler	2	0	0	0	0	0	2
	Gliders	2	0	1	0	0	0	2
Global Argentine (suspended Jan. 2018)	Fixed	2	0	2	3	3	0	5
	Profiler	2	0	0	0	0	0	2
	Gliders	2	0	1	0	0	0	2
Global Southern (suspended Jan. 2020)	Fixed	4	0	2	4	3	1	5
	Profiler	2	0	0	0	0	0	2
	Gliders	2	0	0	0	0	0	2
Coastal Pioneer (NES)	Fixed	6	0	3	6	3	3	3
	Profiler	0	7	0	0	0	0	7
	Gliders	2	0	2	0	0	0	2
Coastal Endurance	Fixed	13	0	6	13	9	4	8
	Profiler	5	2	5	1	1	0	8
	Gliders	2	0	0	0	0	0	2
Cabled	Fixed	4	0	0	2	0	0	2
	Profiler	2	2	2	2	2	0	6
OOI program total		64	11	27	37	25	9	72

1.2.2 Charge of the OOI program

OOI's charge is to deploy and maintain the arrays and to ensure the delivery of the data, but intentionally separates data analysis and interpretation from the operation of the program. The OOI data delivery is supported by a wide range of activities, and the purview of the OOI includes:

- Designing, building, and deploying all platforms within the arrays
- Regular turn-around cruises (1 or 2 times per year) to service each array, including recovery of previously-deployed moorings and gliders, deployment of new moorings and gliders that will collect data over the next deployment period, and collection of data and discrete samples from CTD casts co-located with OOI assets
- Maintenance of all sensors deployed on the array platforms, including ensuring that instruments are calibrated and tested in accordance with vendor-supplied instructions
- Making publicly available to the community all data collected by all deployed sensors, data collected during shipboard deployment/recovery cruises, and documentation of metadata related to the platforms, instruments and operation of the arrays

This final point, making the data publicly available, is the key product of the OOI program. It is important to note, however, that OOI's purview does not include curating the data to produce final analysis-ready data products. However, OOI does undertake preliminary data processing steps for all sensors, with raw sensor data undergoing a standard set of processing and QA/QC procedures prior to being served to end-users via the OOI cyberinfrastructure. These steps are summarized in Section 1.3 below.

Particularly for the biogeochemical sensors and data types covered in this document, development and implementation of comprehensive QA/QC procedures needed to produce analysis-ready data products requires significant human-in-the-loop (HITL) effort. A 2012 report from the COL-NASA Data QA/QC Workshop, convened as OOI prepared for initial operations, estimated that 20 people's full-time effort would be necessary to provide this comprehensive QA/QC for the entire OOI program; however, funding constraints prevented inclusion of this level of effort internal to the OOI program. Recognizing the importance of this work, OOI has recently increased the amount of full-time equivalent effort from 4 to 5.85 people. The efforts of the OOI BGC Working Group and the information provided in this document are therefore aimed at supporting end users in understanding the data processing and QA/QC that is completed internally by OOI, as well as implementing additional QA/QC steps to produce their own analysis-ready data products for scientific applications.

1.3 Internal to OOI Data Processing and QA/QC procedures

Although it is beyond the OOI scope to produce their own analysis-ready data products for the BGC sensors, OOI does undertake initial data processing as well as some basic QA/QC. In this section we describe the data processing and QA/QC steps that are uniform across all sensor types, with sensor-specific details contained within each of the individual data type chapters.

1.3.1 Overview of processing steps common to all sensor types

Here we provide an overview of the approach that OOI applies to processing all data types, with detailed biogeochemical sensor-specific information on these data processing steps covered within each of the following chapters. **Flowcharts summarizing the internal-to-OOI data processing to produce user-ready output from each BGC sensor type are included in each chapter.**

Data are received by OOI either streaming via the cabled network in real time, or in batches telemetered to shore via an Iridium or broadband (e.g. cellular) modem. Additional data may be downloaded from the instruments after the platforms and instruments are recovered during normal maintenance cruises. OOI parses all data received into tables in an internal-only Cassandra database, with each table in the database corresponding to a specific instance of a site, node, sensor, data delivery method and stream name. These individual tables constitute the unique data sets representing the more than 800 instruments served by OOI.

The received data are parsed and stored into the tables essentially unchanged, although in some cases date/time strings are converted into time in seconds since 1900-01-01, or bitmaps are decoded to access error flags. Additionally, copies of the unparsed data are stored and archived by OOI in the Raw Data Archive. Within each parsed data set, individual parameters (or variables) may be identified by OOI as a Science Data Product. Each of the Science Data Products may have one or more processing levels associated with it. OOI classifies those processing levels as L0, L1, or L2.

- **Level 0**, or L0 Data Products are unprocessed parameters within a data set reported by the instrument in native instrument/sensor units and resolution. No QC tests are applied to L0 data products.

  Example: raw chlorophyll-a fluorescence data reported by the Sea-Bird Electronics ECO-Triplet (FLORT) sensor in counts (CHLAFLO-L0).

- **Level 1**, or L1 Data Products may be derived from L0 data products. They may also be reported directly by the instrument in the data set. They represent parameters that have been processed using vendor-provided calibration values and/or algorithms, or values derived from pre-deployment procedures, and that are in scientific units. L1 data products have usually also undergone some QA/QC procedures, utilizing simple automated techniques and human inspection. More information on QA/QC procedures routinely performed by OOI are outlined in detail below. The processing steps that were performed to transform an L0 Data Product to an L1 Data Product are recorded in the associated metadata of the L1 data product if OOI performed the conversion.

  Example: L0 count data from fluorometers are converted to the L1 fluorometric Chlorophyll-a concentration (CHLAFLO-L1) using manufacturer's conversion factors.

- **Level 2**, or L2 Data Products are derived quantities created via an algorithm that draws on multiple L1 Data Products. L2 data products may be based on combined L1 data

products from the same instrument or a combination of L1 and/or L2 data products from separate instruments.

Example: L1 conductivity (CONDWAT-L1), temperature (TEMPWAT-L1), and pressure (PRESWAT-L1) data products are used to calculate the L2 practical salinity data product for the CTDs used by OOI. For multiple of the BGC sensors, L1 temperature (TEMPWAT-L1) and L2 practical salinity (PRACSAL-L2) data products are then used in conjunction with BGC sensor L0 and/or L1 output to produce an L2 product.

It is important to understand how OOI uses the term Science Data Product, and their associated processing levels, as their usage differs from common expectations. Rather than considering an entire data set as an L0, L1, or L2 data product with distinct processing steps and levels between each data set, individual parameters within the data set are identified as data products. It is possible for a data set to contain multiple parameters identified as Science Data Products, with potentially multiple processing levels for those parameters included, in addition to other parameters that are not classified as data products. For example, data sets from Aanderaa dissolved oxygen sensors deployed on surface moorings contain 21 parameters, only 4 of which are identified as Science Data Products (2 L0's, 1 L1, and 1 L2). **The user-ready output for each sensor type is described in each chapter and highlighted in the flowchart summarizing the internal-to-OOI data processing for that sensor.**

1.3.2 Current OOI QC procedures for BGC data streams

Most L1, and some of the L2, data products undergo a series of QC procedures. Here we describe the QC procedures currently in use by OOI, with a particular focus on data produced by BGC sensors. It is important to note that OOI is in the process of converting the existing QC tests applied to ones that are designed to meet the U.S. Integrated Ocean Observing System (IOOS) Quality Assurance of Real Time Ocean Data (QARTOD) quality control standards (Toll, 2012), so the information provided here is current to the time of writing, but will likely change over time.

Here we list some of the QARTOD standards for which this section and the following 'Human in the Loop (HITL) Data Annotations' section will describe OOI's implementation:

- Every real-time observation must be accompanied by a quality descriptor
- All observations should be subject to automated real-time quality tests
- Quality flags and test descriptions must be included in the metadata
- Observers should describe methods/calibration in the metadata
- Observers should quantify level of calibration accuracy and expected error

The OOI data team is responsible for ensuring that the collected data and metadata delivered by OOI meets community data quality standards. In order to achieve these standards, OOI data are evaluated with both manual and automated QC tests. Manual tests are performed by data evaluators, and include "Quick Look" tests where evaluators perform a first pass evaluation of the data with automated tools, and "Deep Dives" where evaluators more closely inspect data flagged as failed or suspect and may consult further with subject matter experts. Data that are flagged as failed or suspect during manual inspection are annotated as such (data annotations

are described in more detail below). In addition to the manual tests, the data are also processed through six automated QC algorithms. These tests have been in use for several years and most L1 and some L2 data products will have quality flags associated with them based on these tests. As OOI transitions towards QARTOD style tests and flags, these existing tests will be deprecated. Table 1.2, below, describes these existing tests and their QARTOD equivalents.

Table 1.2. OOI automated QA/QC tests currently either implemented or under review. Links have been included to the internal OOI documentation for each of the listed OOI tests. Note that not all of these tests are implemented for all data at the time of writing, with tests for BGC data products currently in development.

OOI test	OOI Description	QARTOD Equivalent	Level of OOI Implementation
Global Range Test	Data are flagged unless they fall within valid world ocean ranges or instrument limits (whichever is more restrictive)	Sensor range in the Gross Range test	Currently operational.
Local Range Test (similar to Climatology Test)	Data are flagged unless they fall within locally valid site-specific or depth ranges. Interpolates thresholds between depth and season intervals	User range in the Gross Range test	OOI local ranges are still being established and will be based on data collected to date as test limits are developed.
Spike Test	Deviation from mean compared to neighboring points	Spike	Not operational.
Trend Test	Data are flagged as having a trend if the standard deviation of the residuals to a polynomial curve < original data, multiplied by some factor. Designed to test for sensor drift.	N/A	Not operational.
Stuck Value Test	If 2 neighboring values differ by less than the resolution of the sensor for more than N repetitions, data are flagged	Flat-Line	Currently operational.
Gradient Test	If d(data)/d(t) between two points is greater than a set threshold, all following points fail until one falls within an absolute limit (TOLDAT). First data point is assumed good unless a “good” starting data (STARTDAT) point is defined.	Gradient	Not operational.

Data products processed through the automated QC procedures listed above are assigned data flags to indicate data quality. For every data product tested, the parameter will have two additional parameters added sharing the same parameter name with _qc_results and _qc_executed added containing the results of each test and which tests were performed, respectively. These quality flags are presented as integer numbers that represent a bitmap of the individual test results and a bitmap of which tests were performed. These quality flags, however, fail to meet one of the QARTOD standards described above, where “Quality flags and test descriptions must be included in the metadata.” This failure is one of the reasons OOI is transitioning to QARTOD based QC tests.

Where implemented, data products that have been tested using QARTOD style tests will have two additional parameters added with _qartod_results and _qartod_executed added to the parameter name. The _qartod_results variables are integer values between 1 and 9 that follow the QARTOD flagging conventions, where:

- **1 = Pass:** Data have passed critical real-time QC tests and are deemed adequate for use as preliminary data.
- **2 = Not Evaluated:** Data have not been QC-tested, or the information on quality is not available.
- **3 = Suspect or of High Interest:** Data are considered to be either suspect or of high interest to operators and users. They are flagged as “suspect” to draw further attention to them by operators.
- **4 = Fail:** Data are considered to have failed one or more critical real-time QC checks. If they are disseminated at all, it should be readily apparent that they are not of acceptable quality.
- **9 = Missing Data:** Data are missing; used as a placeholder.

The _qartod_executed variables are strings with each value in the string representing the results of each test performed. The metadata associated with each of these variables includes descriptions of the tests, the results, and the order of the results. At this time, OOI is focused on implementing the Gross Range and Climatology tests from QARTOD for all sensors. As of December 2022, these tests have been implemented for CTDs, pH, and pCO_2 sensors, with tests for dissolved oxygen and fluorometric chlorophyll sensors in process. The test limits used for the older QC tests and the newer QARTOD tests can be found in this GitHub repository.

1.3.3 Human-in-the-Loop (HITL) Data Annotations

Automated QC flags are supplemented within OOI’s internal data processing by Annotations that provide technical notes or qualitative data assessments of the instrument added by staff from the institutions operating the sensors. They represent the first human-in-the-loop (HITL) quality control review of the data coming from the sensor. Annotations are ideal for removing known and identified bad data from a dataset before further processing. While it is not within the purview of OOI to comprehensively flag all such issues, existing annotations will provide valuable and time-saving information to support end user analysis.

Annotations contain important information about the state of the sensors and/or the platforms they are deployed on. This can include known data gaps, sensor or platform failures, power or communications disruptions, changes in instrument configuration or sampling rate, data quality, and suspect/interesting data values. There are a number of pre-set flags which may be associated with these annotations to indicate their topic: fail, not_available, not_evaluated, not_operational, pass, pending_ingest, suspect, or "note" (which means it doesn't fit the other categories). Depending on the type of information being annotated, annotations may either be open-ended, with a start time and no end time, or may have both a start and end time. Similarly, some annotations will indicate issues that affect an entire platform, while others will indicate more specific concerns with a particular sensor or data stream.

End users are recommended to review all annotations that appear in a search for their target BGC data stream to evaluate any potential impacts. Annotations are "live" in that they are constantly being updated/added to as deployments start/stop and/or instruments are reviewed. End users might therefore want to revisit the annotations from time to time to see if there have been any issues noted since the last time they requested the data.

1.4 OOI Data Access

The goal of this best practices and user guide document is to serve as a resource to users accessing data through any of OOI's supported approaches, including through future planned changes in OOI's supported data access tools. **This document is therefore not intended as a guide for how to access OOI data.** The OOI website includes a starting page on "How to Access Data" that synthesizes a variety of different approaches for accessing the data collected by the OOI program. This page is regularly updated and should be a user's primary resource for information and guidance on how to access OOI data.

There are two main sensor data access tools we recommend and use in examples throughout the rest of this document. In 2021, OOI deployed a new tool, the **Data Explorer**, intended as a new primary means for users to access and visualize OOI data.The **OOINet Data Portal** was the primary means of accessing OOI data before the Data Explorer was developed. As of December 2022, not all data streams were available through the Data Explorer; however, all data streams will eventually be available. The Data Explorer and OOINet Data Portal are the primary sources for sensor data and HITL annotations discussed throughout the rest of this document, although other data access options may be preferred for some user applications. Both the Data Explorer and OOINet Data Portal provide data visualizations that can assist end users in the preliminary steps of identifying data that may be suitable for their target analyses. Once target BGC data has been identified, however, end users should download these datasets for further inspection and processing to prepare the datasets for analysis.

The GitHub repository for Worked Examples provided as a supplementary resource to this document (see Section A5) also includes two data access examples for users who would like to see an application of current data access methods for OOI BGC datasets.

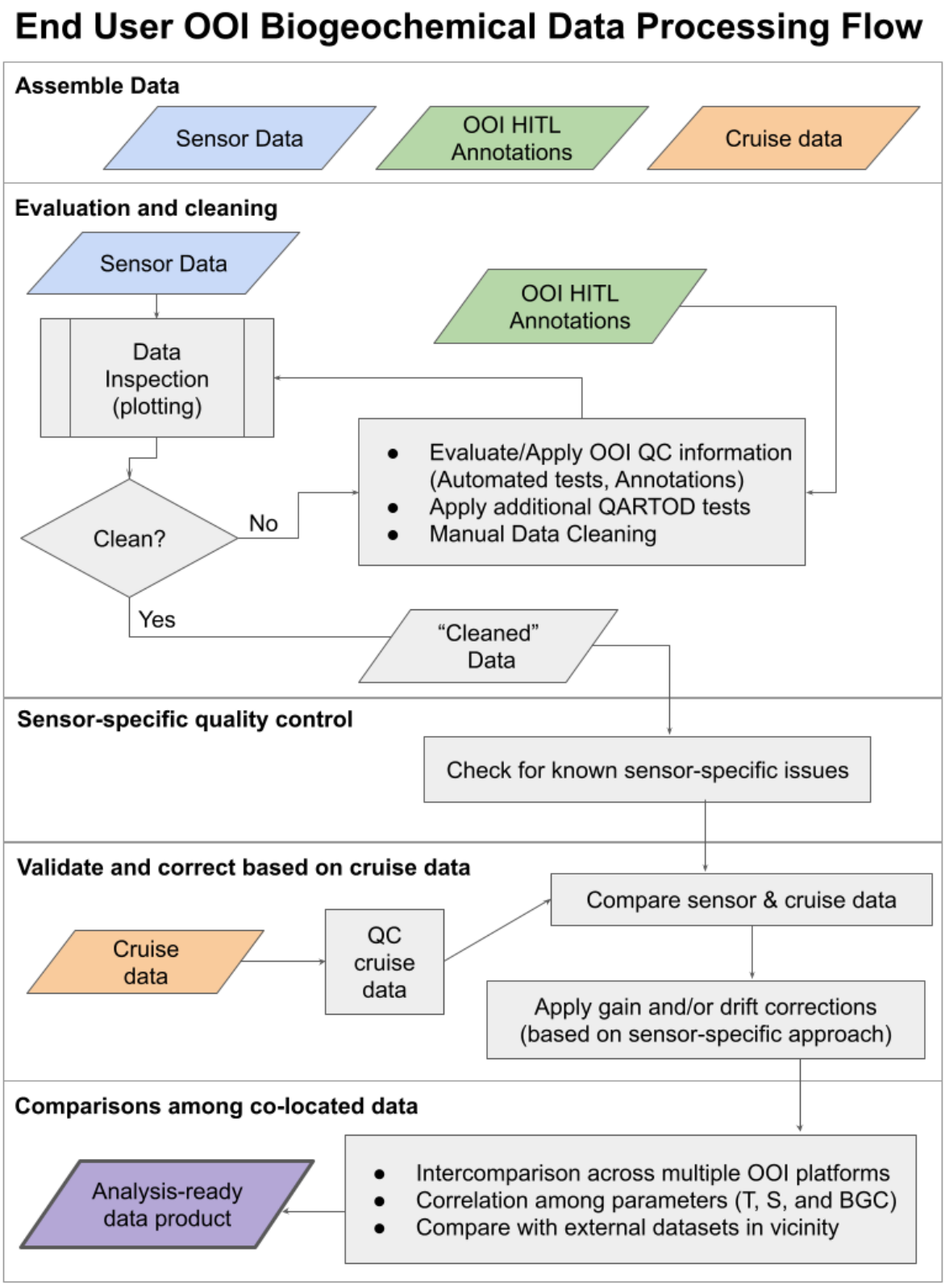


Figure 1.1. Summary of recommended end-user quality control and data processing steps common to all of the OOI biogeochemical sensors.

1.5 Overview of end user QA/QC recommended for all sensor types

To prepare analysis-ready data products from OOI BGC sensor data, end users must undertake additional QA/QC and calibration steps beyond those completed internally by OOI. We describe here recommended end-user quality control and data processing steps common to all of the BGC sensors, which are summarized by an accompanying flowchart (Figure 1.1). Chapters 2-5 provide detailed instructions on the sensor-specific application of these processing steps, with additional flowcharts in each chapter illustrating the sensor-specific application of these processing steps.

1.5.1 Assemble data

The end user data processing workflows recommended in this document are predicated on users assembling three types of data that are needed to create analysis-ready biogeochemical datasets. This data assembly step can be completed using any OOI-supported data access approach of the end user's choice (see Section 1.4 on OOI Data Access).

1. **Sensor data.** This includes:
    a. The BGC sensor data the user intends to prepare for analysis
    b. Flags from OOI's currently-implemented QC tests (see Section 1.3.2), which should be automatically included along with the BGC sensor data
    c. Data from a co-located CTD (temperature, salinity, pressure/depth, and timestamp for alignment with the BGC sensor data), deployed alongside all BGC sensors in the OOI program. In cases where CTD variables are used in OOI's internal processing of the BGC data stream, those variables will be provided along with the BGC sensor data stream. For many BGC variables, however, some or all CTD data are not included and the end user will need to separately download and align the BGC and CTD data streams.

2. **OOI HITL Annotations** (see Section 1.3.3) for all sensor data streams. End users can access OOI's data Annotations through both the Data Explorer and OOINet Data Portal. These annotations can be viewed directly in designated "Annotations" tabs within the web interfaces for these tools, as well as downloaded for later review. In Data Explorer, annotations can be directly downloaded as a csv table from the Annotations tab. In the OOINet Data Portal, the user can check the box for "Download Annotations" when downloading their desired data streams.

3. **Turn-around cruise data.** Discrete sample data collected on turn-around cruises are available through the Data Explorer, and can also be accessed through an OOI managed document storage system called **Alfresco**. The Alfresco server is a repository of all turn-around cruise data and array documentation, including: vendor documentation on delivery, recalibration, and refurbishment of instruments, OOI's internal design documents and pre- and post-deployment testing documentation for all instrument and

platforms, cruise reports for all turn-around cruises to service the arrays, and all data collected on those cruises, including CTD cast and discrete bottle sample data.

Sampling logs and summary sheets for discrete water samples collected upon deployment and recovery can be found on the Alfresco Server using the following path:

```
OOI > {Array ID} > Cruise Data > {Cruise ID} > Ship Data >
Water Sampling
```

Further details about the turn-around cruise data and documentation on accessing data and documentation from the Alfresco server are provided on OOI's "Cruise Data" webpage.

1.5.2 Evaluation and cleaning

The initial steps in recommended end user processing for all OOI BGC sensor datasets focus on preparation of a "cleaned" dataset, applying both automated and HITL QA/QC. This encompasses end user application of the QA/QC steps completed as part of the internal-to-OOI data processing described above, as well as recommendations for additional processing steps not yet implemented by OOI or that require special attention for working with BGC sensor data.

The flowchart summary of this step (Figure 1.1) emphasizes the iterative nature of the process of preparing a "cleaned" dataset ready for further processing. We recommend that users begin by plotting and inspecting the sensor data, and then following each of the subsequent steps iteratively to identify data that may need to be filtered or removed.

1. **Evaluate and apply the OOI-provided HITL Data Annotations.** The Data Annotations process completed internally within OOI is the first HITL step in identifying commonly known issues. End users should precede their own data inspection by reviewing these annotations, which will identify many known issues, such as from platform malfunctions or instrument failures.

2. **Evaluate and apply the OOI-provided data quality flags.** These tests are a valuable initial step in identifying suspect data points. Note that data flagged by these QC algorithms are not necessarily bad, but "of interest" and require further human scrutiny. These data could be flagged by the system because of a sensor issue or because of a previously unobserved phenomenon. At the time of writing this document, the majority of these QC algorithms are still in development by the OOI data team for automated application to BGC data.

3. **Apply additional QA/QC algorithms based on published QARTOD recommendations**. We recommend that end users currently working with these data streams apply their own automated algorithms for QARTOD tests listed in Table 1.2 that have not yet been implemented by the OOI data team. Chapters 2-5 provide more

detailed information, context and recommendations for development and application of automated QARTOD algorithms for each of the biogeochemical sensor data types.

4. **Manually inspect data to identify and address commonly-known issues.** Since comprehensive HITL inspection of all datastreams is not within the OOI purview, additional end user HITL data cleaning is needed. This is critical to identify issues such as biofouling, offsets from factory-supplied sensor calibration, or drift over the course of a deployment. Each of the sensor-specific chapters that follow within this document highlight known issues which are especially common for these sensor types, or to which these data types are especially sensitive.

1.5.3 Sensor-specific quality control

Chapters 2-5 provide descriptions and examples of sensor-specific issues that require additional inspection by the end user to identify and address to produce a clean, analysis-ready final dataset. End users should be sure to reference the sensor-specific flowcharts and section on Common Data Quality Issues within each chapter, and inspect their own datasets for these issues before proceeding.

1.5.4 Validate and correct sensor data based on OOI turn-around cruise data

OOI turn-around cruises to recover previously-deployed assets and deploy replacements include routine CTD casts co-located with moored platforms and just-deployed mobile assets, and the collection of discrete samples from Niskin bottles sampled on those casts. **For all sensor types, we recommend that end users identify and include in their analysis relevant co-located CTD cast and discrete sample bottle data.**

Bottle data provide an opportunity to validate the sensor measurements against independently-collected measurements with the accuracy and precision of wet chemistry laboratory analyses. For some sensor types (e.g., oxygen), CTD casts will routinely include a comparable sensor attached to the CTD collecting full depth profile data as well as discrete bottle samples, whereas for others (e.g., carbonate chemistry and nitrate), these validation data are restricted only to bottle samples. Full details of the types of validation data routinely collected by OOI for each biogeochemical variable are discussed in detail in the individual chapters of this document.

Interpretation and use of turn-around cruise data to validate and/or correct sensor measurements will depend on spatial and temporal variability of the array site at the time that these data were collected. This is especially important in highly dynamic and heterogeneous environments, such as coastal surface waters. We therefore provide guidelines for key factors to consider in applying these data rather than a prescriptive recipe for validation and correction:

- ➔ End users must provide their own quality control on the calibration cast CTD and discrete sample data from OOI turn-around cruises. Discrete water sampling data from each turn-around cruise is accompanied by a README file (found in the same sub-folder within the Alfresco server as the summary and individual variable data). We recommend that end users review this README information about how discrete samples were collected and perform their own manual inspection and quality control of discrete sample data prior to comparing with the sensor data.

- ➔ CTD data can be used to assess alignment between data collected from cruise CTD casts and the mooring and/or mobile asset data they are intended to validate. All biogeochemical sensors on OOI platforms are co-located with CTDs, facilitating this comparison. We recommend that comparison of temperature and salinity be included alongside the biogeochemical variable of interest in comparison of validation samples from cruise CTD casts (both from biogeochemical sensors and bottle samples) with mooring and/or mobile asset data. OOI-provided CTD data has not been corrected for alignment with discrete salt samples, so end users may want to make their own salinity corrections to the CTD data.

- ➔ Shipboard data collected from the continuous underway seawater system on turn-around cruises (available on OOI's Alfresco server and R2R) provides an additional valuable data source for assessing spatial variability. Note when applying these data that attention to careful cleaning and calibration of shipboard underway seawater systems may vary, and attention should also be paid to QA/QC of these data.

- ➔ Approaches to comparison between turn-around cruise data and OOI-deployed sensors will differ for moored sensors deployed at fixed depths versus for profiling assets (profiler moorings and gliders).
    - ◆ *For profiling assets*, more robust comparisons are possible by using full profiles, incorporating validation data from multiple bottle samples and/or full-depth profiles from sensors included in the ship's CTD package (illustrated in the Nitrate and Bio-Optics Worked Examples; see Sections 3.6 and 5.6). We recommend that comparisons incorporate full depth profiles, using associated temperature and salinity data in combination with depth to ensure alignment of water masses. In making these comparisons, it is important to account for sensor response time and resultant lag issues, which for some sensors will require correction prior to making these comparisons.
    - ◆ *For fixed depth moored sensors*, the OOI discrete sampling program prioritizes collection of discrete bottle samples at depths matching the depths of moored sensors. To assess alignment between these individual bottle samples and the moored sensor, we recommend that end users incorporate full CTD depth profiles from the cast where the bottle sample(s) were collected, as well as temporal variability as recorded by the moored sensor in question.

➔ In some cases, samples collected on turn-around cruises can be used to correct for offsets in the sensor measurements (e.g., a gain correction). This is especially important in cases where sensors are known to drift between the time of factory calibration and deployment and where the sensor does not include a reference standard (e.g., oxygen sensors). However, especially in cases where the number of samples available for comparison is limited (e.g., fixed depth sensors being compared with a single depth-aligned bottle sample), differentiation between potential sensor offsets requiring a gain correction and expected spatial and temporal variability in the system may not be possible. In these cases, samples should be used for validation that the moored sensor measurements and cruise sample(s) fall within a mutually-consistent range given the spatial and temporal variability at the time of sampling.

1.5.5 Comparisons among co-located data

An advantage of the design of all OOI arrays is the opportunity to intercompare observations across multiple simultaneously-deployed sensors and platforms. **We recommend that end users identify and leverage these opportunities for intercomparison**:

➔ During turn-around cruises, OOI standard operation for all uncabled arrays is to deploy new moorings prior to recovering those currently in operation. This routinely creates an opportunity to compare ~1-3 days of simultaneously collected data from consecutively-deployed moorings at the same location. However, note that this overlap data may be absent in cases where ship operations were modified to adapt to adverse weather conditions, and cases of instrument or platform failure or power loss prior to recovery. Data accessed via Data Explorer are stitched together across multiple deployments, removing information about individual deployments. Data accessed via the Data Portal separately provide data for each subsequent deployment and include a “deployment” variable, and so are recommended for assessment of overlap periods. In most cases, the co-located calibration cast from the turn-around cruise will be timed during the period when both moorings are in the water, providing additional context for this overlap period.

➔ Throughout each deployment, comparison is also possible among sensors measuring the same variable at different locations within the array (Table 1.1, Figures A.1-A.6). The inclusion of gliders that transit between the moorings throughout each deployment provides an opportunity to compare measurements from glider profiles collected in proximity to the moorings with both fixed-depth and profiler mooring measurements. In some cases, multiple fixed depth sensors may also be measuring the same water mass (e.g., multiple measurements within a deep surface mixed layer), providing an opportunity to check sensor agreement. Careful assessment of spatial and temporal variability is again necessary to contextualize and identify appropriate opportunities for these comparisons.

- ➔ All BGC sensors are deployed alongside CTDs providing temperature and salinity data, and all OOI arrays include multiple BGC sensors. In addition to providing rich opportunities to address scientific questions, these combinations of sensors are also useful in data QA/QC. We recommend that end users leverage the combinations of multiple variables from simultaneously-deployed sensors (e.g., property-property plots, correlation analysis, etc.) to provide additional validation of sensor performance. This may be especially helpful in cases of unusual signals, where a co-variation across multiple variables measured on multiple sensors would be expected from a real environmental signal whereas an anomaly in a single variable may instead reflect a sensor malfunction.

- ➔ External-to-OOI datasets in the vicinity of the OOI arrays can provide additional measurements of the BGC variables measured by OOI (e.g., discrete samples from GO-SHIP repeat hydrography or Line P cruises to the Station Papa array). These data can provide an opportunity to compare additional validation data beyond that provided by the OOI program. We recommend that end users consider whether any such external datasets are available and incorporate them into validation and QA/QC of OOI data where close-enough alignment with such an external dataset makes this possible.

1.6 Outlook and Concluding Remarks

This guide aims to present the current best practices for the use of OOI BGC sensor data at the time of its writing in 2022, while recognizing that both the OOI program and the BGC community will continue to develop new best practices over time. As noted above, at the time of writing this guide, plans are underway for the relocation of the Pioneer Array from the New England Shelf where it is currently deployed, to the Southern Mid-Atlantic Bight. This move will entail changes to the configuration of the array, with the possibility of the existing sensor suite being deployed in different ways. However, in this case, the contents of this document will still apply. It is also possible that new sensors will be added to the arrays over time, and that new array configurations with respect to the spacing between assets or the timing overlap between deployments may require considerations beyond the scope of this document, which is primarily concerned with the existing individual elements of the arrays (i.e., the sensors). However, we also envision that this guide will be reviewed and updated as needed at appropriate intervals, taking these changes into account.

Chapter 2: Dissolved Oxygen

2.1 Introduction to OOI Dissolved Oxygen Sensors

There are two types of dissolved oxygen (DO) sensors within the OOI program. The first type, the Fast Response Dissolved Oxygen (DOFST, contraction of DO FAST) instrument manufactured by Sea-Bird Electronics, is used to measure dissolved oxygen concentration on shallow coastal profilers through strong oxygen gradients. The second type, the Stable Dissolved Oxygen (DOSTA, contraction of DO STABLE) instrument manufactured by Aanderaa, is used on mobile assets, deep profilers, fixed-depth risers, and moorings.

Table 2.1: Manufacturer, model and internal-to-OOI instrument class-series (six letter reference indicator) for the four models of oxygen sensors operated by OOI, and the platforms on which they are deployed. For the class-series, the class DOFST represents “Dissolved Oxygen Fast Response” (Sea-Bird sensors) and DOSTA represents “Dissolved Oxygen Stable Response” (Aanderaa optodes). The last letter of the class-series is an internal classification by OOI and represents the series (determined by specifications related to the sampling rate, deployment duration, deployment depth, etc.). For a summary of where oxygen sensors are deployed across the full OOI program, see Table 1.1. For deployment locations within each array, see Figures A.1-A.6.

Manufacturer	Model	OOI Class-Series	Platforms
Sea-Bird	SBE 43	DOFSTA	Subset of Cabled & Endurance Profiler moorings
	SBE 43F	DOFSTK	Pioneer Profiler moorings
Aanderaa	Optode 4330	DOSTAN	Pioneer AUVs
		DOSTAL	Global sub-surface Profiler moorings
	Optode 4831	DOSTAD	Fixed depths on all array moorings; subset of Cabled & Endurance Profiler moorings
		DOSTAJ	Endurance surface-piercing Profiler moorings
		DOSTAM	Gliders

The OOI deploys two different models of Sea-Bird SBE instruments (SBE 43 and SBE 43F, 2-5 sec response time to 63%) and two different models of Aanderaa Optodes (optode 4330 and optode 4831, <25 sec response time to 63%). An overview of the measurement principles for the Sea-Bird and Aanderaa sensors is available through the IOCCP and BONUS INTEGRAL course for ocean biogeochemical sensors and a review paper by Bittig et al. (2018). Briefly, the SBE sensors use a membrane-covered polarographic probe (Clark cell) that determines dissolved oxygen concentration by measuring the flux of oxygen molecules per second that diffuse through the membrane to the electrode, and which consumes oxygen as it is measured.

The SBE 43 is designed for use with pumped CTDs (sample water pumped through a tube) to provide optimal correlation with CTD measurements, whereas the 43F is easily integrated onto moored profiling platforms. More information on the differences between SBE 43 and SBE 43F is available on the manufacturer's website. The Aanderaa optode operates on dynamic luminescence quenching, which measures oxygen through a non-consumptive method. The two Aanderaa optode models are similar but have different connectors, optimal operating depths, and response times, and are therefore deployed on different platform types. These operating differences are described on the manufacturer's website. The manufacturer and model of oxygen sensors on each type of OOI platform is summarized in Table 2.1.

2.2 OOI standard practices for oxygen sensor deployment and calibration

2.2.1 Calibration information

When purchased, both Aanderaa and Sea-Bird sensors undergo a multipoint calibration across different combinations of temperature and dissolved oxygen saturation by their respective manufacturers. Multi-point calibrations are required to characterize the nonlinear response of the oxygen optode at different oxygen concentrations and temperatures.

Aanderaa Optodes

When purchased, Aanderaa optodes models 4330 and 4831 undergo a multipoint calibration across a surface of 40 different combinations of temperature and dissolved oxygen saturation by the manufacturer (Aanderaa, Norway). Once optodes undergo this multi-point calibration, Bittig et al. (2018) has demonstrated that timely two-point calibrations (0 and 100% saturation) in conjunction with the original multi-point calibrations are sufficient for maintaining accurate calibration of sensors. Optodes within the OOI undergo a two-point calibration after being deployed for 6 months or after two years from previous calibration, whichever comes first. If this 2-point calibration results in a large slope or intercept, the optode is returned to the manufacturer for a multipoint calibration. The original multipoint calibration sheets provided by the manufacturer and the 2-point calibrations performed by OOI are available on the Alfresco server:

```
OOI > Instrument & Platform Documents > Calibration and Repair
> {Array type - Coastal-Global or Cabled} > DOSTA
```

Multi-point calibration coefficients are uploaded to a calibration folder within OOI's asset-management Github repository as CSV files for use in the data processing pipeline. If any sensor undergoes a subsequent two-point calibration, a new calibration file is created, adding the slope and intercept from the two-point calibration.

Sea-Bird Scientific Sensors

At the factory, Sea-Bird performs a multi-point freshwater calibration in a water bath across a surface of different temperatures and DO concentrations (six different water temperatures and three or four different oxygen concentrations at each temperature) in conjunction with a pumped CTD. All calibrations are performed by Sea-Bird, and sensors are sent back to the manufacturer for multi-point calibrations between deployments. The multipoint calibration sheets provided by the manufacturer are are available on the Alfresco server:

```
OOI > Instrument & Platform Documents > Calibration and Repair
> {Array type - Coastal-Global or Cabled} > DOFST
```

2.2.2 Sensor turnaround information

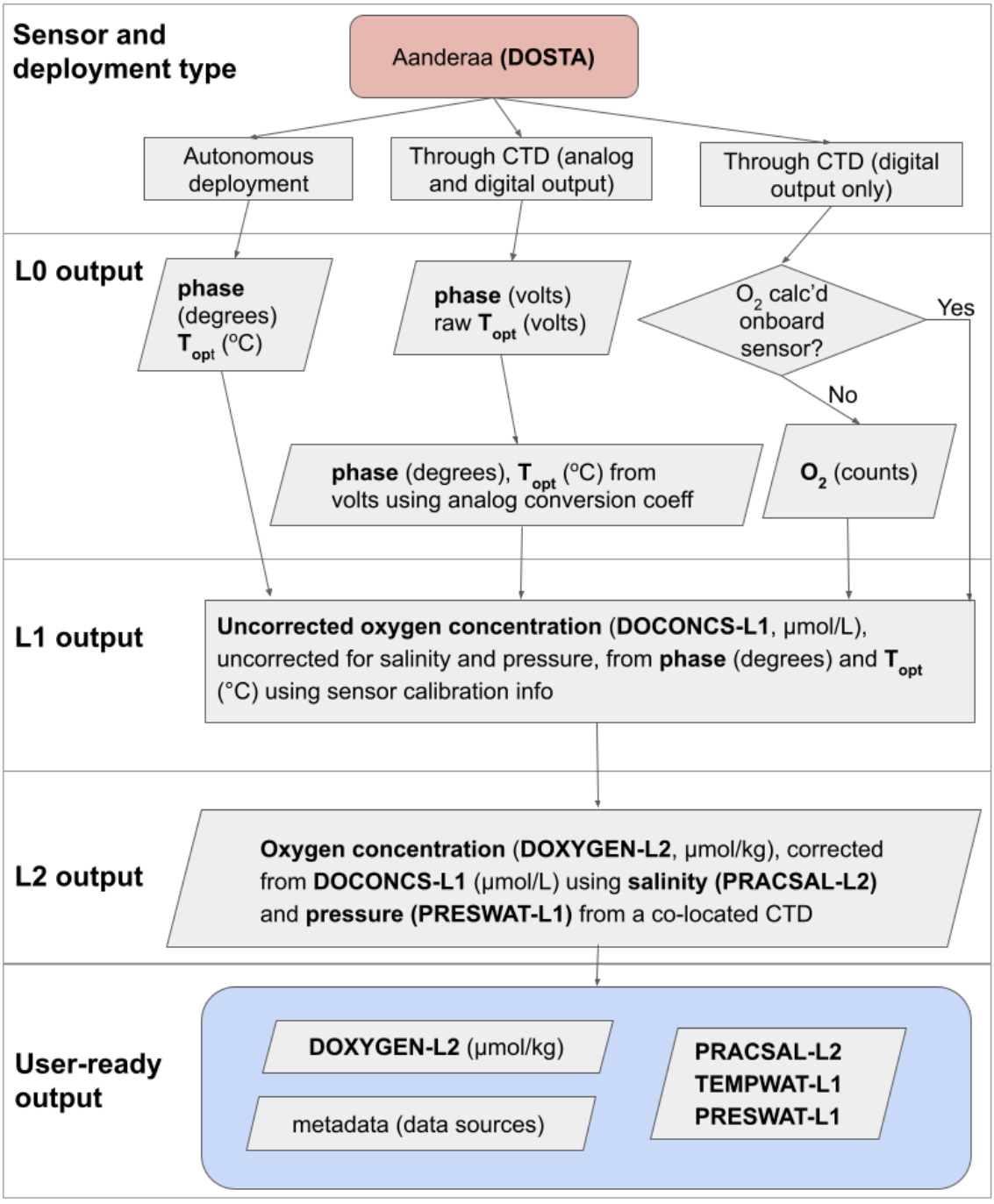
Before being deployed by the OOI, instruments are checked for ground faults and reasonable data output. They are then powered up and run shoreside (for hours to two weeks) to check for any electrical issues that may come up. Aanderaa optodes are kept wet and covered with a sponge during this time period. The sponge and cover are removed on deck right before deployment. Sea-Bird oxygen sensors, which are integrated into McLane profilers, are delivered to OOI from McLane dry. OOI fills the tube connecting the CTD and Sea-Bird DO sensor with water. The plugs on either end are removed right before deployment. Deployments take up to a few hours, depending on the platform.

Upon recovery, sensors are photographed for signs of damage or biofouling. Sea-Bird CTD and DO sensors are cleaned and returned to the vendor for calibration. Since mid-2021, OOI has been performing internal two-point calibration verifications upon recovery of most Aanderaa optodes (see Calibration Information above). OOI stores Aanderaa optodes dry and in the dark in between deployments.

2.3 Internal to OOI Data Processing Workflow

Figures 2.1a-b summarize OOI's internal processing workflow for data from both Aanderaa and Sea-Bird oxygen sensors. Data processing for oxygen data products depends on the type of instrument (Sea-Bird vs Aanderaa) and the type of deployment configuration (e.g., autonomously deployed, deployed through a CTD) of the oxygen sensors used. Briefly, oxygen data are transformed from an unprocessed, Level 0 (L0) data product in raw instrument units, to an intermediate Level 1 (L1) data product, to a OOI-provided Level 2 (L2) data product in µmol kg^{-1} that has been temperature-salinity-and pressure corrected using open-source algorithm code available on GitHub.

OOI Internal Oxygen Processing Flow: Aanderaa (stable response oxygen sensor)



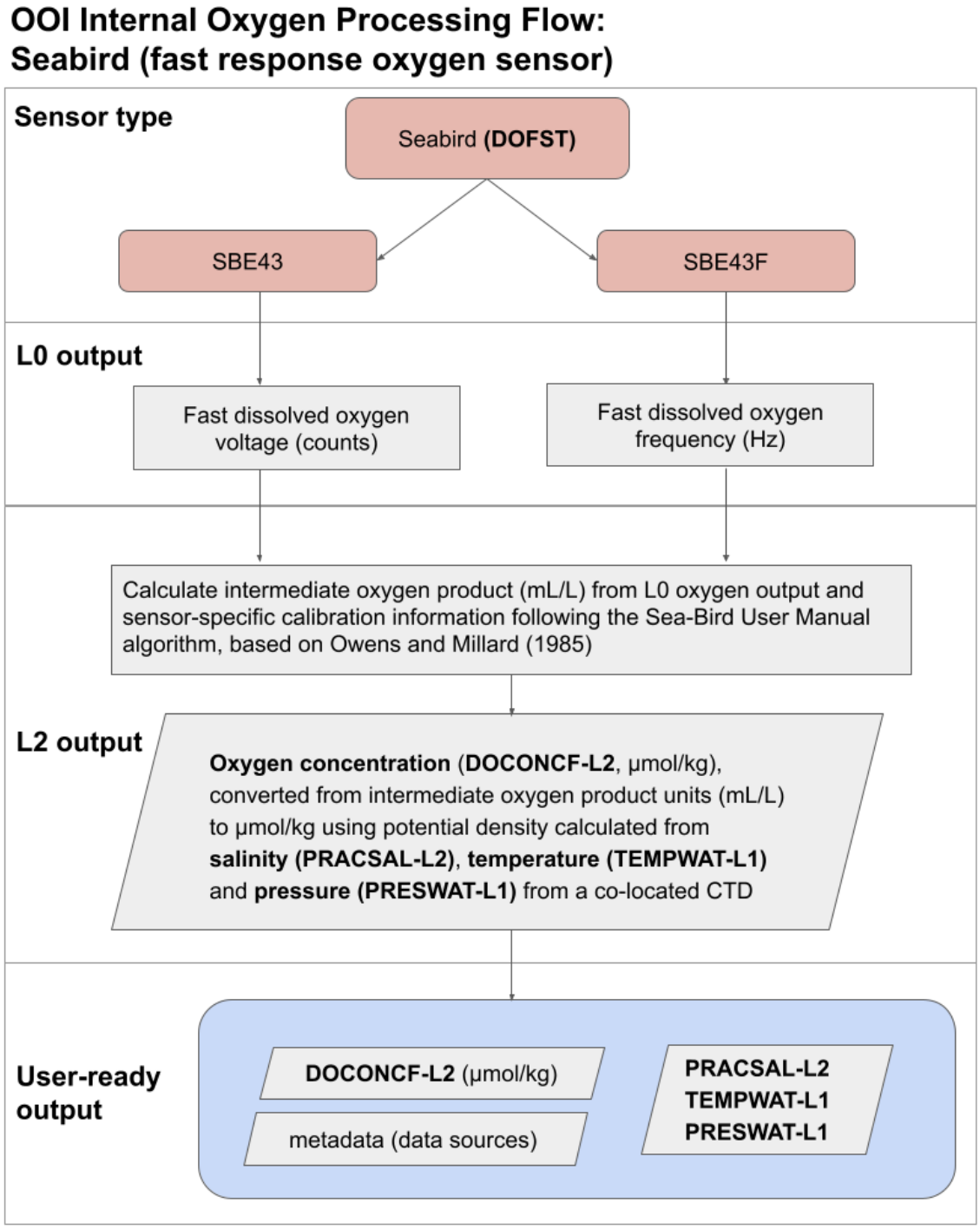


Figure 2.1a-b: Summary of the internal-to-OOI processing workflow for oxygen data from a) Aanderaa (DOSTA, previous page) and b) Sea-Bird (DOFST, this page) oxygen sensors. The user-ready data products provided by OOI are not yet analysis-ready oxygen data and require further processing by the end user to evaluate, clean, and apply corrections to the data. The user-ready output from OOI's internal data processing in these figures is the starting point sensor data for the recommended end user data processing outlined in Section 2.5 and summarized in the flowchart in Figure 2.5.

Detailed OOI documentation describes the computations used to calculate the L2 product for Fast Dissolved Oxygen Sensors (Sea-Bird sensors, sensor code = DOFST; OOI Document Control Number 1341-00521) and for Stable Response Oxygen Sensors (Aanderaa optodes, sensor code = DOSTA; OOI Document Control Number 1341-00520). Briefly, Sea-Bird oxygen products are calculated using an algorithm based on that of Owens and Millard (1985) that incorporates raw voltage data from the Sea-Bird Electronics SBE 43 and 43F Dissolved Oxygen Sensor family of instruments along with L1 and L2 data products from the co-located CTD instruments. End-user products from Sea-Bird sensors are always corrected for salinity, temperature and pressure from the co-located CTD, whereas oxygen products from Aanderaa optodes are corrected for temperature using the sensor's optode foil temperature (T_{opt}) measured by the optode's thermistor. T_{opt} is the preferred parameter because the permeability of the sensor's foil for oxygen is sensitive to the temperature directly at the sensor's foil. After this T_{opt} correction, Aanderaa oxygen products are corrected for salinity and pressure and converted from µmol L^{-1} to µmol kg^{-1} using the co-located CTD data. For both the Aanderaa and Sea-Bird data processing workflows, instrument-specific calibration coefficients are applied during the calculation of the Level 1/Intermediate oxygen products.

2.4 Common Data Quality Issues

2.4.1 Oxygen Sensor Drift

Both types of oxygen sensors deployed by OOI, the Aanderaa optodes (DOSTA) and Sea-Bird SBE43 membrane sensors (DOFST) are subject to drift from factory-calibrated values. **It is critical to check and correct for sensor drift prior to any quantitative interpretation of oxygen data.** Drift in these oxygen sensors has previously been widely documented and routinely corrected for by the oceanographic community (e.g. Nicholson et al., 2008, Bushinsky et al., 2016, Bittig et al., 2018). Aanderaa optodes are expected to experience less drift than the SBE43, hence their designation by OOI as the "stable" oxygen sensor and more widespread deployment across the OOI program (see Tables 1.1 and 2.1). The tradeoff is that the Aanderaa optodes have a slower response time than the SBE43, designated by OOI as the "fast" oxygen sensor and used in place of the Aanderaa optode on coastal profiler moorings.

There are two types of oxygen sensor drift that need to be corrected to produce accurate analysis-ready oxygen data: 1) pre-deployment drift between the time sensors are calibrated in the factory and the time they are deployed (also called *ex situ* or storage drift; example shown in Figure 2.2), and 2) deployment drift over the course of a sensor deployment (also called *in situ* drift, example shown in Figure 2.3). Oxygen optodes drift more rapidly during storage than during deployment (order –5% $year^{-1}$ during storage, as compared to ~0.5-1 % $year^{-1}$ observed during deployment on Argo floats; Bittig et al., 2018). Drift of SBE43 Clark electrode oxygen sensors has been less extensively characterized by the community due to the preference for the stability of optodes during long-term deployment, but has previously been reported as 5-10% over a 100 day deployment (Nicholson et al., 2008).

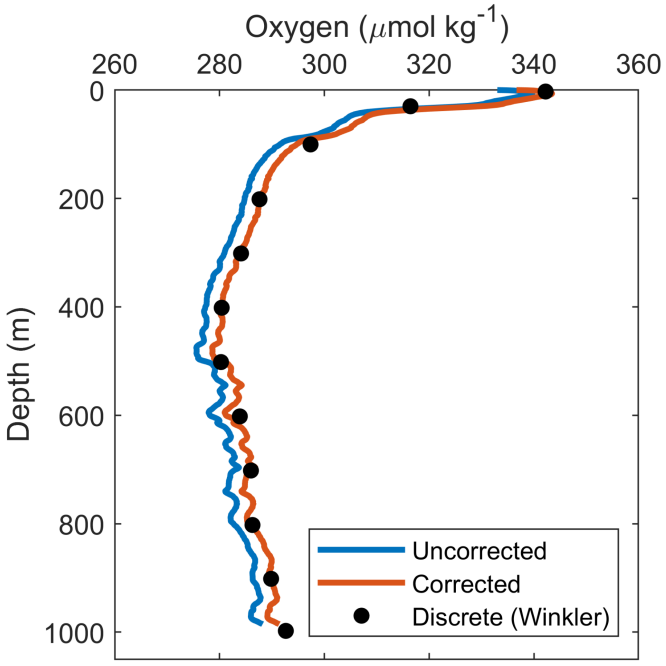


Figure 2.2. Example of a correction for pre-deployment storage drift on an Aanderaa optode deployed on an Open Ocean Glider at the Global Irminger Sea Array in June 2018. Discrete (Winkler) sample oxygen measurements from a co-located CTD cast shortly after deployment do not match the uncorrected sensor-measured Level 2 (L2) dissolved oxygen from the glider. This mismatch is the result of storage drift, which can be corrected for by using Winkler measurements to calculate a gain correction applied to the full profile (for more information on gain corrections see Section 2.5.4).

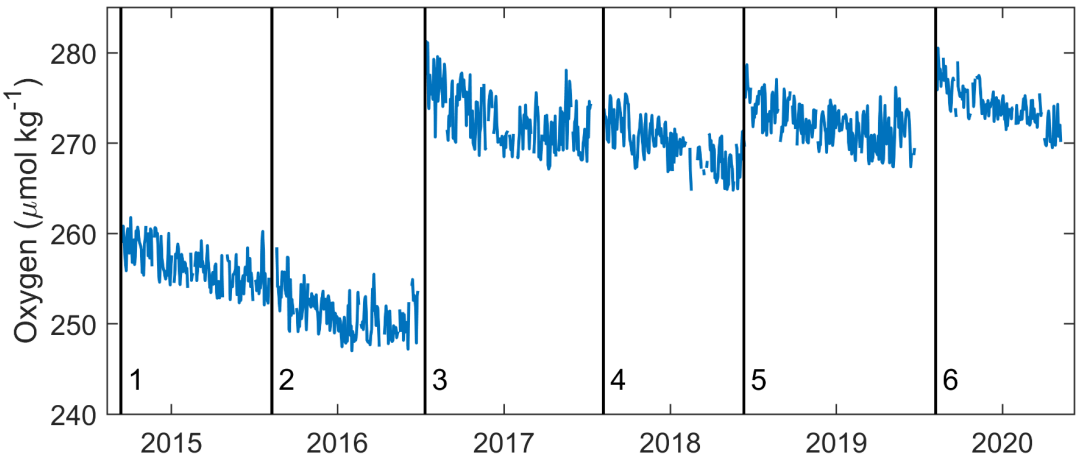


Figure 2.3: Example of *in situ* oxygen sensor drift during deployment of Aanderaa optodes deployed on the Apex Profiler Mooring at the Global Irminger Sea Array. Data shown are Level 2 (L2) dissolved oxygen data along 2.5°C isotherm (~2300 m) for each of the first six years of deployment. Oxygen at these depths is expected to remain stable and observed change over time at these depths can be used to determine the rate of sensor drift (Takeshita et al., 2013; Palevsky and Nicholson, 2018). The lower measured oxygen values in the first two years of deployment as compared to later years is the result of pre-deployment storage drift, which was lessened from deployment year 3 onward by using older, pre-conditioned optodes. For more information on corrections for *in situ* drift, see section 2.5.4.

Drift in Aanderaa oxygen optodes is attributed to changes in the sensing foil that decrease luminescence quenching, which can result from changes in oxygen accessibility to the sensing

foil (i.e., migration of the luminophore molecules within the foil) or oxygen diffusivity within the sensing foil (Bittig et al. 2018). Drift in optodes is most commonly downward, reflecting a decrease in sensitivity over time, but upwards drift has also been observed (Bushinsky et al., 2016). Drift has been observed to occur more rapidly when foils are new, with slower drift observed in older optodes and those that have undergone a burn-in process to pre-condition prior to calibration (Bittig et al., 2018). Aanderaa pre-treats foils with a burn-in process before the foils are mounted on optodes for their initial multi-point calibration. To help reduce oxygen sensor drift, OOI has added supplemental sensor burn-in as part of their routine pre-deployment preparation for optodes with new foils, and reuses optodes unless they have been mechanically damaged. For a comprehensive discussion of storage and deployment drift of oxygen optodes, as well as suggested methods for minimizing and correcting errors due to drift, see Bittig et al. (2018), and references within. Note that oxygen optode *in situ* drift has been most robustly characterized in the literature for deployment on BGC-Argo floats, with drift generally within ~0.5-1% per year; however the more frequent oxygen sampling on OOI platforms can lead to faster drift rates (>10% per year, see Table S2 in Palevsky and Nicholson, 2018, and the Worked Example in this chapter).

Drift in the SBE43 Clark electrode oxygen sensors has different mechanistic drivers than Aanderaa optode drift, though with a similar result of primarily downward drift over time. The chemistry of the sensor electrolyte changes as oxygen is measured, resulting in a slow but continuous loss of sensitivity that produces a continual, predictable drift in the sensor calibration with time. Downward drift also results from membrane fouling, which contributes to drift by altering the oxygen diffusion rate through the membrane, thus reducing sensitivity. Drift can occur both during storage and deployment, with storage drift reduced if the sensor is kept in an oxygen-free environment. For a more detailed discussion of SBE43 dissolved oxygen sensor deployment recommendations, cleaning, and storage to minimize drift, see Sea-Bird Application Note 64.

OOI-provided Level 2 (L2) oxygen data are not currently corrected for pre-deployment storage drift or in situ drift, so end users must calibrate oxygen data released by OOI to correct for drift based on concurrent reference data. Reference data can include discrete sample oxygen measurements (from shipboard titrations), hydrographic profiles calibrated using discrete sample oxygen measurements, and/or air measurements during surface intervals on profiling assets (Bittig et al. 2018). For further details on recommended approaches to correct for both types of drift, see Section 2.5.4.

2.4.2 Oxygen Sensor Time Response

Oxygen sensors have slower response times to changes in ambient conditions than temperature or conductivity sensors, with considerably slower response times from Aanderaa optodes than SBE electrodes (hence the designation by OOI of the SBE oxygen sensors as “fast” for coastal profiling applications in strong oxygen gradients). The oxygen electrodes on the Sea-Bird SBE instruments (SBE 43 and SBE 43F) have a 2-5 sec manufacturer-reported response time to 63%, whereas the manufacturer-reported response time of Aanderaa optodes

is <25 sec to 63%. The longer response time of the optode is due to the diffusion of oxygen through the boundary layer at the water-sensor interface and across the sensor foil. Sensor response times are temperature dependent, with faster response times in warmer waters, and are also dependent on the deployment application (see Sea-Bird Application Note 64, Bittig and Kortzinger 2017, and Bittig et al. 2018 for further detailed discussion and characterization of oxygen sensor response times).

Oxygen data collected on moving platforms (profiler moorings, gliders, and AUVs) needs to be corrected for oxygen sensor response times prior to analysis. Comparison of oxygen profiles from upward-moving and downward-moving profiles on the same platform provides *in situ* data that can be used to characterize oxygen sensor response times (e.g. Bittig et al. 2014, 2018, Bittig and Kortzinger, 2017, Gordon et al. 2020). Availability of paired upward-moving and downward-moving profiles to use for response time corrections will vary across OOI platforms. Due to the need to conserve battery power to last throughout full deployments, some OOI moving platforms such as gliders and the Coastal Surface Piercing Profilers (CSPP) sample only in one direction. The current OOI practice is to sample on both upwards and downwards profiles during the initial days of a new deployment to characterize the response time of the sensors. We recommend that users calculate the sensor- and platform-specific response time correction based on the data collected during these periods (see Section 2.5.3 for further details).

2.4.3 Biofouled and scratched sensors

Damage or scratches to the sensor foil and biofouling can both cause an unrealistically large oscillation in dissolved oxygen on a 24-hr basis. Scratches on the foil can expose photosensitive foil materials leading to increased diel oscillations. When algae form a biofilm over the foil, their photosynthesis and respiration can cause larger diel variability in oxygen to occur at the surface of the foil (i.e., a dramatic increase in dissolved oxygen during daylight hours and corresponding decrease at night, see example in Figure 2.4)., compared to the conditions in the water column. This enhanced diel signal, when present, usually begins to occur a month or two after a mooring is deployed. This issue occurs for near-surface moored instruments (i.e., in the euphotic zone).

Since this issue was identified, OOI started deploying UV lamps adjacent to the dissolved oxygen sensor on moorings and has demonstrated the effectiveness of this approach in preventing a biofilm from forming on the sensor (Figure 2.4). The UV approach is currently employed on sensors in the euphotic zone on the Endurance, Pioneer, and Global arrays and will continue to be used in the future; however, many of the early years of mooring data are affected by biofouling. UV systems were added to euphotic zone fixed depth oxygen sensors across the Endurance Array beginning in 2017-2018, with UV lights added to Coastal Surface Piercing Profilers at the Endurance inshore sites in 2019. Beginning in 2018, UV lights began being deployed alongside fixed depth oxygen sensors across the Pioneer and Global arrays. Full details about the addition of these UV lamps to mitigate biofouling are included in OOI's HITL Data Annotations. If using oxygen data in the euphotic zone collected prior to implementation of the UV lamps, end users should scrutinize the data for potential biofouling signals.

2.4.4 Issues with ancillary data streams

Level 2 (L2) oxygen products are corrected for temperature, salinity, and pressure using ancillary data products from co-located CTDs. Oxygen measured using an Aanderaa optode is corrected for temperature at the foil surface on the Aanderaa optode and then with salinity and pressure from co-located CTDs. Oxygen measurements made with Sea-Bird sensors are corrected with temperature, salinity and pressure from the co-located CTD. As a result, errors in temperature, salinity, and pressure products are incorporated into L2 oxygen products. For example, see OOI Document #3408-30001: Global Wire Following Profiler Salinity Drift Correction for discussion of a salinity data issue, which also affects the L2 oxygen data. End users may need to investigate the quality of ancillary data streams if their scientific purposes require error less than what may be introduced by issues with ancillary data products.

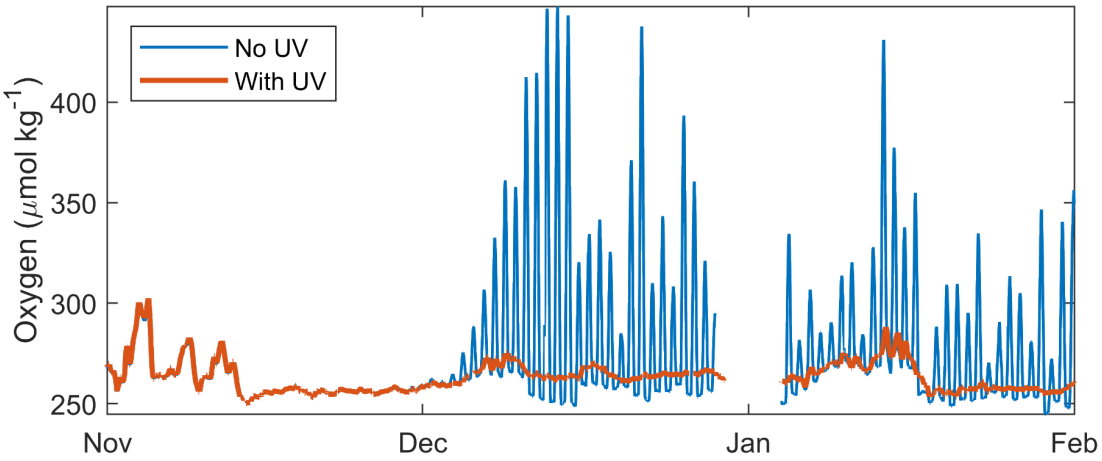


Figure 2.4: Dissolved oxygen concentrations from the near-surface instrument frame at the Oregon Shelf Surface Mooring of the Coastal Endurance Array. Two Aanderaa optodes were deployed simultaneously, one with UV lights (red line) and one without UV lights (blue line). After about a month of deployment, the Aanderaa optode without UV lights (blue line) reflects oxygen signals associated with some combination of foil photosensitivity and photosynthesis/respiration of biofouling growth. The addition of the UV light to the second Aanderaa optode has mitigated biofouling growth on this dissolved oxygen sensor (red line). Also, post-recovery photos (not shown) reveal that the optode without the UV light biofouled and the one with the UV light showed no apparent biofouling.

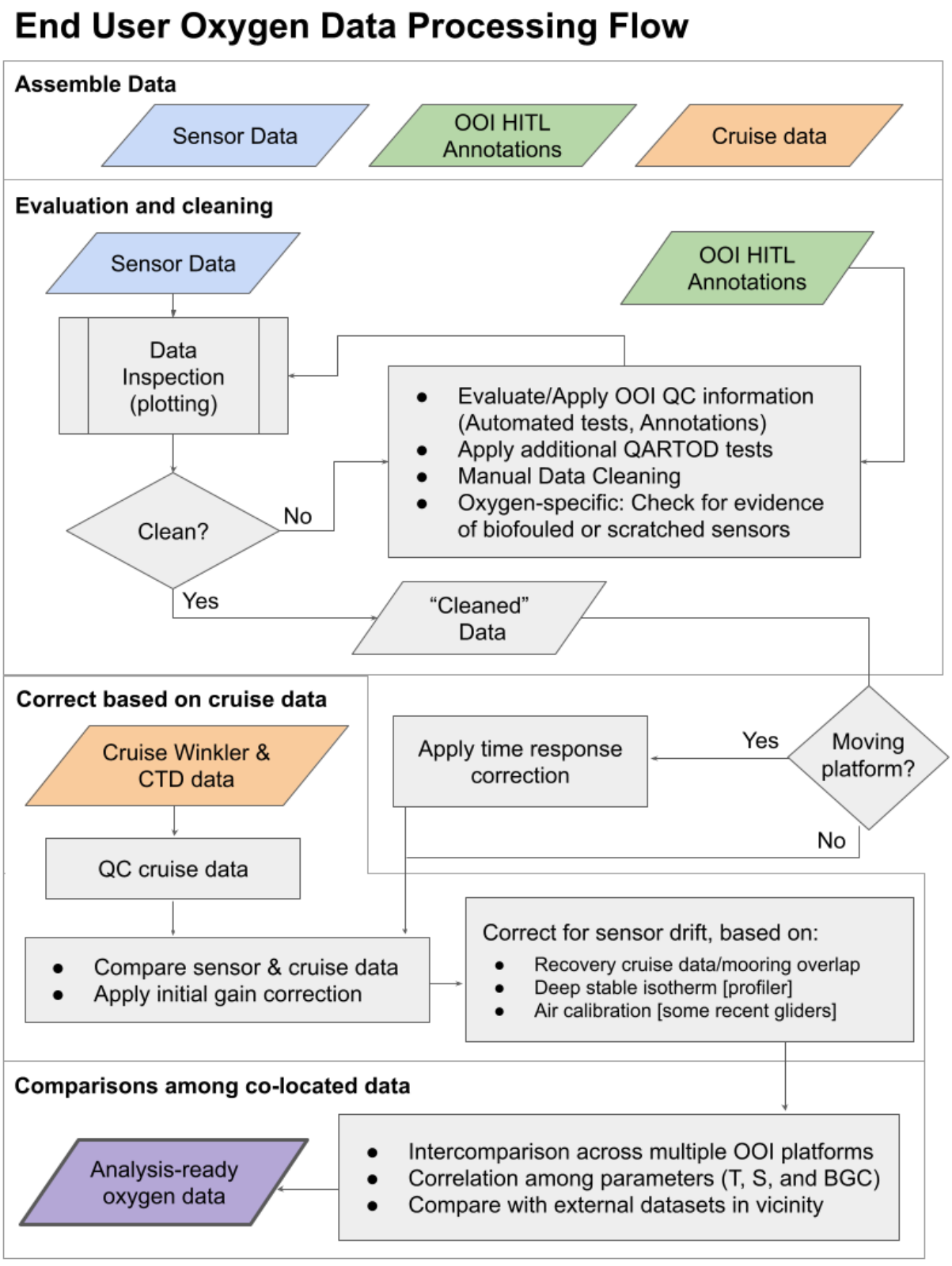


Figure 2.5. Summary of recommended end-user quality control and data processing steps for OOI oxygen sensor data. These steps will allow the user to evaluate, clean, and apply corrections to the sensor data provided by OOI (see Figure 2.1a-b) to prepare analysis-ready oxygen data products.

2.5 Recommended End-User Data Processing

The end user oxygen data processing flowchart (Figure 2.5) summarizes additional data processing that must be performed by the end user to evaluate, clean, and apply corrections to the OOI-provided oxygen data. These steps are essential to prepare the OOI-provided oxygen data for scientific applications and analyses, especially those involving quantitative interpretation.

Chapter 1 of this document provides an overview of QA/QC procedures recommended for all OOI biogeochemical sensors (Section 1.5), and provides a high-level walk-through and context for each step in the recommended end user oxygen data processing summarized in Figure 2.5. We recommend that users intending to work with OOI oxygen data use the flowchart in Figure 2.5 and instructions in Section 1.5 as a starting point and reference for each data processing step. Here, we walk through each of the steps outlined in Figure 2.5, which synthesize the approaches end users can take to correct OOI data products for oxygen sensor-specific behaviors and data quality issues described in Section 2.4.

2.5.1 Assemble data

Users will need to assemble oxygen sensor data, accompanying OOI HITL annotations, and corresponding turn-around cruise data to use in preparing their final analysis-ready oxygen data. See Section 1.5.1 for details of each of these components.

Ancillary CTD data (pressure, temperature, salinity) co-located with the oxygen sensor to be analyzed will in most cases already be merged with the oxygen data stream by OOI since these data are used in preparing Level 2 (L2) oxygen data. Unless issues are identified when reviewing these ancillary CTD data (see Section 2.4.4), the OOI-provided L2 Oxygen product should be used for further data treatment.

OOI turn-around cruises routinely conduct CTD casts near each deployed platform, where casts include a SBE43 oxygen sensor on the CTD sensor package as well as discrete samples collected from Niskin bottles where oxygen is measured via shipboard Winkler titration. We recommend that end users assemble both the discrete bottle data and full depth-resolved CTD cast data to include in their analysis.

2.5.2 Evaluation and cleaning & sensor-specific quality control

The initial step in recommended end user OOI oxygen sensor data processing is to prepare a “cleaned” dataset, applying both automated and human-in-the-loop (HITL) QA/QC to evaluate the data and identify points that may need to be filtered or removed. Section 1.5.2 summarizes the recommended steps for OOI BGC sensor end user data evaluation and cleaning. Here we provide additional context on the application of these steps specifically for OOI oxygen data:

1. Evaluate and apply OOI-provided HITL Data Annotations. Annotations by the OOI data team identify many platform-wide issues users may need to be aware of (power failure-caused

data gaps, etc.), as well as oxygen sensor-specific issues. Users of oxygen sensor data should particularly check the Annotations for information about the implementation of UV lamps to mitigate biofouling (see Section 2.4.3).

2 & 3. Apply QA/QC based on published QARTOD recommendations. As of December 2022, automated data quality flags for oxygen are in development within OOI (see Table 1.2). Users are encouraged to check for and apply OOI's automated oxygen data quality flags once they become available. If OOI implementation of these automated tests is not available, end users are also encouraged to implement their own version of the tests listed in Table 1.2 and/or to manually inspect data to determine whether any of these issues are present in their dataset.

4. Manually inspect data to identify and address commonly-known issues. Oxygen data users should check especially for evidence of biofouled and/or scratched sensors (see Section 2.4.3). While some examples of this issue will be flagged by OOI-provided Data Annotations, users should carefully inspect all data in preparation for their own analysis. This issue often develops part-way through a deployment (e.g. Figure 2.4), so data prior to fouling may be salvageable for analysis even if later data are compromised.

2.5.3 Response time corrections

Oxygen data from moving platforms (gliders, profiler moorings, and AUVs) requires sensor response time corrections prior to further analysis. The magnitude of the response time correction and its influence on the final data will depend on the sensor type (Sea-Bird DOFST vs. Aanderaa optode DOSTA - see Table 2.1), the deployment configuration and vertical velocity, and strength of the oxygen concentration gradients through which the data are collected (see section 2.4.2).

- ➔ If both upward-moving and downward-moving profiles are collected throughout the full deployment (usually the case for Wire-Following Profilers and AUVs) we recommend using data throughout the full deployment to calculate the sensor response time and assessing whether there are any changes in response time during the deployment period.

- ➔ If paired upward-moving and downward-moving profiles are only collected for a subset of the deployment (for gliders and Coastal Surface Piercing Profilers), we recommend using the period with paired profiles (usually at the beginning) to calculate the sensor response time to apply throughout the deployment. The most accurate oxygen optode response time corrections will be derived by calculating the temperature-independent boundary layer thickness at the sensing foil, which can be used to calculate the response time at the in situ temperature (Bittig et al. 2014, Bittig and Kortzinger, 2017).

- ➔ For some deployments earlier in the OOI program history, initial paired upwards and downwards data may not have been collected. In these cases, we recommend applying a best-estimate lag correction based on information about the sensor type, vertical velocity speed, and deployment configuration, drawing on other deployments on the

same OOI platform and/or other previously-published estimates on similar platforms (e.g. Bittig and Kortzinger, 2017 for gliders) though this will lead to greater uncertainty in the corrected data.

For an example of how to calculate optode sensor response time using paired upward-moving and downward-moving profiles, along with accompanying MATLAB code that can be applied across a range of moving platforms, see Gordon et al. 2020. For further information on the range of community-endorsed best practices for oxygen sensor response time corrections on gliders, see section 8.3 of the OceanGliders Oxygen SOP (López-García et al., 2022).

2.5.4 Correct based on Turn-Around Cruise Data

Due to the known tendency of oxygen sensors to drift from their factory calibrations over time, oxygen data collected on turn-around cruises are critical for preparing final analysis-ready oxygen data. Section 2.4.1 provides background on oxygen sensor drift, and Section 1.5.4 summarizes recommendations for using OOI turn-around cruise data to validate and correct sensor data across all BGC sensors. Here we provide recommendations for how to use the data collected by OOI to correct for both pre-deployment storage drift and *in situ* drift during oxygen sensor deployments.

Quality control cruise data for comparison with sensor data

Before using OOI-provided oxygen data from turn-around cruises for sensor calibration, end users should apply their own quality control to these datasets. Discrete sample measurements from Winkler titrations can occasionally include anomalous data points, so we caution users to avoid pinning an entire sensor drift correction on a single discrete data point. Evaluation of discrete oxygen data in tandem with CTD cast SBE43 data from complete depth profiles will provide more robust validation.

CTD cast SBE43 data provided from OOI turn-around cruises has not been calibrated using the accompanying OOI-collected discrete samples. End users who wish to use the full depth profile data from the CTD cast SBE43 sensor should use the discrete bottle sample data to calibrate the CTD oxygen data prior to analysis (e.g, Uchida et al, 2010). We recommend this approach, as it will enable the best possible alignment between calibration samples and the deployed sensors.

Apply initial gain correction

To correct for storage drift between the time the sensor was factory- or lab-calibrated and the time of deployment, oxygen measurements at the beginning of a deployment must be corrected by comparison with concurrent reference data. The discrete bottle sample and CTD profile data collected on OOI turn-around cruises provide the reference data needed for this correction. End users should align the turn-around cruise oxygen data from discrete samples and/or discrete-sample-calibrated SBE43 data from CTD casts, which can be used to calculate an

initial gain correction factor. This gain correction should be applied to the sensor data throughout the full deployment.

For fixed-depth oxygen sensors, there will usually only be a single calibration data point from a nearby cast available to calculate the initial gain correction, whereas profiling assets will have the benefit of comparison across a full profile. The Worked Example at the end of this chapter (Section 2.6, Figure 2.6) provides an example of the gain correction for a sensor on a fixed-depth mooring, and Figure 2.2 provides an example calculating the initial gain correction for profiling data from a glider.

In applying initial gain corrections, end users should consider the potential impacts of spatial and temporal variability that may affect the match-up between the calibration cast data and the sensor to be calibrated. Directly comparing a discrete Winkler bottle sample with measurements from the various OOI platforms (moorings, gliders, etc.) may be challenging for platforms at depths/locations with strong vertical temperature/salinity gradients, since the Niskin bottle may not have sampled the exact same water mass as the platform. In some cases, it may be preferred to use the discrete sample oxygen data to calibrate the SBE43 oxygen sensor on the turn-around cruise CTD sensor package and then find the CTD cast data from the cruise that best align with the platform's observations (in terms of time, location, depth, temperature, and salinity properties). See the Nitrate Worked Example in Chapter 3 for an example of how to identify turn-around cruise samples and deployed sensor calibration data match-ups based on temperature rather than depth to facilitate water mass alignment.

Correct for sensor drift

In addition to applying the initial gain correction, oxygen sensor data also needs to be corrected for drift over the course of the deployment. End users will need to assess their particular dataset to determine which of the following correction mechanisms are applicable. If more than one of these approaches are possible for a given dataset, use of multiple methods of assessing drift and intercalibration among assets corrected with different methods will improve the robustness of the final drift correction.

- ➔ In cases where an OOI-deployed oxygen sensor successfully collects data for the full deployment period, it will still be collecting data at the time of the turn-around cruise where it is recovered. Turn-around cruise CTD cast and discrete sample data collected at the end of the deployment can be used to calculate a final gain correction following the same approach as the initial gain correction described above. If no additional data are available to assess drift over the course of the deployment, users should assume a linear sensor drift between the initial and final gain correction[1]. The Worked Example at the end of this chapter (Section 2.6) provides an example of this drift correction for two deployments of a fixed-depth moored oxygen optode.

[1] Note that drift is not always linear, especially for optodes that have not had an extensive burn-in period or been previously deployed. For instance, Wolf et al. 2018 find that an exponential drift correction best fits their deep isotherm reference data.

- ➔ Measurements from optodes deployed on deep profiling platforms that sample water masses that are assumed to have stable oxygen concentrations can also be used to address deployment drift. In this approach, oxygen concentrations at deep reference levels (a specific density or temperature value) are assumed to experience negligible rates of microbial respiration, resulting in relatively stable oxygen concentration such that trends in oxygen concentrations over the course of deployment are attributed to sensor drift. This approach has been applied to both OOI data (Palevsky and Nicholson, 2018) as well as BGC-Argo floats (Takeshita et al., 2013; Wolf et al., 2018). Figure 2.4 shows an example of oxygen data on a deep isotherm measured by the Apex Profiler Mooring at the OOI Irminger Sea Array. This observed drift can be corrected by assuming that the oxygen concentration measured on the stable deep isotherm at the time of initial gain corrections remains stable throughout each deployment. This deep isotherm reference approach is only possible if the profiling platform measures a deep isopycnal or isotherm with relatively stable oxygen concentration, and requires the end user to assess the validity of the assumption of oxygen stability to assess the uncertainty introduced by this approach.

- ➔ Starting in 2018, OOI engineers began configuring OOI gliders at the Irminger Array to enable oxygen air calibration measurements during deployments, supported by an ancillary NSF award (Nicholson, WHOI and Palevsky, BC). The use of oxygen air calibrations during surfacing intervals has been demonstrated to effectively correct for both storage and in situ drift and become widely implemented for BGC-Argo floats (Bittig and Kortzinger, 2015, Johnson et al. 2015, Bushinsky et al., 2016, Claustre et al. 2020), and has also been demonstrated to also provide effective calibration for gliders (Nicholson and Feen, 2017). Following the initial deployment of air-calibrating oxygen measurements on Irminger Sea gliders in 2018, OOI approved an Engineering Change Request in 2020 to apply this modified optode configuration across all gliders in the OOI program. Gliders modified to enable oxygen air calibration have subsequently been deployed regularly at the Irminger array and at the Pioneer array beginning in 2021, but this change has not yet been fully implemented across all OOI gliders.

While the first drift correction method is applicable to all OOI oxygen sensors that successfully complete their deployments, there are often cases where sensors and/or platforms do not last through the full deployment period prior to recovery and neither the deep isotherm or glider air calibration method are applicable, since these latter two only apply to a subset of profiling assets. In these cases, or cases where discrete sample measurements are not available (e.g., due to COVID-19 restrictions in spring-summer 2020) oxygen sensor drift may be possible to estimate by intercalibrating with another asset - for example, comparing data from a fixed-depth moored sensor with aligned profiles from an air-calibrating glider. If no such intercomparison is possible, the oxygen sensor data may not be correctable for drift, though can still be usable for applications less sensitive to sensor stability.

2.5.5 Comparisons among co-located data

Comparisons among co-located data provide valuable opportunities to improve constraints on oxygen sensor calibration. Section 1.5.5 provides context on four types of recommended intercomparisons, all of which are valuable for improving oxygen data calibration and validation:

- ➔ Comparison of simultaneously collected data from consecutively-deployed moorings at the same location. This can be especially useful in calculating/validating oxygen sensor gain and drift corrections.

- ➔ Comparison among multiple sensors measuring the oxygen at different locations within the array (Table 1.1, Figures A.1-A.6). Oxygen sensors are deployed on nearly every location on the OOI arrays that includes BGC sensors, and are on all profiling assets, providing especially rich opportunities for intercomparison among multiple oxygen sensors within the same array.

- ➔ Comparison of oxygen sensor data with co-located CTD temperature and salinity data, deployed alongside every oxygen sensor within the OOI program, as well as other BGC variables (see Figures A.1-A.6 for where oxygen sensors are deployed alongside other BGC sensor types). See chapters 3-5 for details on recommended end user data processing for other BGC sensors needed to facilitate such intercomparison.

- ➔ Comparison with external-to-OOI datasets in the vicinity of the OOI arrays. Cruises in the vicinity of OOI arrays during deployments have the potential to provide valuable supplementary calibration information. This can be especially useful for constraining oxygen sensor drift in cases where the sensor didn't continue to collect data all the way to the recovery turn-around cruise. Climatological datasets (e.g., the World Ocean Atlas) have also previously been used to estimate drift on stable deep isotherms (Takeshita et al. 2013) and could be used as a supplement to the turn-around cruise data in constraining and validating deep isotherm drift corrections at OOI Global arrays.

2.6 Worked Example

We provide here an example of how to apply the end user processing workflow described in Section 2.5 and summarized in the flowchart in Figure 2.5. This example prepares an oxygen product for scientific use from Level 2 (L2) oxygen data (DOXYGEN-L2, processed by OOI as shown in Figure 2.1a) from an Aanderaa optode on the near-surface instrument frame on the Oregon Offshore Surface Mooring of the Coastal Endurance Array across two consecutive deployments (Figure 2.6).

This example illustrates the importance of using discrete, shipboard-titrated Winkler oxygen measurements from turn-around cruises to correct for the storage (*ex situ*) drift and deployment (*in situ*) drift to prepare accurate final oxygen data. Turn-around cruises provide discrete sample oxygen measurements corresponding to the location of the oxygen sensor at both the beginning and end of each deployment, enabling these corrections. Comparison between the Level 2 oxygen data (colored lines for Deployments 8 and 9 in Figure 2.6) and the discrete samples from the turn-around cruises shows storage drift between the time of the factory calibration and beginning of each deployment as well as *in situ* drift over the course of the deployment, with the storage drift especially pronounced for Deployment 9 and *in situ* drift especially pronounced for Deployment 8.

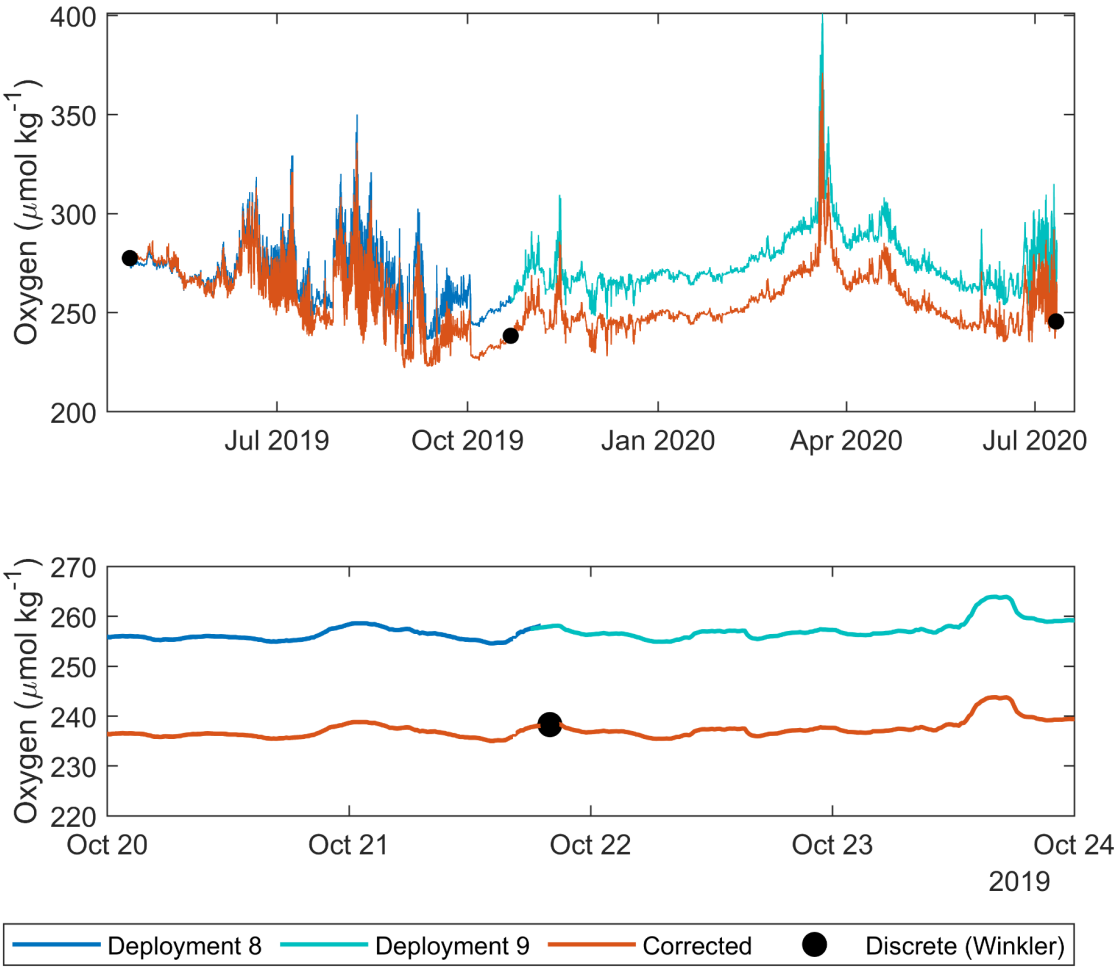


Figure 2.6: Dissolved oxygen worked example from the Coastal Endurance Array. Top panel shows dissolved oxygen data from two consecutive deployments (deployment 8 in blue and deployment 9 in teal) for the Oregon Offshore Surface Mooring. Black circles indicate oxygen concentrations from Winkler titrations made on discrete bottle samples taken near the near-surface instrument frame (7 m depth). Dissolved oxygen concentrations, corrected for storage drift and deployment drift, are shown in red. Bottom panel reveals a closer look at the transition from deployment 8 to deployment 9 during an October turn-around cruise.

Pseudo-Code

The pseudo-code provided below provides each step in the data processing pipeline for this worked example preparing the analysis-ready corrected oxygen dataset shown in Figure 2.6, with steps organized following the sequence given in Figure 2.5 and the text in Section 2.5. This pseudo-code is intended to support end users in developing their own data processing sequence following these recommended steps using any programming language or OOI data access method of their choice. The MATLAB code used to implement this example, including annotations to assist users in following the script, is provided as a supplementary resource, but is intended solely as a reference and not as a template for end user data processing code.

Assemble data:

```
Review available oxygen data on Data Explorer;
Review OOI HITL annotations for oxygen and co-located CTD data;

Download oxygen and co-located CTD data of interest;

Download co-occurring turn-around cruise discrete sample oxygen data
from shipboard Winkler titrations;
```

Inspect and evaluate the oxygen sensor data:

```
Plot time series of Level 2 oxygen and CTD data;
     Variables analyzed are user-ready output from Figure 2.1

Calculate median value of dissolved oxygen burst sampling from Level
2 oxygen data;

Plot time series of median oxygen data;

Check for evidence of biofouled or scratched sensors;
     OOI HITL Data Annotations report that UV light sensors were
     installed prior to these deployments, mitigating biofouling
```

Quality control discrete sample data from turn-around cruises:

```
Review OOI HITL Annotations for discrete sample data;

Plot cruise discrete sample oxygen values (from shipboard Winkler
titrations) and SBE43 oxygen values from CTD casts;

Evaluate Winkler-titrated discrete sample oxygen values to look for
outliers by comparing with SBE43 oxygen profiles;
```

Look for mis-match between bottle temperature/salinity data from shipboard CTD and temperature/salinity data from oxygen sensor co-located CTD;

Calculate gain corrections from turn-around cruise discrete samples:

Calculate instantaneous gain correction at times when discrete oxygen samples were collected on turn-around cruises:

Convert Winkler discrete sample oxygen values to µmol/kg:
Winkler (µmol/kg) = Winkler (ml/l)*44.661*1000/water density

T1 = time of turn-around cruise at beginning of sensor deployment
T2 = time of turn-around cruise from end of sensor deployment

$Gain_{T1}$ = Winkler value (µmol/kg) at T1 / Sensor value at T1 (µmol/kg)
$Gain_{T2}$ = Winkler value (µmol/kg) at T2 / Sensor value at T2 (µmol/kg)

Correct oxygen sensor data for gain and drift:

Calculate gain correction for each sensor data time point between T1 and T2 based on linear drift between $Gain_{T1}$ and $Gain_{T2}$:

Tn = time point between T1 and T2
$Slope_{T1/T2}$ = ($Gain_{T2}$ - $Gain_{T1}$) / (T2 - T1)

$Gain_{Tn}$ = $Gain_{T1}$ + $Slope_{T1/T2}$*Tn

Apply corrections to Level 2 oxygen sensor data:
$Oxygen_{L2}$ = time series of uncorrected Level 2 oxygen data (µmol/kg)
$Oxygen_{CORR}$ = corrected oxygen data (µmol/kg)

$Oxygen_{CORR}$ = $Oxygen_{L2}$ * $Gain_{Tn}$

Plot analysis-ready oxygen data

Chapter 3: Nitrate

3.1 Introduction to OOI Nitrate Sensors

The OOI nitrate data is collected by the SUNA V2 and ISUS instruments. The ISUS and SUNA V2 instruments are functionally the same sensor and the SUNA V2 has replaced the ISUS in OOI, so for readability we will refer only to the SUNA V2 in this document. See Table 3.1 for more specifics on deployment information. The SUNA V2 is a commercial instrument that uses a 1 cm pathlength UV spectrometer with a deuterium lamp to collect spectral absorption data which is then inverted to generate an *in situ* nitrate concentration. The technique is generally most sensitive to interferences from bromine absorption which is a function of seawater temperature and salinity, dissolved organic matter absorption and attenuation by particulates.

In order to be useful for scientific or management applications, nitrate data from extended autonomous deployments should be carefully evaluated for accuracy, which may be impacted by calibration issues, sensor drift, biofouling or chemical interference. With proper corrections applied, the data can be high quality.

Descriptions of the instrument and the techniques are in the literature (Finch et al., 1998, Johnson and Colletti, 2002, SUNA V2 users manual, 2018). ISUS sensors were used on most fixed (moored) OOI platforms, until they started being phased out and replaced by the SUNA V2 in 2017. SUNAs were always used in profiling applications, and are also used on some mobile platforms (gliders, AUVs). Table 3.1 outlines which sensor types are deployed at which OOI locations.

Raw variables measured

The SUNA measures light absorption between 190 - 370 nm using a photodiode array. The absorption spectra is a result of the presence of any chemical species in seawater that absorb UV light including: bromide, nitrate, nitrite, dissolved organic matter, and hydrogen sulfide. Johnson and Coletti (2002) described the spectral deconvolution method for calculating nitrate concentration using known absorption coefficients for bromide and correcting each wavelength for interference from dissolved organic matter (DOM) and particulates. Sakamoto et al. (2009) refined the calibration procedure to account for the temperature dependence of bromide absorption. **Note that interference from nitrite or hydrogen sulfide is presumed small and therefore not accounted for in the processing of the nitrate data by OOI.**

Final calculated target product

Data sourced from the OOI data sets as nitrate concentration has been corrected by OOI using the Temperature Compensated Salinity Subtracted (TCSS) algorithm from Sakamoto et al. (2009), which uses data from the co-located CTD sensor. The raw spectral data necessary for this correction are available on OOI's raw data server. Future advancements or new developments in the inversion technique may lead to re-calculation of the nitrate data from the

spectral data collected by the SUNA, so OOI nitrate users are encouraged to review current literature for best practices. In all events the user should record and report all post-processing steps.

Table 3.1. OOI nitrate sensor array and platform deployment locations, and internal-to-OOI instrument class-series. For the class-series, all nitrate sensors are designated NUTNR and are Sea-Bird/Satlantic SUNA V2 and ISUS instruments. The last letter of the class-series is an internal classification by OOI and represents the series (determined by specifications related to the sampling rate, deployment duration, deployment depth, etc.). For a summary of where all BGC sensors are deployed across the full OOI program, see Table 1.1. For deployment locations within each array, see Figures A.1-A.6.

Array	Platforms	Sensors	OOI Class-Series
Global Argentine Basin Array (suspended 2018)	Global Profiling Glider	SUNA (2016 - 2017)	NUTNR-M
	Near Surface Instrument Frame (NSIF) & Subsurface Buoy	ISUS (Mar 2015 - Jan 2018)	NUTNR-B
Global Irminger Sea Array	Global Profiling Glider	SUNA (2014 - present)	NUTNR-M
	Near Surface Instrument Frame (NSIF) & Subsurface Buoy	ISUS (Sep 2014 - Jun 2018) SUNA V2 (Jun 2018 - present)	NUTNR-B
Global Southern Ocean Array (suspended 2020)	Global Profiling Glider	SUNA (2015 - 2016)	NUTNR-M
	Near Surface Instrument Frame (NSIF) & Subsurface Buoy	ISUS (Feb 2015 - Dec 2018) SUNA V2 (Dec 2018 - Jan 2020)	NUTNR-B
Global Station Papa Array	Global Profiling Glider	SUNA V2 (2013 - present)	NUTNR-M
Regional Cabled Array	Shallow Profiler Mooring	Deep SUNA (2014 - present)	NUTNR-A
Coastal Endurance Array	Surface Piercing Profiler	SUNA V2 (2014 - present)	NUTNR-J
	Near Surface Instrument Frame (NSIF)	ISUS (2014 - 2018) SUNA V2 (2018 - present)	NUTNR-B
Coastal Pioneer Array	Coastal AUV	SUNA (2015 - present)	NUTNR-N
	Surface Piercing Profiler (suspended 2016)	SUNA V2 (2014 - 2016)	NUTNR-J
	Coastal Profiling Glider	SUNA (2015 - present)	NUTNR-M
	Near Surface Instrument Frame (NSIF)	ISUS (2013 - Mar 2018) SUNA V2 (Mar 2018 - present)	NUTNR-B

Expected accuracy and precision

Using the methodology described above, nitrate concentration can be determined with an accuracy of < 2 µM and precision (and limit of detection) of 0.3 µM over the range of concentrations typical for ocean ecosystems (0-50 µM). Offsets in SUNA data are common, so accuracy of the data is often a function of the availability of *in situ* calibration data sets.

Sampling rate protocol

The SUNA can take a measurement every 1 second, and can be run in either continuous mode, where measurements are made continuously, or in polled mode, where the sensor is running on an intermittent schedule rather than continuously. SUNA measurements are split into either light or dark measurements. Light values are measurements that are made with the lamp shutter open, whereas dark values are measurements made with the lamp shutter closed. These dark measurements are primarily used to space out the light measurements, prevent the lamp from heating up too much, and monitor abrupt changes in the instrument. Dark measurements are not regularly used in data post-processing, but are available in the recovered instrument data stream on the OOI raw data server if desired by the user. The SUNA can be configured to measure light and dark values on varying schedules. Table 3.2 provides details on the SUNA sampling protocols for different OOI platforms and arrays.

Table 3.2. SUNA configuration and sampling setup

Configuration	Surface mooring SUNA	Coastal glider SUNA	Global glider SUNA	AUV SUNA	Cabled Array Shallow Profiler
Sampling mode	Continuous	Polled	Polled	Polled	Polled
Power/sampling frequency	On for 3 min, every 15 min	On every surface dive & 7th dive of segment down to 200m or bottom of dive	On every surface dive down to 200m or bottom of dive	On for duration of mission	On twice per day (local noon, local midnight) for the profiler upcast
Light averages	3	1	1	1	1
Dark averages	100	1	1	1	1
Light samples	1	4	4	1	1
Dark samples	1	1	1	11	36
Output type	Full ascii (output and log)	Reduced binary (output) Full binary (log)*	Reduced binary (output) Full binary (log)*	Full binary (output) Full ascii (log)	Full ascii (output and log)

*Glider SUNAs were configured to log data in reduced binary prior to 2021

For most OOI surface moorings the SUNA is powered up every 15 minutes, runs continuously for 3 minutes and reports approximately 4 to 6 sets of average values during that time interval. One light average is reported for every dark average reported, with each light average made up of 3 measurements with the lamp shutter open, and each dark average made up of 100 measurements with the lamp shutter closed. On the OOI Cabled Array, the SUNAs on the Shallow Profiler Moorings are configured somewhat differently, with sampling done twice per day (local noon, local midnight) and only on the upcast, comprising 36 s of dark measurements and 1 s of light measurements. This results in 1-3 measurements of nitrate, depending on the spectrometer and electronic characteristics, as well as the actual concentrations and background signals every 40s (due to the added processing time), which results in a profiling resolution of 2 m at a profiling rate of 5 cm/sec.

SUNAs mounted on gliders and AUVs operated by OOI are operated in polled mode (i.e., the sensor is not running continuously), meaning that the vehicles determine when the sensors are turned on and sampling. For global profiling gliders, the gliders turn the SUNAs on for the first dive of each segment on the descent to 200 m. Each global glider segment (interval between surfacings) consists of two yos, where a yo includes both a down and up profile (but not necessarily all the way up to the surface, which would be the end of a segment). Coastal gliders operate similarly, but due to the much shallower water depth and resulting dive pattern, the SUNA is configured to sample for the first dive of the first yo per segment (to dive depth or to 200m, whichever comes first). In addition, it samples every seventh yo until that segment ends, at which point the cycle restarts with the first dive of the first yo. Glider SUNAs are configured to output 4 light measurements for every dark measurement and no averaging is applied. AUV SUNAs are also set to polled mode, but are turned on for the duration of the vehicle mission. They output 1 light measurement for every 11 dark measurements and no averaging is applied.

3.2 OOI standard practices for SUNA deployment and calibration

3.2.1 Preparation for deployment

In practice, OOI deploys calibrated instruments from the manufacturer with minimal modifications. Vendor calibrations are performed before every cruise for moored instruments and on a less frequent “as-needed” schedule for instruments on mobile platforms. Instrument factory calibration information is available on the Alfresco server and calibration files are posted on receipt of instruments returning from vendor servicing. The path to find factory calibration information for sensors deployed at one of the coastal or global arrays, for example is:

```
OOI > Instrument & Platform Documents > Calibration and Repair
> Coastal-Global Arrays > NUTNR
```

All nitrate sensor vendor calibrations are stored in that same folder structure (for both ISUS and SUNA V2 sensors) and specific calibrations can be found by searching for a specific sensor serial number and date. All instruments deployed to the field should have their most recent

factory calibrations available on Alfresco (see Table 3.3 below for numbering scheme for all test documents).

In addition, a reference update to the vendor calibration is performed by OOI prior to each deployment (generally in the week leading up to deployment) and uses deionized water (DI) as the nitrate free endmember (SUNA V2 users manual, 2018). This reference update procedure provides an offset to the vendor calibration and helps ensure that the instrument reads 0 ± 2 µM in DI water.

To perform this, the instrument sample area is thoroughly cleaned before being enclosed with parafilm and filled with DI water. The instrument is connected to a computer using Sea-Bird's UCI program and nitrate values are recorded. Regardless of the magnitude of nitrate readings, a reference update is performed. This step creates a new calibration file that applies an offset to the factory calibration in order to bring the nitrate readings within 0 ± 2 µM. A reference check is performed a second time with the updated reference values to check for short term stability, i.e., that the updated reference file produces a nitrate reading within 0 ± 2 µM using the same target water. A second reference check is performed after drying, recleaning, and re-inclosing the sample area with new parafilm and DI water. This additional step helps ensure that the reference update was not adversely affected by the presence of any debris or bubbles in the sample area. During the procedure, the lamp output is checked by evaluation of the maximum light counts recorded. This step is used to evaluate that the light output is sufficient to produce data over the expected duration of the deployment. These procedures are performed in a cold room (4-10°C for WHOI and OSU sensors, 12°C for Cabled Array sensors) in order to mimic deployment temperatures, whereas vendor calibrations are done at room temperature. The calibration temperature is captured in the results sheet for that reference update (the 3305-00327-xxxxx or 3305-00527-xxxxx document; see Table 3.3).

3.2.2 Information collected upon recovery

After recovery of the platform, data are downloaded directly from each instrument. Notes are made about any significant biofouling or physical observations (corroded pins, worn cables, etc.). Any clock offset is also noted, as compared against a computer synced to the shipboard timeserver. These recovered status notes and observations are recorded in the data download checklists found on the raw data server under each mooring site and recovery folder. For example, the data recovery checklist for the Pioneer Central Surface mooring recovered in spring 2021 can be found from the main file tree structure by navigating to:

> `CP01CNSM > R00013 > Data_Recovery_Checklist_CP01CNSM-00013.xlsx`
> (https://rawdata.oceanobservatories.org/files/CP01CNSM/R00013/)

A post-deployment reference update is also performed utilizing the same procedure as done pre-deployment. This creates a new calibration file after applying an offset to bring the instrument to read 0 ± 2 µM in DI water. This helps to determine drift associated with changes in the output of the deuterium lamp over the course of the deployment.

Table 3.3. OOI Test Documentation Numbering Scheme for nitrate sensors on the Coastal and Global Arrays. The first 9 numbers are a fixed indication of the test and sensor type, and xxxxx refers to the sequential document number for a specific test record. These documents can all be accessed on OOI's Alfresco server using the path: `OOI > Instrument & Platform Documents > Test Documents > Instruments > Coastal-Global Arrays > NUTNR`

Document Number	Description
3305-00100-*xxxxx*	Instrument incoming inspection. Brief record of instrument condition and associated calibration files after receipt from vendor
3305-00108-*xxxxx*	ISUS quality conformance test (QCT) record. Basic functionality check performed after receipt from vendor
3305-00127-*xxxxx*	SUNA quality conformance test (QCT) record. Basic functionality check performed after receipt from vendor
3305-00308-*xxxxx*	ISUS pre-deployment reference update. Reference update to vendor calibration performed in the week or so leading up to a deployment
3305-00327-*xxxxx*	SUNA pre-deployment reference update. Reference update to vendor calibration performed in the week or so leading up to a deployment
3305-00508-*xxxxx*	ISUS post-deployment reference update. Reference update to vendor calibration performed in the week or so after recovery
3305-00527-*xxxxx*	SUNA post-deployment reference update. Reference update to vendor calibration performed in the week or so after recovery

Table 3.3 summarizes the numbering scheme for pre- and post-deployment for nitrate sensors on the Coastal and Global Arrays, all of which are publicly available on the Alfresco server. The same documents are also available for sensors deployed on the Cabled Array, but as of writing, are not all available on the Alfresco server yet, and do not follow the same numbering scheme. Users of Cabled Array nitrate data are recommended to check the equivalent repository on Alfresco for Cabled Array documents:

```
OOI > Instrument & Platform Documents > Test Documents >
Instruments > Cabled Array
```

If users are unable to find the relevant documents, they should submit a request on the OOI Discourse Forum specifying the instrument and type of document they're looking for.

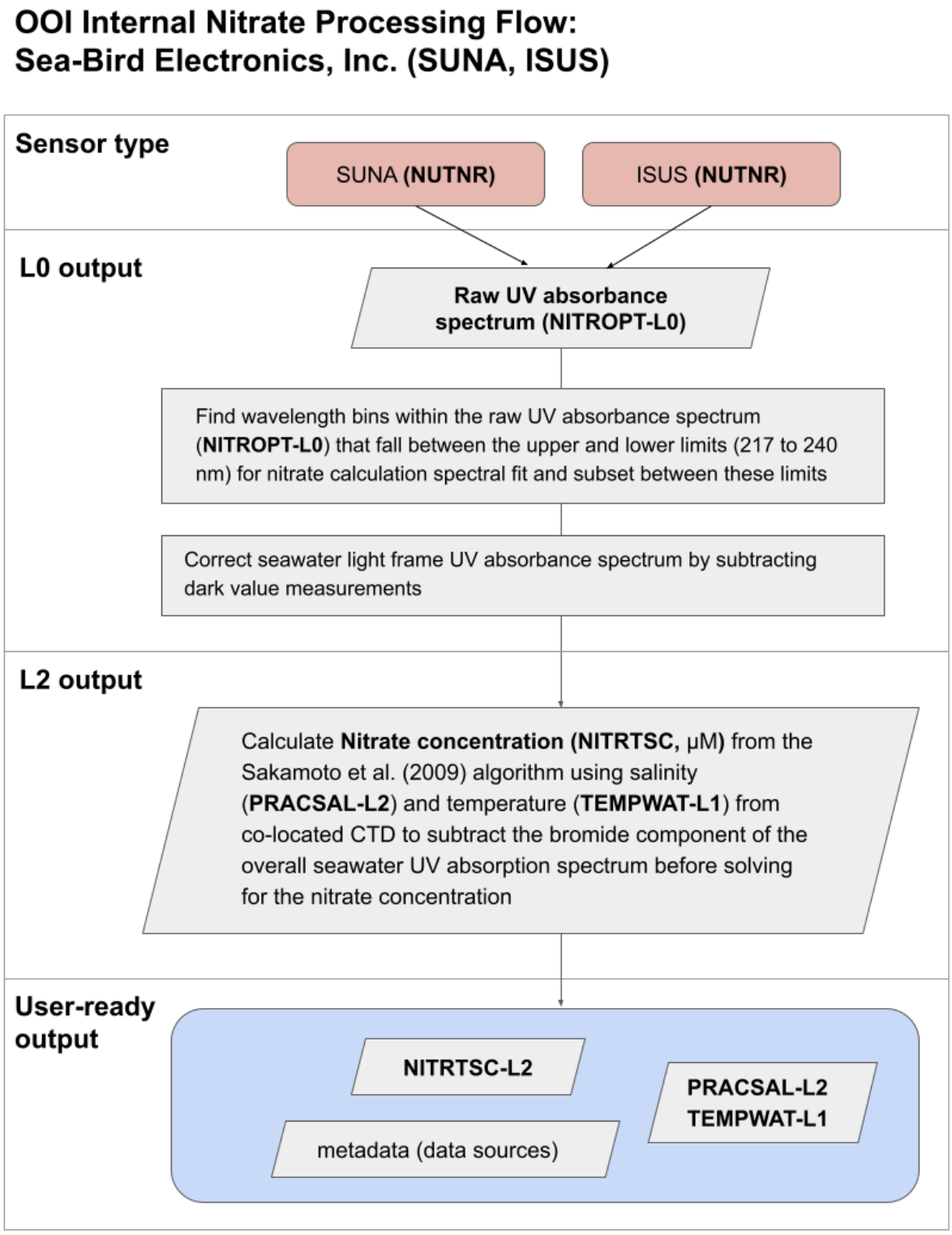


Figure 3.1: Summary of the internal-to-OOI processing workflow for nitrate data. The user-ready data products provided by OOI are not yet analysis-ready nitrate data and require further processing by the end user to evaluate, clean, and apply corrections to the data. The user-ready output from OOI's internal data processing in this figure is the starting point sensor data for the recommended end user data processing outlined in Section 3.5 and summarized in the flowchart in Figure 3.2

3.3 Internal to OOI Data Processing Workflow

Figure 3.1 summarizes OOI's internal processing workflow for data from both ISUS and SUNA nitrate sensors, with data from both sensors processed in the same way. Nitrate data are transformed from an unprocessed, Level 0 (L0) raw UV absorbance data product to a OOI-provided Level 2 (L2) data product in µM that has been temperature and salinity corrected. The OOI data processing workflow to transform raw nitrate sensor data to the L2 nitrate data product uses the Temperature Compensated Salinity Subtracted (TCSS) algorithm from Sakomoto et al., 2009. Briefly, the Sakamoto et al. (2009) algorithm uses L1 and L2 data products from the co-located CTD instruments to subtract the bromide component of the overall seawater UV absorption spectrum before solving for the nitrate concentration. The Python script used to process nitrate data by OOI is available on GitHub here.

3.4 Common Data Quality Issues

Nitrate sensors deployed by OOI are subject to drift from factory-calibrated values, as well as interference from other dissolved compounds (e.g., CDOM). Both of these can result in a decrease in the photon flux at the sensor's detector which does not stem from an increase in nitrate concentration in the target water relative to the calibration blank. **It is therefore critical to check and correct for sensor drift and/or interference prior to any quantitative interpretation of nitrate data.**

For moored instruments it is common to see an increase in nitrate concentration over time while other water mass characteristics, e.g., temperature and salinity, are relatively invariant. The pre- and post- deployment discrete sample nitrate concentrations are used to characterize this change over time. For example, the initial *in situ* instrument concentration should be within ± 2 µM of the pre-deployment discrete sample concentration, whereas the post-deployment discrete sample concentration may be much lower than the instrument recorded concentration. This is generally most apparent in near-surface data where nitrate concentrations are near zero and at depth where temporal, spatial and vertical gradients are small. Intermediate depth comparisons and variability over time may be more difficult to interpret due to the additional variability introduced by internal waves, changes in mixed layer depth, biological uptake and regeneration.

Best practices for correcting *in situ* nitrate sensor data sets from moorings and autonomous platforms evolve over time, so we encourage OOI nitrate data users to check with subject matter experts and recent literature sources. Our intent here is to present a non-exhaustive framework for understanding common data anomalies and corrections that can be applied when using the OOI data sets.

3.4.1 Biofouling

Biofouling due to biofilm development and/or soft or hard fouling will decrease light output from the lamp side, and light collection on the detector side, resulting in a decrease in UV photon flux relative to the calibration values. Biofilms are almost always present at some time past

deployment. For biofouling, it is reasonable to presume that biofouling increases with time and that the rate of increase is relatively smooth. In contrast, interferences from DOM and particulate attenuation are generally associated with changes in water mass, water column biological activity, advection and turbulent energy and are more episodic within a given time series. The exception from relatively smooth and continuous biofouling impact on the data is when stringers and fronds attached to the mooring or instrument enter the sensing volume of the instrument. This is less common for the nitrate sensors than open faced optical instruments due to the short pathlength of the nitrate sensors. Note that the ISUS instruments that were deployed with a copper guard prevented this, but OOI engineers found that design also tended to collect sediment at certain locations due to restricted flow.

3.4.2 Drift Due to Lamp Output

The data produced by the SUNA nitrate sensors is subject to uncertainty due to decay in the lamp output over time. Stability of the lamp is an issue during deployment and transportation. Calibration of the instrument with nitrate free water effectively compensates for output changes since the last calibration, but *in situ* data should be evaluated relative to the time since the last calibration and the duty cycle of the instrument. Because the decrease in light intensity of the lamp is also a relatively continuous function with time, using the pre- and post-deployment bottle samples to adjust for changes over time effectively accounts for both biofouling and lamp degradation, as described above in Section 3.4.1.

3.4.3 Dissolved Organic Matter and Other Interferences

DOM absorbs UV light and the absorption coefficient is inversely related to the wavelength of the light. The SUNA internally compensates for this effect by using the absorbance at wavelengths greater than the wavelengths used in the NO_3^- determination to subtract a baseline absorbance due to DOM. However, because the shape of the absorbance due to DOM is variable in the wavelength range used to determine the nitrate concentration, interference from DOM is often underestimated. Further complicating the measurement is that changes in the DOM pool are usually associated with changes in water mass characteristics, such that changes in DOM are associated with a change in temperature as well as a potential change in the nitrate concentration itself. This is particularly a problem in areas of riverine influence as higher DOM concentrations may be associated with higher nitrate concentrations.

Other than DOM, bromide, hydrogen sulfide (H_2S) and nitrite (NO_2^-) are other known sources of interference in the SUNA data. H_2S and NO_2^- concentrations are generally negligible in most of the ocean but are particularly high in low-oxygen environments. The spectral attenuation characteristics for these chemical species can be derived in the laboratory and used to correct the observed spectra, as for bromide. Nitrate concentrations can then be estimated after correcting the observed spectra for interfering chemical species (e.g. Meyer et al., 2018).

3.4.4 Variable Interferences

The UV absorption methodology for measuring nitrate in the ocean works best when the range of variance associated with the absorption of other substances can be reliably constrained as small relative to the nitrate variance in time and space (Johnson et al., 2013). Correcting the data stream for linear offsets is practical because nitrate concentrations in the surface ocean often approach a near zero minimum within a profile, and the time variability of nitrate concentration with depth can be assumed to be near zero (Johnson et al., 2013). This allows a linear adjustment to the dataset by correcting the near-surface mixed layer data to a presumed minimum concentration. Time and space variance can then be estimated from data at depth and as a check of the assumption of the surface concentration. This type of baseline adjustment can also be a first order methodology to adjust for the use of different instruments over time.

Where the nitrate concentration and/or other absorbing components have high variability then it helps to scale the potential impact on the SUNA-derived nitrate data from known sources of variability. The OOI data set includes data that can be used to constrain anomalous data. We can functionally constrain NO_3^- variability between the actual *in situ* NO_3^- concentration and other factors by using co-located UV fluorometer and backscatter sensors (see Chapter 5 for further information on the OOI bio-optical sensors) as well as spectral data from the SUNA that is not used in the calculation of NO_3^-. To correct for variable interference from DOM and/or particulate matter, the absorbance spectra can be characterized at the deployment site, then subtracted from the measured absorbance spectra in post-processing as described in Johnson and Colletti (2002). Generally, this correction only needs to be applied in regions with high productivity or sediment loads and so will be most relevant for some of the SUNAs deployed on moorings and gliders at the Coastal Arrays.

3.5 Recommendations for End-User Data Processing

The end user nitrate data processing flowchart (Figure 3.2) summarizes additional data processing that must be performed by the end user to evaluate, clean, and apply corrections to the OOI-provided nitrate data. These steps are essential to prepare the OOI-provided nitrate data for scientific applications and analyses, especially involving quantitative interpretation.

Chapter 1 of this document provides an overview of QA/QC procedures recommended for all OOI biogeochemical sensors (Section 1.5), and provides a high-level walk-through and context for each step in the recommended end user nitrate data processing summarized in Figure 3.2. We recommend that users intending to work with OOI nitrate data use the flowchart in Figure 3.2 and instructions in Section 1.5 as a starting point and reference for each data processing step. Here, we walk through each of the steps outlined in Figure 3.2, which synthesize the approaches end users can take to correct OOI data products for the most common nitrate sensor-specific behaviors and data quality issues described in Section 3.4.

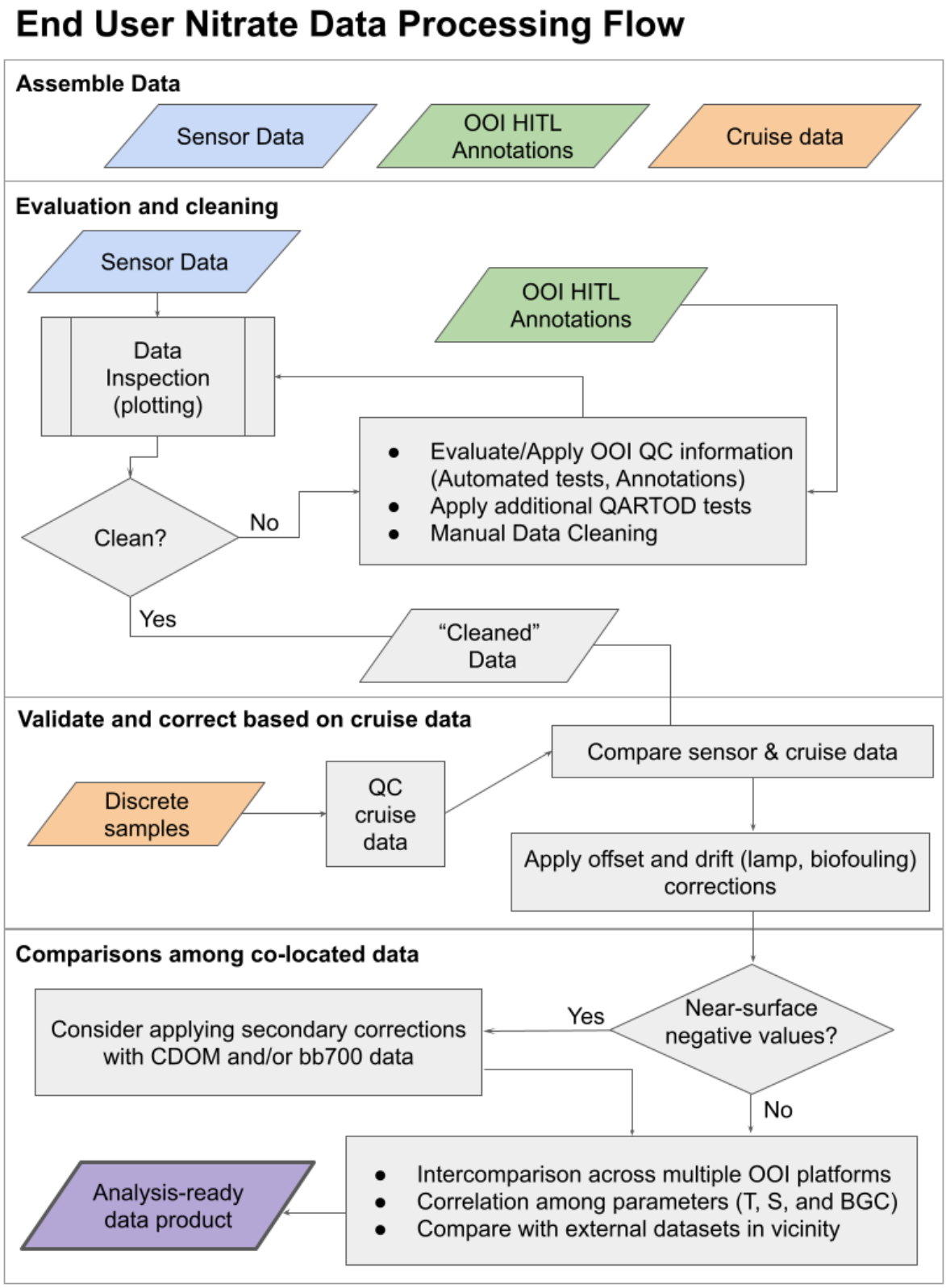


Figure 3.2. Summary of recommended end-user quality control and data processing steps for OOI nitrate sensor data. These steps will allow the user to evaluate, clean, and apply corrections to the sensor data provided by OOI (see Figure 3.1) to prepare analysis-ready nitrate data products.

3.5.1 Assemble Data

Users will need to assemble nitrate sensor data, accompanying OOI HITL annotations, and corresponding turn-around cruise data to use in preparing their final analysis-ready nitrate data. See Section 1.5.1 for details of each of these components.

Ancillary CTD data (pressure, temperature, salinity) co-located with the nitrate sensor to be analyzed will in most cases already be merged with the nitrate data stream by OOI since these data are used in preparing Level 2 (L2) nitrate data. Unless issues are identified when reviewing these ancillary CTD data, the OOI-provided L2 Nitrate product should be used for further data treatment.

Nitrate sensor data are validated using data from water samples collected from Niskin bottles on CTD casts during OOI turn-around cruises. Discrete samples for sensor data validation should be collected during CTD casts just before sensor deployment and immediately after recovery, to account for possible offset and drift effects described above. Discrete water samples are analyzed for nitrate either onboard or at offshore laboratories using GO-SHIP protocols (Becker et al., 2020). It is necessary for end users to assemble the discrete bottle data to include in their analysis.

3.5.2 Evaluation and cleaning & sensor-specific quality control

The initial step in recommended end user OOI nitrate sensor data processing is to prepare a “cleaned” dataset, applying both automated and human-in-the-loop (HITL) QA/QC to evaluate the data and identify points that may need to be filtered or removed. Section 1.5.2 summarizes the recommended steps for OOI BGC sensor end user data evaluation and cleaning. Here we provide additional context on the application of these steps specifically for OOI nitrate data:

1. Evaluate and apply OOI-provided HITL Data Annotations. Annotations by the OOI data team identify many platform-wide issues users may need to be aware of (e.g., power failure-caused data gaps, etc.), as well as nitrate sensor-specific issues.

2 & 3. Apply QA/QC based on published QARTOD recommendations. As of December 2022, automated data quality flags for nitrate are planned but not yet in development within OOI (see Table 1.2). Users are encouraged to check for and apply OOI’s automated nitrate data quality flags once they become available. If OOI implementation of these automated tests is still not available, end users are strongly encouraged to implement their own version of the tests listed in Table 1.2 and/or to manually inspect data to determine whether any of these issues are present in their dataset.

Nitrate data end users are particularly recommended to implement a spike test, as there should not be any spikes in observed NO_3^- concentrations. Spikes in the data may be caused by abrupt water mass changes, evident in temperature, salinity and density, as well as particulates which can be inferred from bio-optical backscatter data (discussed in Chapter 5). For coastal data sets, frontal dynamics with resuspension of bottom sediments can result in suspect data that

may or may not be correctable, but should be largely diagnosable by looking at other data as described above.

4. Manually inspect data to identify and address commonly-known issues. Nitrate data users should check especially for evidence of biofouled sensors (see Section 3.4.1) or lamp output drift (see Section 3.4.2). While some examples of this issue will be flagged by OOI-provided Data Annotations, users should carefully inspect all data in preparation for their own analysis. Lamp drift and biofouling both tend to result in a reduction in lamp output over time. OOI L2 nitrate concentration data sets may also contain negative concentrations and/or other unlikely values and should be evaluated by the user for suitability. In many cases, these issues can be corrected using time varying and constant offsets based on discrete sample data, described in more detail in Section 3.5.3, below.

3.5.3 Correct based on Turn-Around Cruise Data

Data from pre- and post-deployment bottle samples can be used to correct for sensor drift issues caused by biofouling and changes in lamp output over time, as well as offsets due to interferences. Section 3.4 provides background on common SUNA data quality issues, and Section 1.5.4 summarizes recommendations for using OOI turn-around cruise data to validate and correct sensor data across all BGC sensors. Here we provide recommendations for how to use the data collected by OOI to correct for both sensor drift and offsets observed during nitrate deployments.

Quality control cruise data for comparison with sensor data

Before using OOI-provided nitrate data from turn-around cruises for sensor calibration, end users should apply their own quality control to these datasets. Discrete sample measurements can occasionally include anomalous data points, so users should evaluate the discrete nitrate sample data prior to using it for sensor data calibration and correction.

Biofouling and Lamp Drift Correction

A single drift correction that will effectively account for both lamp drift and some degree of biofouling can be applied to nitrate sensor data. Initially, biofouling will result in a gradual decrease over time of the light that the sensor can detect. Lamp drift will also result in a decrease in the light intensity of the lamp that is a relatively continuous function with time, so using the pre- and post-deployment bottle samples to adjust for changes over time effectively accounts for both biofouling and lamp degradation. The simplest method of correction for biofouling and/or lamp drift is to apply a linear correction with respect to time from the start of the mooring deployment to the end:

$$\Delta\,NO_3/\text{day} = ((NO_{3,\text{bottle}} - NO_{3,\text{Suna}})_{\text{deployment}} - (NO_{3,\text{bottle}} - NO_{3,\text{Suna}})_{\text{recovery}})/(\text{Recovery date} - \text{Deployment date})$$

and:

$$NO_{3,\text{adj}}(t) = NO_{3,\text{recorded}}(t) - \Delta\,NO_3/\text{day} * t$$

where deployment and recovery represent the paired bottle and instrument data from the pre-deployment and post-recovery profiles, and the mooring start and end samples respectively, $NO_{3,adj}(t)$ is the adjusted data at a given time (t) and $NO_{3,recorded}(t)$ is the *in situ* sensor data recorded at time (t).

For instruments for which pre- and post-deployment bottle data are not available, the pre- and post-deployment calibrations can be used to derive the time rate of change in nitrate concentration due to lamp degradation using the same relationship.

$$\Delta\, NO_3/\text{day} = (NO_{3,\text{DI deployment}} - NO_{3,\text{DI recovery}})/(\text{Pre-deployment calibration check date} - \text{Post-deployment calibration check date})$$

and:

$$NO_{3,adj}(t) = NO_{3,recorded}(t) - \Delta\, NO_3/\text{day} * t$$

Note that neither of these methods will necessarily correct a heavily biofouled instrument if the biofouling growth rate is exponential, which may occur towards the end of a deployment. If biofouling is heavy, the lamp fails or other sources of interference cause the detected light to fall below a threshold value, and the instrument will return a nitrate value of 0.

Offset corrections

Although using the drift calibration methods outlined above should also correct for any offsets between observed nitrate sensor values and nitrate for discrete samples, it is also helpful to assess the magnitude of offsets between the sensor and discrete sample data. Offsets can be most easily assessed in near-surface data where nitrate values are frequently undetectable (i.e., below the detection limit of the sensor), and in deep water data where water mass variability and biological influences are minimal and nitrate concentrations can be presumed constant along density surfaces with time (see Figure 2.3 for an example of how deep water data can be used to identify sensor offset and drift, using deep oxygen data). Offsets can be assessed by examining various combinations of data types, but are most robust when comparing the nitrate sensor data relative to discrete water samples. Users may also assess offset by comparing the nitrate sensor data to data from overlapping instruments (e.g., from a mooring turn-around) and nearest neighbor instruments (e.g., from overlapping glider missions and/or nearby moorings). Offset adjustments are generally applied as step functions between data sets.

Secondary corrections

While the above drift and offset corrections are likely to account for most sources of error in the nitrate sensor data, there may still be information in the data, post-correction, that does not match with expected values or variances. This could be specific to a particular instrument, or a particular section of a time series, and in that case one might simply ignore the data and move on with the rest of the large data set. However, as outlined in Sections 3.4.3 and 3.4.4 above, certain compounds may be present that interfere with the sensor's method of nitrate estimation. Although we do not provide a specific method for dealing with such interferences here,

end-users should investigate the data further to understand more about the instrument, the methodology and the ocean conditions within which the sensor is deployed.

3.5.4 Comparisons among co-located data

Comparisons among co-located data provide valuable opportunities to improve constraints on nitrogen sensor calibration. Section 1.5.5 provides context on four types of recommended intercomparisons, all of which are valuable for improving nitrate data calibration and validation:

- ➔ Comparison of simultaneously collected data from consecutively-deployed moorings at the same location. This can be especially useful in calculating/validating nitrate sensor drift and offset corrections.

- ➔ Comparison among multiple sensors measuring nitrate at different locations within the array (Table 1.1, Figures A.1-A.6). Nitrate sensors are deployed widely on the OOI arrays including on profiling assets and gliders, allowing for intercomparison between multiple nitrate sensors within the same array.

- ➔ Comparison of nitrate sensor data with temperature and density from co-located CTD data, as well as other BGC variables (see Figures A.1-A.6 for where nitrate sensors are deployed alongside other BGC sensor types). See chapters 3-5 for details on recommended end user data processing for other BGC sensors needed to facilitate such intercomparison.

- ➔ Comparison with external-to-OOI datasets in the vicinity of the OOI arrays, such as the World Ocean Atlas or GO-SHIP repeat hydrography cruises.

3.6 Worked Example

We provide here an example of how to apply the end user processing workflow described in Section 3.5 and summarized in the flowchart in Figure 3.2. This example validates for scientific use Level 2 (L2) nitrate data (NITRTSC-L2, processed by OOI as shown in Figure 3.1) from the Deep SUNA on the Coastal Endurance Oregon Offshore Cabled Shallow Profiler Mooring, from a deployment that collected data during July and August of 2020 (see Figure A.4 for Endurance Array map schematic).

The example illustrates a comparison of the OOI-provided Level 2 *in situ* nitrate data from the profiler mooring with discrete bottle samples collected during the deployment period. The discrete sample data selected for comparison are from the turn-around cruise cast most closely co-located with the profiler mooring subsequent to its deployment (500 m away), which are compared to the profiler data collected closest in time to the cast where discrete samples were collected (2 hours difference). The data comparison between the profiler SUNA L2 data and the discrete samples suggests that the goodness of fit is very reasonable (Figure 3.3). The deepest profiler values from the SUNA are slightly low relative to the bottle samples but well within the accuracy specification for the nitrate sensor (± 2 µM). The agreement with near surface values

(at 10 m and 20 m) is not as good, with a difference of about 3 µM at 20 m and almost 8 µM at 10 m. To address possible mis-matches due to spatial and temporal variability between the profiler and cast where discrete samples were collected, particularly the surface, the data are also compared with respect to temperature rather than depth (Figure 3.3, right hand panel). Temperature serves as a proxy for density, as temperature usually decreases relatively monotonically with depth, and has the benefit of being provided alongside the data download for the profiler, and does not require end user validation/calibration prior to use (which is needed for salinity). Plotting nitrate against temperature rather than depth is therefore recommended for identifying match-ups between discrete samples and sensor data from profiling assets, which is especially true in areas with steep gradients in nitrate (and/or other variables), where internal waves are frequent, and where water masses are highly variable.

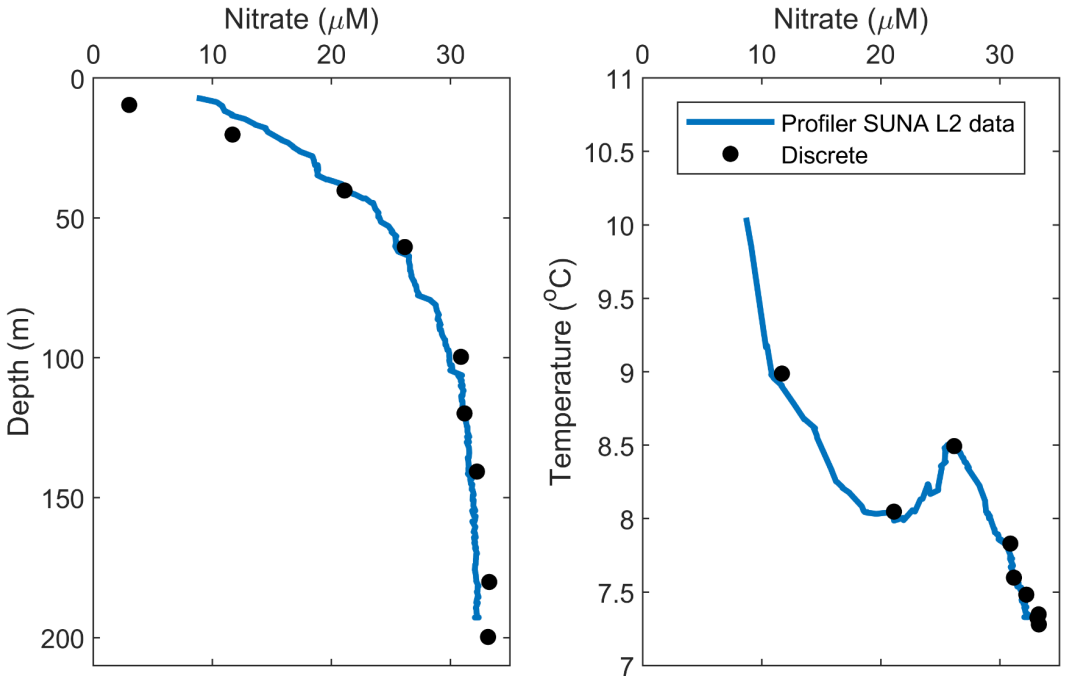


Figure 3.3. Comparison between discrete sample bottle data (black circles) and Level 2 (L2) nitrate data (blue line) collected by the Deep SUNA on the Coastal Endurance Oregon Offshore Cabled Shallow Profiler Mooring on August 11, 2020. The left hand panel is plotted as nitrate concentration with depth, and the right hand panel is plotted as nitrate concentration against temperature. The discrete sample data is Niskin bottles sampled from the CTD cast taken on August 11, 2020 at 23:39 (UTC), and the profiler data is taken from the mooring profile completed on August 11, 2020 at 21:51 (UTC). Note that the right hand panel excludes the surface-most discrete sample, which at >13°C is far above the maximum temperature sampled by the profiler.

The goodness of fit between SUNA derived nitrate concentrations and bottle samples is assessed using matchups based on both temperature and depth (Figure 3.4). In this example, the goodness of fit is reasonable for both regressions as reflected in the similarity of the regressions R^2 values (0.982 and 0.986, respectively). However, excluding the nominal 10 m sample from the temperature regression data set results in a slope remarkably close to 1, and

an offset (0.3 µM) that is well within the accuracy of the nitrate sensor (± 2 µM). Hence we conclude that the *in situ* nitrate data (the OOI-provided L2 data) are suitable for further analysis without additional correction. For this case, since the profiler mooring only successfully collected nitrate data for a few weeks, there is only a single turn-around cruise matchup with discrete samples and it is therefore not possible to adjust for drift over time. For a short deployment and well-aligned discrete sample match > 1 week after deployment, drift is likely insignificant, though users analyzing data from longer deployments should take care to assess and correct for drift.

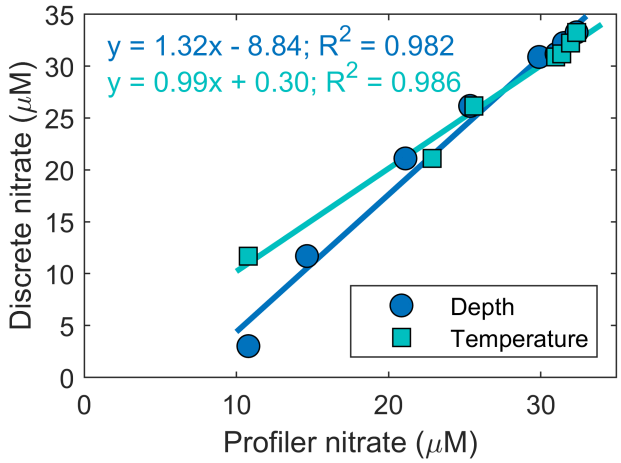


Figure 3.4. Linear regressions of the bottle nitrate and profiler nitrate (SUNA) using matchups based on depth (squares) and temperature (circles). Note that the temperature regression does not include the near surface low nitrate value from outside the depth range sampled by the profiler.

Following from the discrete sample validation, L2 nitrate data from all profiles collected during the August 2020 deployment are plotted in Figure 3.5. We note that near-surface values for nitrate are sometimes negative, though generally within the instrument accuracy specification around zero (± 2 uM). While negative values are impossible as actual nitrate concentrations, this is a legitimate result of the nitrate sensing instrument and technique of the SUNA, which has been widely observed (e.g., Johnson et al., 2017). For further analysis, (e.g., estimates of vertically integrated loads, uptake rates, etc.) the user should consider how to handle these data points uniformly within the analysis and explicitly state the chosen method.

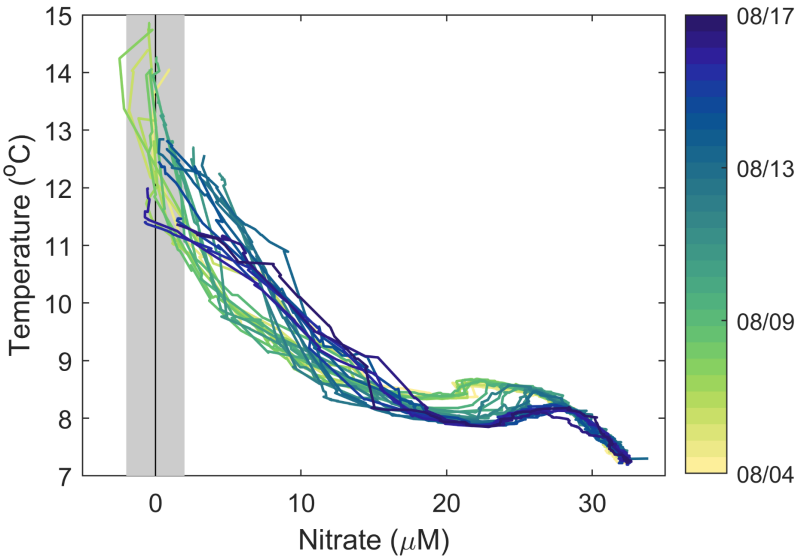


Figure 3.5. All Level 2 (L2) nitrate data collected by the Deep SUNA on the Coastal Endurance Oregon Offshore Cabled Shallow Profiler Mooring during the August 2020 deployment, plotted against temperature. The shaded area is the nitrate concentration accuracy specification (± 2 uM) around zero.

Pseudo-Code

The pseudo-code provided below provides each step in the data processing pipeline for this worked example, with steps organized following the sequence given in Figure 3.2 and the text in Section 3.5. This pseudo-code is intended to support end users in developing their own data processing sequence following these recommended steps using any programming language or OOI data access method of their choice. This example was implemented in Microsoft Excel, and the workbooks used in the data processing are provided as a supplementary resource.

The initial "Assemble data" step of the pseudo-code follows the process for downloading data from OOI's Data Explorer, with further step-by-step details for how to download these nitrate data provided as a supplementary resource. There are multiple ways to find and download data from the OOI servers, described in more detail in the Data Access section of Chapter 1.

Assemble data:

```
Review available nitrate data on Data Explorer;
Review OOI HITL annotations for nitrate and associated CTD data;

Download data of interest including time, depth, salinity, seawater
temperature, nitrate concentration;

Download co-occurring discrete sample nitrate data using the 'find
nearby sample and glider profiles tool' in Data Explorer.
```

Inspect and evaluate the nitrate sensor data:

```
Plot time series of Level 2 nitrate and CTD data;
     Variables analyzed are user-ready output from Figure 3.1

Plot time series of nitrate data from near-surface and/or depth;

Check for evidence of biofouling or lamp drift;
```

Inspect and evaluate the discrete nitrate data:

```
Sort downloaded discrete nitrate date file by time, depth and nitrate
value;

Create a reduced table including at a minimum the CTD depth, CTD
temperature and discrete nitrate data;

Sort reduced table by depth;
```

Plot the discrete nitrate data against CTD depth and evaluate for suspect data points.

Compare nitrate sensor and cruise discrete nitrate data:

Match bottle data with profiler data using both depth, temperature and/or density;

Plot sensor nitrate and bottle discrete nitrate against depth(Figure 3.3) and against temperature (Figure 3.4);

Evaluate for outliers and mis-matched data points;

Look for offsets, particularly at depth;

Look for progressive changes with time;

Plot sensor nitrate against bottle discrete nitrate pairs based on depth and temperature in a property/property plot and apply linear regressions (Figure 3.5);

Evaluate the goodness of fit based on slope, intercept and R^2 value intrinsically and between the depth and temperature pairs;

Exclude suspect data points if any and plot the reduced data set and apply a linear regression;

Evaluate the goodness of fit of the reduced data set compared to the full data set based on slope, intercept and R^2 value.

Apply offset and drift corrections*:

Assess offset and drift (if discrete bottle data available):

Using either a depth matched or density matched data set:

Calculate the offset (intercept, Int) per unit time:

$$dInt/dt = (Int_{start} - Int_{end})/(Time_{start} - Time_{end})$$

Apply offset and drift correction to nitrate sensor data:

For a given sensor nitrate value at depth, z, and time, n:

$NO_3(t_n,z)$ where $t_{start} < t_n < t_{end}$:

$$NO_{3ADJ}(t_n,z) = dInt/dt * (t_n - t_0) + NO_3(t_n,z)$$

**Not implemented in worked example here since corrections were not warranted based on comparison between nitrate sensor and cruise discrete nitrate data, but provided for user reference*

Plot analysis-ready nitrate data

Chapter 4: Carbonate chemistry

4.1 Introduction to the OOI carbonate chemistry sensors

There are four parameters in the carbonate system: Dissolved Inorganic Carbon (DIC), Total Alkalinity (TA), partial pressure of CO_2 (pCO_2), and the concentration/activity of hydrogen ions (pH). The carbonate system has two degrees of freedom, so given any two parameters, the other two can be estimated with relative accuracy. For remotely operated platforms such as morings and AUVs, where automatic measurement systems are necessary, the two most common parameters measured are pCO_2 and pH. As the interactions between the different carbonate system parameters are relatively complicated, thorough quality assurance/quality control (QA/QC) of carbonate data requires a fair bit of carbonate knowledge. For a more detailed background, consult Zeebe and Wolf-Gladrow, (2001) or Dickson et al., (2007). Of the carbonate parameters, OOI platforms include sensors for measuring seawater pH and pCO_2. Details of those measurement systems, including theory of operation, are provided in the sensor section below.

Table 4.1: Manufacturer, model and internal-to-OOI instrument class-series (six letter reference indicator) for the carbonate chemistry sensors operated by OOI, and the platforms on which they are deployed. For the class-series, the class PCO2A represents “pCO2 Air-Sea”, PCO2W represents “pCO2 Water” and PHSEN represents “seawater pH.” The last letter of the class-series is an internal classification by OOI and represents the series (determined by specifications related to the sampling rate, deployment duration, deployment depth, etc.). For a summary of pCO_2 and pH sensors deployed across each of the OOI arrays, see Table 1.1. For deployment locations within each array, see Figures A.1-A.6.

Manufacturer	Model	OOI Class-Series	Platforms
Pro-Oceanus	CO2 Pro-Atmosphere	PCO2AA	Buoy on surface moorings
Sunburst Sensors, LLC.	SAMI-CO2	PCO2WA	Endurance & Cabled shallow Profiler moorings
		PCO2WB	Coastal and Global moorings at fixed depths
		PCO2WC	Global Apex moorings at fixed depths (inductive)
	SAMI-pH	PHSENA	Endurance & Cabled shallow Profiler moorings
		PHSEND	Endurance & Pioneer moorings at fixed depths
		PHSENE	Global Apex surface moorings
		PHSENF	Global subsurface flanking moorings

Sensor types

There are three primary sensor types gathering carbonate system data at OOI arrays: Pro-Oceanus CO_2-Pro Atmosphere, Sunburst SAMI-pH, and Sunburst SAMI-CO2. The Sunburst SAMI-pH provides all pH sensor data collected on the OOI arrays. Both the Sunburst SAMI-CO2 and Pro-Oceanus pCO_2 sensors provide seawater pCO_2 measurements, with the Pro-Oceanus also providing atmospheric pCO_2 measurements and therefore being used on surface buoys, while the SAMI-CO2 sensor is used by OOI for seawater pCO_2 measurements at depth. The manufacturer and model of carbonate chemistry sensors on each type of OOI platform is summarized in Table 4.1, and details of each sensor are described below.

Pro-Oceanus CO_2-Pro Atmosphere

The CO_2-Pro™ Atmosphere instrument measures the partial pressure of CO_2 gas in both surface seawater and air, and is therefore deployed on the OOI moorings/buoys that have a surface expression (Table 4.1, Figures A.1-A.6). The CO2-Pro™ Atmosphere mounts under the buoy, with seawater drawn in for measurement from the instrument location and air from above the buoy drawn in for measurement through a NEMA box. The instrument uses non-dispersive infrared spectroscopy to measure CO_2 in both air and seawater samples. Alternating measurements of pCO_2 in air and water made with the same detector ensure a high level of accuracy for analyses comparing air and seawater pCO_2, such as calculations of air-sea CO_2 flux. In-water measurements of CO_2 are performed in the headspace equilibrated with seawater across the gas-permeable membrane. A water pump is routinely attached to the flow-through head of the membrane equilibrator for faster equilibration with surrounding water. The manufacturer-reported equilibration time (to 63%) is 2.5 minutes for water and 5 seconds for air. The instrument performs internal calibration checks to correct for sensor drift by reading zero “0” CO_2 values in the CO_2-scrubbed gas mixture. Manufacturer-reported accuracy is ±0.5%, with resolution of 0.01 ppm. Various anti-fouling measures, such as a mesh-filter or a copper guard on the sensor water inlet, are sometimes employed to prolong instrument life-time in the field. For further details on the operating principles of this sensor and experimental assessment results, see Jiang et al. 2014.

Sunburst SAMI-pH

The SAMI-pH is a spectrophotometric pH sensor, which uses an indicator dye to detect pH (total hydrogen scale) of seawater. The seawater sample stream is pumped through the instrument and injected with a pulse of the pH-sensitive dye meta-cresol purple (mCP). The reagent is introduced using a 50 µL solenoid pump that enables the same amount of reagent to be added for each sample pH measurement. The mCP-sample mixture is then pumped through a flow cell, where two LEDs send alternating pulses of light with wavelengths matching peak optical absorbances of the protonated (HI^-, measured at 434 nm) and unprotonated (I^{2-}, measured at 578 nm) forms of the indicator. A reference photodiode tracks any changes in the LED light sources over time. The sample pH can be calculated based on the ratio of absorbance at the two wavelengths, measured both with the added indicator dye (sample) and without added dye (blank), the known reagent batch-specific molar absorptivities, and temperature measurements from an onboard thermistor (see Section 4.3 and Figure 4.1 for OOI’s internal data processing

procedure). The manufacturer-reported accuracy of the pH measurement is ±0.003 pH units, with resolution of <0.001 pH units, long term drift of <0.001 pH units per six months, and a response time of 3 minutes. For further information on the operating principles and experimental validation of the SAMI-pH sensor, see Martz et al. (2003), Seidel et al. (2008), Lai et al. 2018, and the Sunburst SAMI-pH user manual.

Sunburst SAMI-CO2

The SAMI-CO2 measures the partial pressure of carbon dioxide (pCO_2) in seawater by equilibrating a pH sensitive indicator solution (Bromothymol Blue) to the sampled seawater. Unlike the SAMI-pH, in the SAMI-CO2, the sample does not mix directly with the indicator, but instead aqueous carbon dioxide in the seawater sample diffuses across a permeable silicon membrane equilibrator, producing a color change in the indicator solution. Similar to the SAMI-pH, the equilibrated indicator solution is then pumped through an optical cell where two LEDs send alternating pulses of light with wavelengths matching the peak optical absorbances for the protonated (HI^-, measured at 434 nm) and unprotonated (I^{2-}, measured at 620 nm) forms of the indicator. As for the SAMI-pH, a reference photodiode tracks any changes in the LED light sources over time. Unlike for the SAMI-pH, where seawater blanks are run along with each sample measurement, for the SAMI-CO2, blanks are determined by measuring deionized water in the same manner as seawater samples. On OOI SAMI-CO2 instruments, these blanks are run every 3.5 days (an empirically derived interval). The sample pCO_2 can be calculated based on the absorbance measurements at both wavelengths, normalized to the blank measurements, temperature measurements from an onboard thermistor, and an instrument-specific calibration curve (see Section 4.3 and Figure 4.1a-c for OOI's internal data processing procedure). The manufacturer-reported accuracy of the pCO_2 measurement is ±3 µatm, with precision of <1 µatm, long-term drift of <1 µatm over six months, and a response time of ~5 minutes. For further information on the operating principles and experimental validation of the SAMI-CO2 sensor, see DeGrandpre et al. (1995), Schar et al. (2009), and Lai et al. 2018.

4.2 OOI standard practices for deployment and calibration

4.2.1 Calibration information

All carbonate system sensors are returned to the manufacturer for calibration, maintenance, and replacement of the reagent (for Sunburst sensors) after each deployment, prior to being redeployed. The original calibration sheets are provided by the manufacturer and are accessible on OOI's Alfresco server:

```
OOI > Instrument & Platform Documents > Calibration and Repair >
Coastal-Global Arrays >{Instrument ID}
```

In particular, the calibrated ranges of the SAMI-CO2 sensors vary between locations and have changed over time as the dynamics of each site have been more fully characterized.

Each sensor is inspected and a Quality Conformance Test (QCT) is performed by the OOI team following receipt after factory calibration/maintenance. The QCT procedures for all sensors involve both visual inspection of the sensor components and housing for defects and issues such as corrosion. The QCT for the Sunburst SAMI pH and pCO_2 sensors involves connecting to the instruments via the client software, and performing checks on the instrument communications, firmware, memory, internal thermistor, pumps, internal clock, and battery voltage. For the Pro-Oceanus pCO_2 instrument, the QCT checks that sensor communications function and the air and water pumps are operational, and that the sensor will sample at the specified sampling frequency. The results of the QCTs are also OOI's Alfresco server:

```
OOI > Instrument & Platform Documents > Test Documents >
Instruments > Coastal-Global Arrays > {Instrument ID}
```

Additional pre-and-post deployment calibration validations have been recently implemented for the SAMI pCO_2 and pH sensors deployed on the Pioneer and Global Arrays. This additional validation test is performed in a seawater test tank at the AVAST facility at WHOI, which maintains a constant temperature and salinity. The instruments sample once every 15 minutes for a total duration of three hours. A Sea-Bird SBE37 CTD collects temperature and salinity data alongside the SAMI. Water samples are drawn every 15 minutes from a sampling tube in the same vicinity as the instrument intake. The water samples are collected and preserved according to standard protocols for lab analysis. Once processed, the bottle data are compared against the instrument data.

4.2.2 Deployment and recovery of the sensors

The SAMI-pH and SAMI-CO2 sensors (PHSEN and PCO2W) share a similar pre-deployment procedure. Using the supplied SAMI client software, the instrument configurations and sampling interval are checked against the deployment configuration, the instruments purged with DI water to eliminate air-bubbles, and the sensors launched. A bag of DI water is kept attached to the inlet nozzle until just before the instrument is deployed to avoid any bubbles getting into the sampling tubing used by the PCO2W and PHSEN and causing them to air-lock. Deployment of the Pro-Oceanus pCO_2 sensor (PCO2A) occurs when the surface mooring is deployed and the mooring computer launched. Deployment configuration, deployment, and recovery checklists can be found on OOI's Alfresco server:

```
OOI > {Array ID} > Asset Information > {Mooring ID} >
{Deployment/Recovery Cruise ID}
```

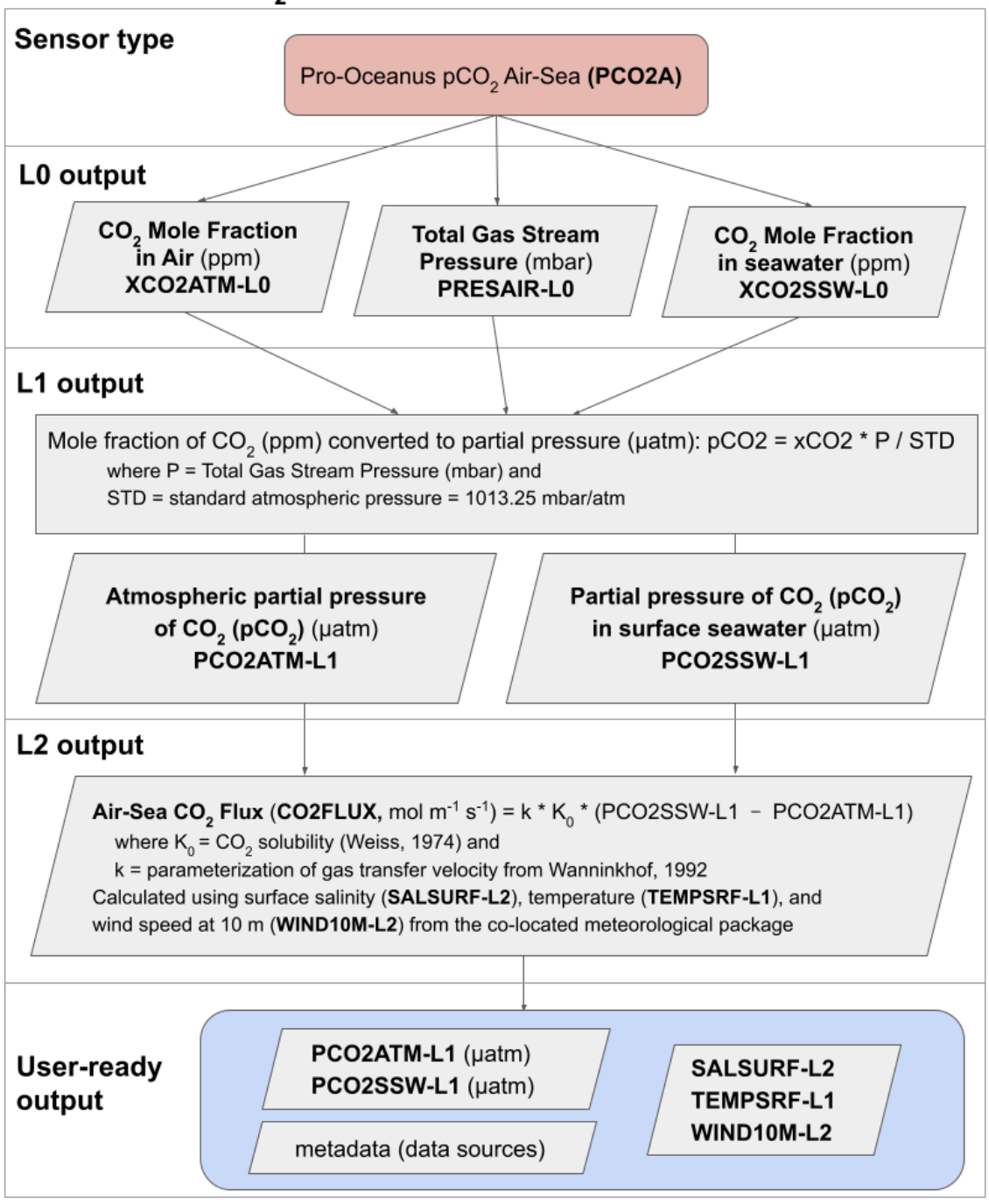
During recovery, when the SAMI-pH or SAMI-CO2 instruments hit the deck, any built-up pressure is released from the sensor housings. Once recovery of the full mooring is complete, the instruments are given a thorough visual inspection with photo documentation, identifying any issues such as corrosion or biofouling. A laptop is used to connect to the instrument using the client software to stop sampling. Internally-recorded data is recovered by downloading from the sensor's internal memory using the supplied client software.

The Pro-Oceanus CO2-Pro Atmosphere instrument does not self-record. Its data is stored on the mooring Data Concentration Logger (DCL) on board the mooring. No data download or special procedures are done on the instrument itself. A visual inspection with photo documentation is performed of the instrument, making note of any instrument damage, corrosion, or biofouling. Data recovery is not performed until the mooring is back at the OOI facility and the mooring well and computer may be accessed.

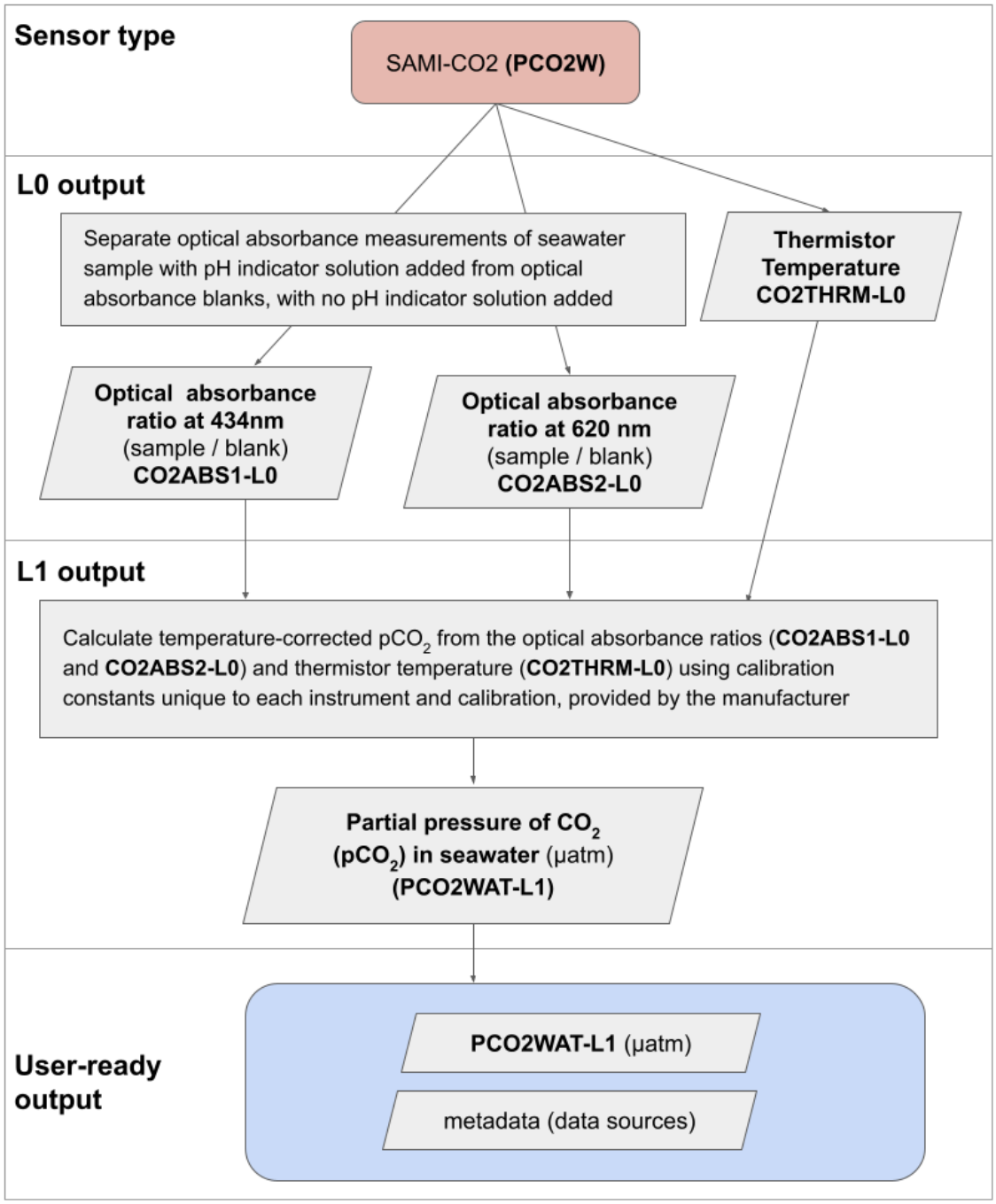
Table 4.2 OOI Test Documentation Numbering Scheme for carbonate chemistry sensors: the Pro-Oceanus CO2-Pro Atmosphere (PCO2A), Sunburst SAMI-CO2 (PCO2W), and Sunburst SAMI-pH (PHSEN).

Document Number	Description
3305-00100-*xxxxx*	Instrument incoming inspection. Brief record of instrument condition and associated calibration files after receipt from vendor
3305-00103-*xxxxx*	PCO2A quality conformance test (QCT) record. Basic functionality check performed after receipt from vendor
3305-00109-*xxxxx*	PHSEN quality conformance test (QCT) record. Basic functionality check performed after receipt from vendor
3305-00110-*xxxxx*	PCO2W quality conformance test (QCT) record. Basic functionality check performed after receipt from vendor
3305-00400-*xxxxx*	Instrument deployment procedure. Final check of instrument settings and launching the instrument.
3305-00409-*xxxxx*	PHSEN deployment procedure. Configuration/settings check, purge, instrument launch
3305-00410-*xxxxx*	PCO2W deployment procedure. Configuration/settings check, purge, instrument launch
3305-00600-*xxxxx*	Instrument data recovery procedure. Shut-down procedures and retrieving data stored on internal memory.
3305-00609-*xxxxx*	PHSEN data recovery procedure. Shut-down of the instrument and downloading internally-recorded data.
3305-00610-*xxxxx*	PCO2W data recovery procedure. Shut-down of the instrument and downloading internally-recorded data.

OOI Internal Carbon Processing Flow: Pro-Oceanus CO_2 Pro Atmosphere (surface moorings)



OOI Internal Carbon Processing Flow: Sunburst SAMI-CO2



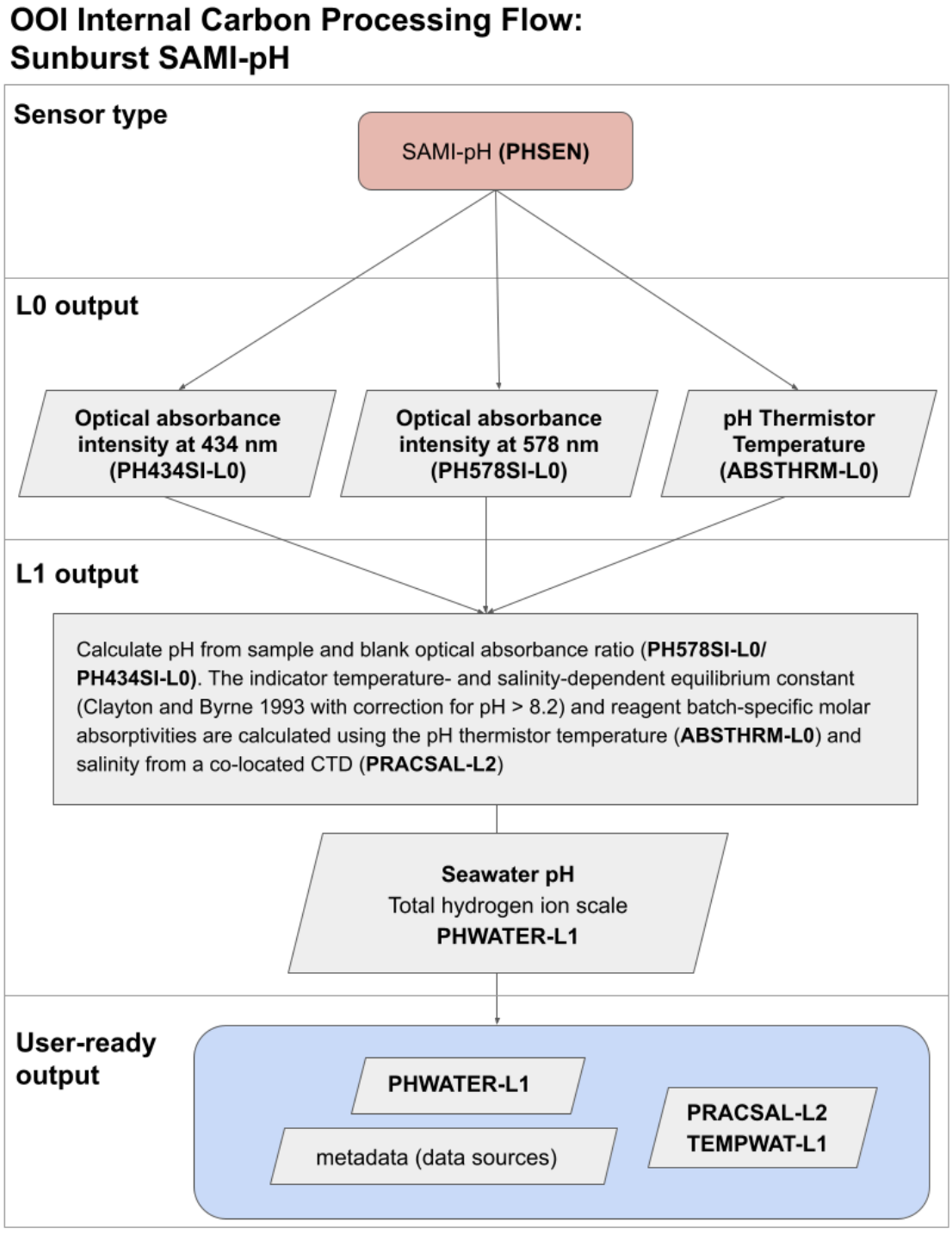


Figure 4.1a-c: Summary of the internal-to-OOI processing workflow for a) the Pro-Oceanus CO_2-Pro Atmosphere, b) the Sunburst SAMI-CO2, and c) the Sunburst SAMI-pH carbonate chemistry sensors. The user-ready data products provided by OOI are not yet analysis-ready data and require further processing by the end user to evaluate, clean, and apply corrections to the data. The user-ready output from OOI's internal data processing in these figures is the starting point sensor data for the recommended end user data processing outlined in Section 4.5 and summarized in the flowchart in Figure 4.4.

4.3 Internal to OOI Data Processing Workflow

Figures 4.1a-c summarize OOI's internal processing workflow for carbon system sensor data from the Pro-Oceanus CO2-Pro Atmosphere, Sunburst SAMI-CO2, and Sunburst SAMI-pH instruments. For all instruments, data are transformed from unprocessed, Level 0 (L0) data products produced by the instrument to user-ready Level 1 (L1) data using OOI open source processing code (Ocean Observatories ion-functions).

For the Sunburst SAMI-pH and SAMI-CO2 sensors, L0 data report optical absorbance intensities collected by the instrument at two instrument-specific target wavelengths, as well as the internal thermistor temperature (see Section 4.1 for summaries of the instrument operating principles). Level 1 data for both seawater pH and pCO_2 from the SAMI instruments are the final user-ready output (PHWATER, seawater pH on the total scale from the SAMI-pH, and PCO2WAT - partial pressure of CO_2 in seawater from the SAMI-CO2). These L1 data incorporate corrections to normalize seawater sample measurements for blanks and temperature- and salinity-dependent corrections using the SAMI thermistor temperature and salinity from a co-located CTD. For further detailed information, see the OOI documentation describing the computations used to calculate the L1 product for the SAMI-pH (sensor code = PHSEN; OOI Document Control Number 1341-00510) and SAMI-CO2 (sensor code = PCO2W; OOI Document Control Number 1341-00490).

For the Pro-Oceanus CO2 Pro Atmosphere, L0 output includes the wet mole fractions of CO_2 measured in both the atmosphere and surface seawater, as well as the gas stream pressure. These data are used to calculate the user-ready L1 output, which are the partial pressures of CO_2 in the atmosphere (PCO2ATM) and in the surface seawater (PCO2SSW, which is directly comparable to PCO2WAT from the SAMI-CO2). For further detailed information, see the OOI documentation describing the computations used to calculate the L1 data products for the Pro-Oceanus CO2 Pro Atmosphere (sensor code = PCO2A, OOI Document Control Number 1341-00260).

OOI also calculates and provides to end users a Level 2 (L2) product, air-sea CO_2 flux, calculated from the Pro-Oceanus L1 data products for surface seawater and atmospheric partial pressures of CO_2 in combination with data from the co-located meteorological package. Full details of this Level 2 CO2FLUX product calculation are provided in OOI documentation (OOI Document Control Number 1341-00270). Since there is no community consensus on the choice of air-sea gas exchange parameterization based on wind speed, and data are not validated or corrected using discrete or external datasets prior to implementing this calculation, this product is not recommended beyond exploratory use. We recommend that end users interested in air-sea CO_2 flux calculate this rate themselves after completing the end user data processing described in Section 4.5.

4.4 Common Data Quality Issues

4.4.1 Issues common across carbonate chemistry sensors

Biofouling & clogging

Biofouling and clogging may look like sensor drift but are essentially impossible to correct for. Biofouling can occur when organisms colonize the interface of the instrument with the seawater/air or begin to grow in the tubing connecting the instrument to the seawater/air and clog or reduce flow. For autotrophic biofouling, we might see consistently low seawater pCO_2 values due to photosynthetic activity of autotrophs. For heterotrophic biofouling, we might see consistently high seawater pCO_2 values due to respiratory activity of heterotrophs. For any type of biofouling, we may see higher variance in the measurements, for example very high diel oscillations (up to 300-500%). The degree of biofouling may be noted in the OOI HITL Annotations (see Section 1.3.3), but not all such instances will be noted and end users will need to perform their own data evaluation to check for biofouling. One approach to identifying biofouling is to compare data with previous years. If variance is much higher than that expected naturally in the historical record, biofouling may be to blame. Biofouling is much more likely to be an issue one or two months into the deployment than when the sensor is just deployed. Additionally, biofouling is a more serious issue at the Pioneer Array and Endurance Array due to high surface ocean productivity at coastal ocean locations. OOI experience using copper guards has not shown the guards to meaningfully reduce biofouling.

Sensor drift

Drift is expected for any sensor, and the SAMI and Pro-Oceanus carbonate chemistry sensors deployed by OOI can drift over the course of deployment. For recommendations for end users on how to identify and correct for sensor drift using turn-around cruise discrete sample data, see Section 4.5.3.

4.4.2 Pro-Oceanus CO2 Pro Atmosphere Sensor-Specific Issues

- ➔ Failure of the solenoid valve can cause air or dissolved gas contamination. This issue can be identified based on the relationship between air and seawater pCO_2 on a per deployment basis. The solenoid valve failure is apparent if air data starts to follow water values (example shown in Figure 4.2). This data should be removed.

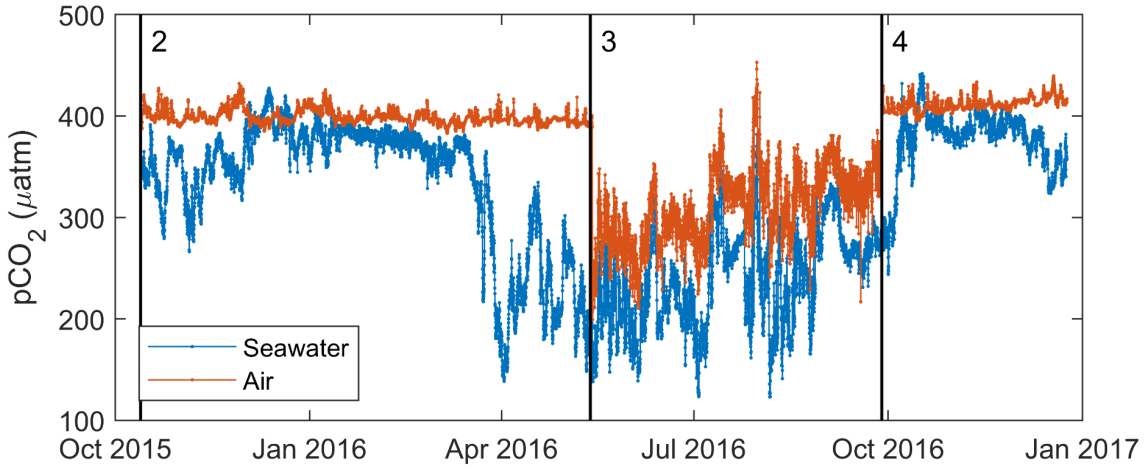


Figure 4.2. Example of solenoid valve failure. Partial pressure of CO_2 in the surface seawater (blue) and atmosphere (red) from the Pro-Oceanus CO_2-Pro Atmosphere at the Coastal Endurance Washington Shelf Surface Mooring Surface Buoy for deployments 2, 3, and 4 (black lines). Air pCO_2 values start to track seawater values at the start of deployment 3, indicating contamination of the atmospheric measurement due to the failure/malfunction of the solenoid valve. These air values are bad and should not be used. A user may choose to use a climatology value instead.

- ➔ Due to a firmware issue, contaminated observations may occur following auto-zeroing. Auto-zero frequency is variable across arrays and deployments, but typically occurs at 6 or 12 hour intervals. This issue can be recognized as low values at regular intervals in some deployments (see Figure 4.3 for example). Contaminated sample points should be removed. These outliers may be reliably detected by calculating the median-absolute deviation for a 24-hour window and identifying any points that fall below the median - 3*median_absolute_deviation.

- ➔ Internal tubing kinking can lead to bad air measurements. This is sometimes caught in the OOI HITL Data Annotations but may have been missed. Air pCO_2 is expected to be relatively stable although coastal arrays may have higher variability due to proximity to land. Abrupt changes or shifts to periods of high variability may be due to internal kinking. Depending on the duration of the issue, the user could choose to use an atmospheric average during this time based on climatology, or treat this as a data gap.

- ➔ Software/firmware issues can cause failure of the instrument to follow sampling schedule or to reset its settings to a default value (including clock, sampling interval, zeroing interval). Some firmware issues may be resolved by re-supplying power to the instrument. These issues are often addressed or annotated before the data is released by OOI.

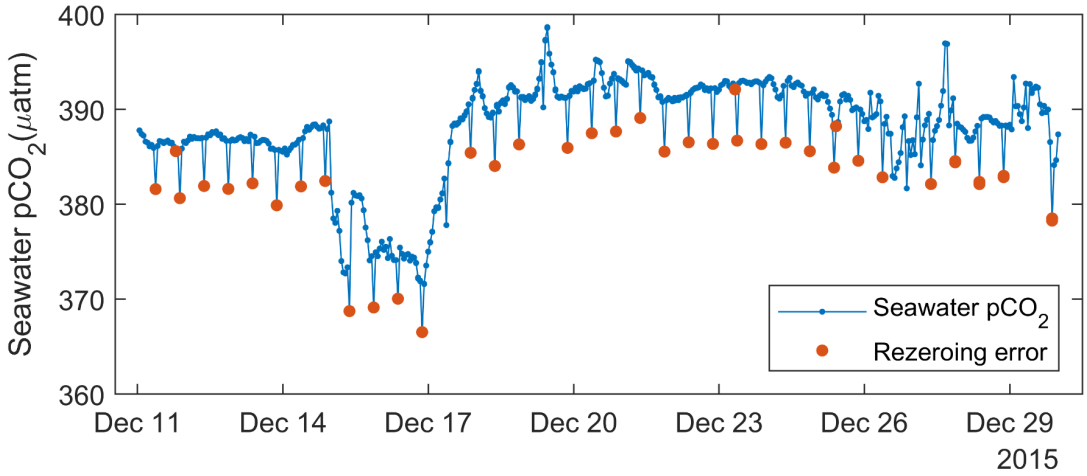


Figure 4.3. Example of auto-zeroing issue. Partial pressure of CO_2 in the surface seawater from December 11 to December 30, 2015, collected by the Pro-Oceanus CO_2-Pro Atmosphere at the Coastal Endurance Oregon Shelf Surface Mooring Surface Buoy. Low values, highlighted in red, were flagged using an algorithm that calculates the median and median absolute difference on a 24-hour rolling window, then flags points that are less than median - 3*median absolute difference. Low values occur approximately every 12 hours due to an auto zeroing issue with the firmware. The filtering algorithm correctly identifies >90% of the expected auto-zeroing error points, with over/under identification during periods of rapid changes

4.4.3 SAMI-CO2 and pH Sensor-Specific Issues

As the SAMI CO_2 and pH systems use similar approaches, many common issues overlap.

- ➔ The most common problem OOI has had with the SAMI pH sensors is they get airlocked on deployment. Air locking occurs if air gets pumped into the instrument, which can happen if it is out of the water during sampling. Upon discovering that a SAMI has airlocked, OOI turns it off for the rest of the deployment, producing a short but unusable dataset and then no available data for the rest of the deployment. Most deployment-long data gaps are due to this problem. End users should check for OOI-provided annotations of this issue, and carefully check data collected immediately before big gaps.

- ➔ Commonly used indicator dyes are m-Cresol Purple (mCP) and thymol blue (TB). mCP has an optimal range of 7.1-8.1 and TB has an optimal range of 7.6-8.6. The user should be aware that measurements in the extremes of this range are less precise. In surface waters, pH is more likely to be near the high end of this range than the low end.

- ➔ SAMI pH instrument and reagent testing has been primarily completed for salinity values > 20 PSU. Measurements with co-located salinity values ≤ 20 PSU should be flagged, as uncertainties in this range will be higher.

- ➔ The sensor reagent can expire or contain impurities. Per Sunburst Sensors, reagent shelf life is expected to be at least 1 year, which spans the majority of OOI sensor deployments. However, prior to 2020, sensors deployed at all arrays except Endurance and the Cabled Array used reagents containing ethylene glycol as an antifreeze which may reduce the shelf life of the reagents over the course of a deployment. Ethylene glycol usage was discontinued after 2019. Impurities in reagents happen more often with the pH sensor than the pCO_2, but the new pH v2 sensors use purified m-cresol which should reduce the problem. This will likely be identified in the pH comparison to pCO_2 (see Section 4.5.4 below) but is otherwise difficult to determine.

- ➔ Occasionally the internal temperature sensor does not measure correctly and may affect internal calculations of pH. The corrected seawater pH (PHWATER-L1) data product pulls salinity and temperature from the closest CTD. A user can determine what salinity and temperature data product was used in the pH calculations by downloading the pH data product from OOINet, which will contain integrated temperature and salinity data as well as the source of those data used in the pH calculations.

- ➔ DI water blanks occur every 3.5 days to monitor and correct for LED drift. If the DI water is inadequately flushed, this can lead to regular pCO_2 spikes immediately following the blanks.

- ➔ Particulates or CDOM can influence the signal and should especially be watched for in coastal or highly productive areas. This could lead to both increased variability in the sensor signal and overall offsets between the measured and true pH values.

4.5 Recommendations for End-User Data Processing

The end user carbonate chemistry data processing flowchart (Figure 4.4) summarizes additional data processing that must be performed by the end user to evaluate, clean, and apply corrections to the OOI-provided carbonate chemistry data. These steps are essential to prepare the OOI-provided pH and pCO_2 data for scientific applications and analyses, especially those involving quantitative interpretation.

Chapter 1 of this document provides an overview of QA/QC procedures recommended for all OOI biogeochemical sensors (Section 1.5), and provides a high-level walk-through and context for each step in the recommended end user carbonate chemistry data processing summarized in Figure 4.4. We recommend that users intending to work with OOI carbonate chemistry data use the flowchart in Figure 4.4 and instructions in Section 1.5 as a starting point and reference for each data processing step. Here, we walk through each of the steps outlined in Figure 4.4, which synthesize the approaches end users can take to correct OOI data products for sensor-specific behaviors and data quality issues described in Section 4.4.

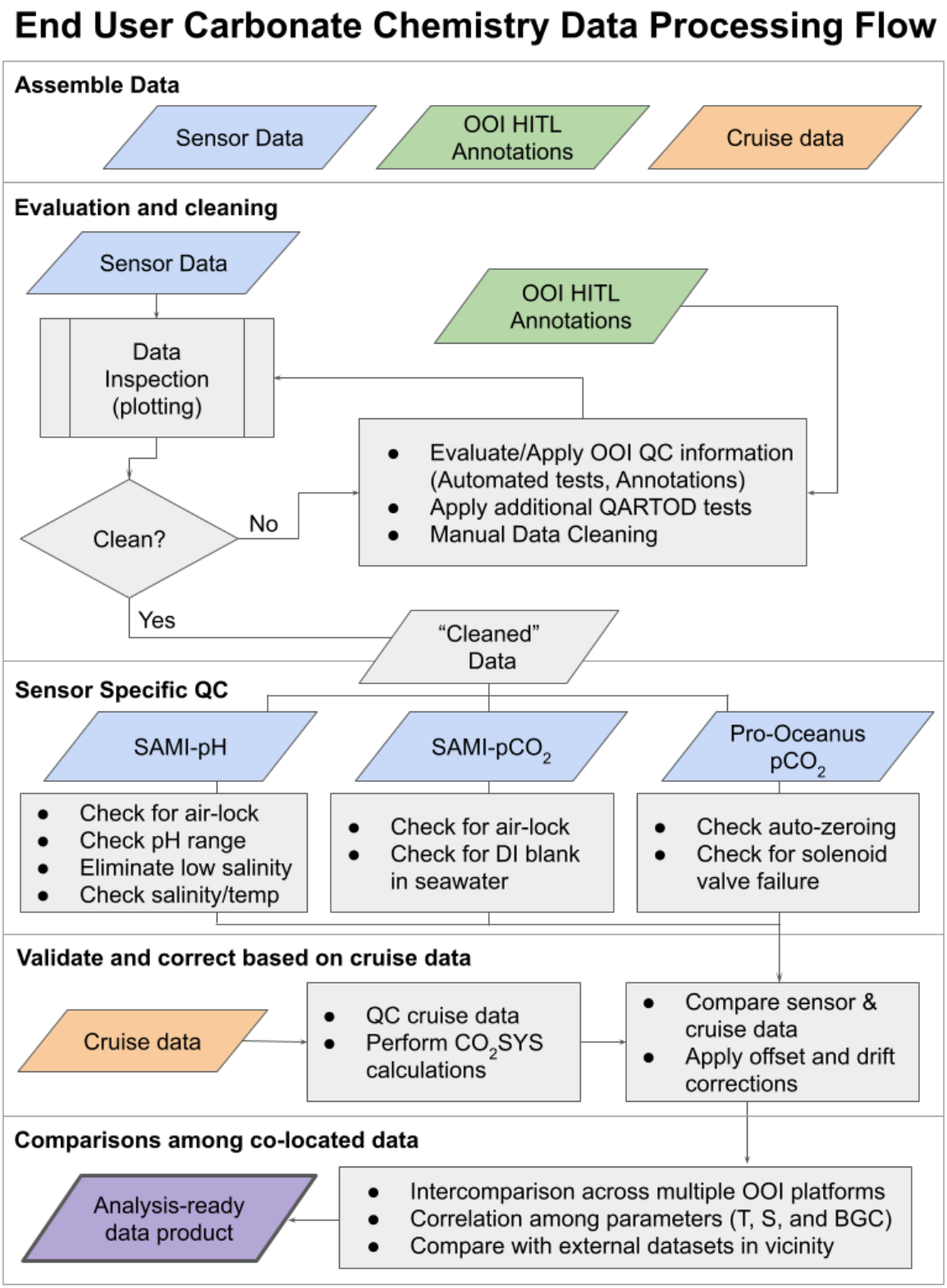


Figure 4.4. Summary of recommended end-user quality control and data processing steps for OOI carbonate chemistry sensor data. These steps will allow the user to evaluate, clean, and apply corrections to the sensor data provided by OOI (see Figure 4.1a-c) to prepare analysis-ready carbonate chemistry data products.

4.5.1 Assemble data

Users will need to assemble pH and/or pCO_2 sensor data, accompanying OOI HITL annotations, and corresponding turn-around cruise data to use in preparing their final analysis-ready pH and/or pCO_2 data. See Section 1.5.1 for details of each of these components.

If not already merged with the carbonate chemistry sensor data, end users should be sure to download pressure, temperature, salinity data, provided by CTDs co-located with all carbonate chemistry sensors in the OOI program. These data will be critical for carbonate chemistry system calculations needed to validate and correct pH and pCO_2 data.

At least two carbonate parameters are measured on discrete samples collected during the turn-around cruises at each array. Table 4.3 indicates what combination of dissolved inorganic carbon (DIC), total alkalinity (TA), pH and pCO_2 measurements are made at each array and the shore-based laboratory that completes these discrete sample analyses.

Table 4.3. Discrete carbonate chemistry water samples taken at each OOI array.

Array	Total Alkalinity	DIC	pH	pCO_2
Cabled Array[C]	–	X	–	X
Coastal Endurance[C]	–	X	–	X
Coastal Pioneer[A]	X	X	X	–
Global Irminger[A]	X	X	X	–
Global Papa*	(1, 2, 3, 6, 7, 8)[A,B]	X[A,B,C]	(6, 7, 8)[A]	(4, 5)[C]
Global Argentine Basin[A]	X	X	X	–
Global Southern Ocean[A]	X	X	X	–

A) Samples processed by Wang CO_2 Lab, Woods Hole Oceanographic Institution
B) Samples processed by Dickson Lab, Scripps Institute of Oceanography
C) Samples processed by Hales Carbon System lab, Oregon State University
*) Samples processed by different labs at different times.

4.5.2 Evaluation and cleaning & sensor-specific quality control

The initial step in recommended end user OOI pH and pCO_2 sensor data processing is to prepare a “cleaned” dataset, applying both automated and human-in-the-loop (HITL) QA/QC to evaluate the data and identify points that may need to be filtered or removed. Section 1.5.2 summarizes the recommended steps for OOI BGC sensor end user data evaluation and cleaning. Here we provide additional context on the application of these steps specifically for OOI carbonate chemistry data:

1. Evaluate and apply OOI-provided HITL Data Annotations. Annotations by the OOI data team identify many platform-wide issues users may need to be aware of (power failure-caused data gaps, etc.), as well as sensor-specific issues. Users of carbonate chemistry sensor data should particularly check the Annotations for the common data quality issues described in Section 4.4. Figure 4.5 shows an example application of Data Annotations to filter and quality control a SAMI seawater pCO_2 dataset.

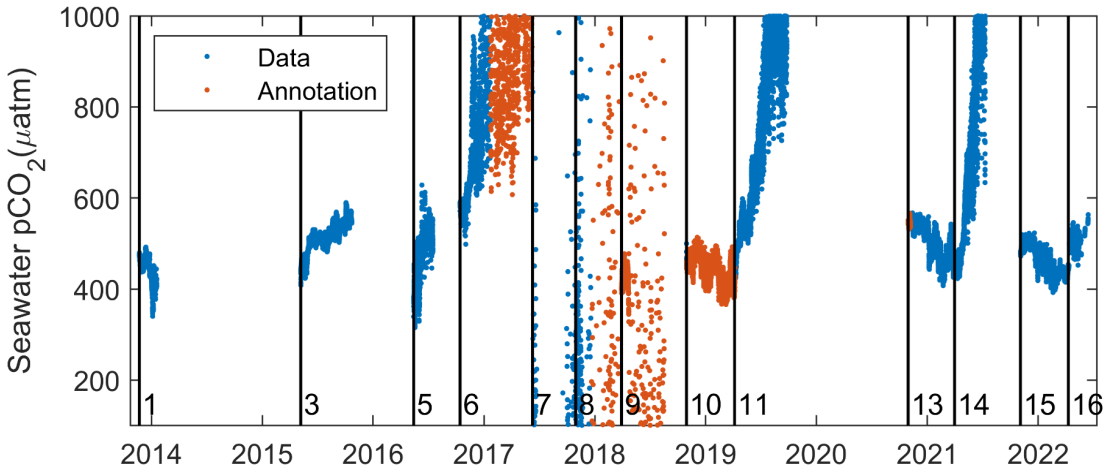


Figure 4.5. Example application of OOI-provided HITL Data Annotations. Partial pressure of CO_2 in the seawater from the Coastal Pioneer Central Surface Mooring Multi-Function-Node Sunburst SAMI-CO2 near the bottom. Red points indicated data that have been flagged in annotations as bad or questionable. The annotation log will contain further information on the issue.

2 & 3. Apply QA/QC based on published QARTOD recommendations. As of October 2022, OOI has implemented automated data quality flags based on Global and Local Range tests for pH and pCO_2 data (see Table 1.2). End users should use the OOI-provided quality flags to identify data that warrant closer evaluation in producing a cleaned and analysis-ready dataset (see Section 1.3.2 for further detail on the OOI QC procedures). End users are also encouraged to implement their own version of the tests listed in Table 1.2 that have not yet been implemented by OOI and/or to manually inspect data to determine whether any of these issues are present in their dataset.

In cases where simultaneously-deployed pH and pCO_2 sensors are either co-located or measuring the same water mass, comparison between these two variables provides an additional opportunity to identify outliers by assessing consistency within the carbonate chemistry system. See Section 4.5.4 below for further details.

Sunburst sensors has publicly available data processing and QC software for the SAMI-pH (see SAMI-pH manual for details). This software examines the raw signal intensity data and flags values where light saturation or higher noise in absorbance interfere with pH measurement (see manual for details). The software will also provide the option to convert the pH data with the *in*

situ or constant salinity values. SAMI QC software flagged values may yield unreliable pH data. OOI has created data processing code in Python to implement the QC procedures on a NetCDF output file from the OOI Data Portal/M2M (API)/THREDDs, but this data processing step is not currently applied to the SAMI-pH data. The raw absorbance data (L0) must be downloaded if a user would like to implement this step by using the Sunburst QC software or the Python code provided by OOI on the ooi-data explorations GitHub repository (see example applying these QC steps in Figure 4.6).

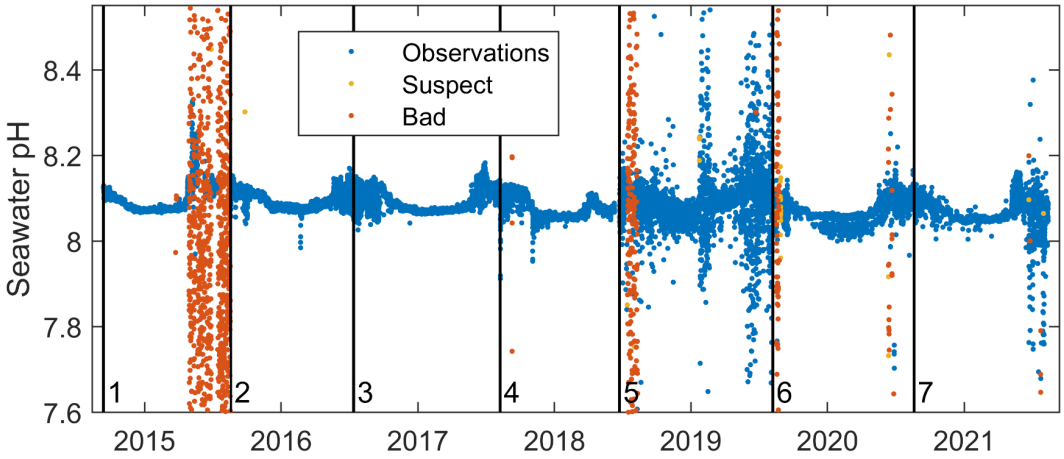


Figure 4.6. Example of SAMI-pH instrument specific Quality Control. Seawater pH measurements (blue) from the Global Irminger Sea Flanking Subsurface Mooring A SAMI-pH. Suspect (yellow) and bad (red) measurements are identified using Sunburst-supplied quality-control algorithms.

4. Manually inspect data to identify and address commonly-known issues. While some examples of the common data quality issues identified in Section 4.4 will be flagged by OOI-provided Data Annotations, users should carefully inspect all data in preparation for their own analysis. Users should be sure to provide their own manual data quality assessments for each of the sensor-specific sets of issues summarized in the flowchart in Figure 4.4.

4.5.3 Validate and correct based on cruise data

Carbonate chemistry discrete sample data collected on turn-around cruises are critical for preparing final analysis-ready data from OOI pH and pCO_2 sensors. Section 1.5.4 summarizes recommendations for using OOI turn-around cruise data to validate and correct sensor data across all BGC sensors. Here we provide recommendations specific to the application of turn-around cruise discrete samples for correction carbonate chemistry data. Discrete samples collected on turn-around cruises include either two or three of the four carbonate chemistry variables (Table 4.3), enabling calculation of the full carbonate chemistry system to provide pH and pCO_2 data that can be used to validate/calibrate measurements from the deployed sensors.

Quality control cruise discrete sample data

QC flags for all carbonate chemistry samples can be found in the summary spreadsheets on the Alfresco server[2]. A description of the QC flag definitions, laboratory methodology, and uncertainty (if provided) can be found in the "readme" files associated with each discrete summary file. When comparing bottle data, a user should filter out any bottle samples that may be flagged as bad, and may want to filter bottle samples flagged as questionable. At minimum, a user should carefully inspect the data from bottle samples flagged as questionable and any metadata notes available to determine if those data should be included. Further QC of the bottle data may be necessary, and we recommend following the data QC procedures outlined in Jiang et al. (2021) for all the bottle data. Some strategies for discrete sample QC could be to evaluate property-property relationships between carbonate data and other ancillary data. For example, salinity and total alkalinity often have a strong linear relationship driven by mixing between water masses and outliers may want to be investigated further. See Section 4.5.4 for additional information on incorporating co-located datasets.

Perform carbonate chemistry system calculations

In order to directly compare bottle data to sensor data users will need to first calculate the full carbonate system to obtain either *in situ* pH or *in situ* pCO_2. Software packages are available in a number of different programming languages (see Orr et al. 2015, Table 1) to perform carbonate system calculations. For carbonate system calculations, we recommend using carbonic acid dissociation coefficients K1 and K2 as calculated from Lueker et al. 2000, the HSO_4 dissociation coefficient from Dickson, 1990, HF dissociation coefficient from Perez and Fraga 1987, total borate concentration as estimated from Lee et al. 2010, and using the total pH scale for direct comparisons with the SAMI-pH sensor.

When directly measured, bottle sample pH is measured and reported at 25 degrees celsius and 1 atm of atmospheric pressure, and on the total pH scale, while sensor pH is measured and reported at *in situ* temperature and pressure, and on the total scale. In order to compare pH sensor data to measured pH data from bottle data, users will need to recalculate the bottle data to *in situ* temperature and pressure using a carbonate system software package (e.g. CO2SYS, seacarb, etc.) and bottle pH and either DIC or TA. If nutrient measurements are available, a user can include those as input variables for the carbonate system calculations, otherwise the software packages will assume nutrients are negligible.

When bottle pCO_2 is directly measured, it is measured at one atm of atmospheric pressure and at ambient lab temperature, but the lab also reports values after recalculation using *in situ* temperature and pressure and thus can be directly compared to sensor pCO_2. If it is not directly measured, pH and/or pCO_2 will need to be derived using measurements of two other carbonate system parameters (e.g. DIC/TA, DIC/pH, TA/pH, or DIC/pCO_2). If a choice of sample pairs is available to use to quantify the carbonate system (e.g., a subset of the samples processed at

[2] Summary sheets for turn-around cruise discrete water samples can be found on the Alfresco Server using the following path: `OOI > {Array ID} > Cruise Data > {Cruise ID} > Ship Data > Water Sampling`

WHOI) we recommend using either the DIC/pH or the TA/pH pair to quantify pCO_2 as this pair has been shown to have the lowest calculation uncertainty for pCO_2 (see Orr et al. 2018 for more details). Add-ons to commonly used carbonate system packages to propagate uncertainty in carbonate system packages are available in Orr et al. 2018 and we recommend users evaluate the error in calculated parameters prior to comparison with *in situ* sensor data. We also recommend that end users investigate all sample pair combinations for a robust estimation of the uncertainty in pCO_2 and/or pH calculations.

Apply corrections to sensor data

Once discrete samples have been quality controlled and used to calculate *in situ* pH or pCO_2, they can be directly compared to the sensor data to assess if sensor corrections are needed (see Worked Example and Figure 4.8). If there is an offset, sensor data can be corrected by quantifying the offset between the discrete sample data and the sensor data, and correcting the sensor data by this offset. In cases where an OOI-deployed pH or pCO_2 sensor successfully collects data for the full deployment period, it will still be collecting data at the time of the turn-around cruise where it is recovered, providing discrete sample data for calibration/validation at the end as well as at the beginning of the deployment. If the offset between the discrete samples and sensor data changes from the beginning and end of the deployment, this sensor drift can be corrected by assuming that the offset varies linearly over the length of the deployment. If multiple instances of bottle samples exist (at multiple different time points) a linear regression correction can be applied instead of a linear interpolation.

In applying these corrections, end users should consider the potential impacts of spatial and temporal variability that may affect the match-up between the turn-around cruise discrete sample data and the sensor to be calibrated. Carbonate chemistry sensors are primarily deployed on fixed depth platforms within OOI, which will generally only provide 1-2 discrete location- and depth-aligned samples for calibration on each turn-around cruise. Caution should be exercised to assess whether a correction is warranted to improve accuracy. In some cases, offsets between discrete samples and sensor data may best be explained by spatial variability in the system causing the bottle data to not be entirely representative of the water measured by the instrument. For coastal sites, especially for surface measurements, water can be very heterogeneous over small spatial scales. Offset corrections may not be warranted in cases where the discrete sample measurements (including their calculated uncertainty envelope propagated in CO2SYS calculations) overlaps with the envelope of variability measured by the sensor.

4.5.4 Comparisons among co-located data

Comparisons among co-located data provide valuable opportunities to improve constraints on carbonate chemistry sensor calibration and identification of anomalous data. Section 1.5.5 provides context on recommended intercomparisons valuable across all BGC sensor types. Here we provide additional context specific to working with carbonate chemistry data.

Data comparison across pH and pCO_2 sensors

In cases where simultaneously-deployed pH and pCO_2 sensors are either co-located or measuring the same water mass, we recommend that end users complete a combined assessment of these two variables. The Coastal Pioneer, Coastal Endurance, and Regional Cabled Arrays all include directly co-located pH and pCO_2 sensors, and all arrays include both pH and pCO_2 sensors in near-surface waters that may in some cases provide comparable measurements across different depths during periods with deeper mixed layers (see schematic drawings of the OOI arrays, Figures A.1-A.6). A high degree of anti-correlation is often observed for seawater pCO_2 and pH due to the influence of primary production and respiration in the absence of strong abiotic processes and events that alter equilibrium of the carbonate system of seawater, e.g. air-sea exchange during storms. This relationship can be used as a tool for data QC, although this approach requires the user to consider the influences of these processes in the observed relationships and observed deviations.

For each of the comparison tests below, we recommend that end users flag outliers as varying more than three standard deviations away from the mean residual value, consistent with QARTOD standards:

$$\text{Outlier} = [\text{residuals} < \text{mean(residuals)} - 3 * \text{std(residuals)}] \mid [\text{residuals} > \text{mean(residuals)} + 3 * \text{std(residuals)}]$$

These tests will detect cases where either the pCO_2 or pH data is an outlier. For any data point where the relationship flags an outlier, additional context and user analysis will be needed to identify which of the sensors to flag as the source of the potentially erroneous data.

1) pH and pCO_2 relationship:
   - ➔ Perform a linear regression between the natural log of seawater pCO_2 (as the independent variable) and seawater pH (as the dependent variable) (modification of the procedures in Fassbender et al. 2017) for co-located measurements (or measurements from the same water mass as identified in T-S space), and use this relationship to identify outliers.
   - ➔ Use the linear regression to predict pH from observed pCO_2 values and compute the residuals (differences) between original pH and predicted pH for each co-located (in time and space) pCO_2 sw and pH measurement.

2) H^+ and pCO_2 relationship: use linear regression to predict H^+ from pCO_2 and H^+ in nmol/kg
   - ➔ Perform a linear regression between the seawater pCO_2 (as the independent variable) and seawater H^+ concentration (as the dependent variable). H^+ concentration (in nmol/kg) is calculated as $H^+ = 10^{-pH} * 10^9$
   - ➔ Use the linear regression to predict H^+ from observed pCO_2 values and compute the residuals (differences) between original pH and predicted pH for each co-located (in time and space) $pCO_{2,sw}$ and H^+ measurement.

3) pH and xCO_2 relationship. xCO_2 is directly measured by the Pro-Oceanus sensor and can either be downloaded directly or computed from pCO_2 and total gas stream pressure.
    - ➔ Perform a linear regression between the natural log of seawater xCO_2 (as the independent variable) and seawater pH (as the dependent variable) (modification of the procedures in Fassbender et al. 2017), and use this relationship to identify outliers.
    - ➔ Use the linear regression to predict pH from observed xCO_2 values and compute the residuals (differences) between original pH and predicted pH for each co-located (in time and space) seawater pCO_2 and pH measurement

4) Identify outliers by comparing calculated pH from estimated Total Alkalinity and measured pCO_2
    - ➔ Calculate pH from sample pairs including TA or DIC and observed pCO_2 (see section 4.5.3 above). High resolution TA or DIC is not available at the OOI arrays, but published relationships exist to estimate TA from salinity and temperature. A data user can estimate TA from the methods in Carter et al. (2018), Broullón et al. 2019, or generate their own relationships using observed discrete samples.
    - ➔ Compute the difference between original and calculated pH based on estimated TA and observed pCO_2 and assign outliers as above

Comparison with other co-located measured parameters

pH and pCO_2 sensor data can be compared with co-located CTD temperature and salinity data, deployed alongside every BGC sensor within the OOI program, as well as other BGC variables (see Figures A.1-A.6 for where carbonate chemistry sensors are deployed alongside other BGC sensor types). See chapters 2, 3, and 5 for details on recommended end user data processing for other BGC sensors needed to facilitate such intercomparison.

Seawater pCO_2 and pH may be correlated with temperature, salinity, dissolved oxygen, and/or chlorophyll. Temperature and salinity affect the solubility of CO_2 in seawater whereas oxygen and chlorophyll can indicate efficiency of air-sea gas exchange (oxygen) and biological activity (oxygen and chlorophyll). These correlations (or anti-correlations) will vary by region, but for stationary sites, enough data can exist to identify trends. We recommend the user compare the full time series of seawater pCO_2 and pH with co-located parameters to see whether such correlations exist. These comparisons are also useful in instances where discrete data is limited. The correlations can also be used to evaluate periods of potential instrument issues. For example, a large change in surface seawater pCO_2 over a short period may indicate instrument issues, however, if sea surface temperature also changes rapidly, following the correlation, this could instead indicate a new water mass has entered the region and the seawater pCO_2 values are accurate. Similarly, a spike in seawater pCO_2 accompanied by a spike in temperature or salinity may indicate a good measurement rather than an instrument issue. This analysis step could help in interpreting data flagged with 3 (questionable or interesting).

Comparison with external-to-OOI datasets

External-to-OOI datasets can provide an opportunity to compare with additional validation data beyond that provided by the OOI program. Some public data products are available that may be useful for generating climatologies, understanding realistic values, and developing regional property-property relationships. These datasets may not be found in very close proximity to the OOI arrays, thus decreasing their utility for direct assessment of instrument accuracy, but may be useful to identify outliers or periods of time where the sensors were not functioning properly. Some suggested databases include the Surface Ocean CO_2 Atlas (SOCAT), the Rolling Deck to Repository (R2R), the Marine Boundary Layer Reference dataset (MBL), the NOAA Global Greenhouse Gas Network (CCGG), and the Lamont-Doherty Earth Observatory underway database (LDEO). This list is not exhaustive and an OOI data user may find other useful datasets in the Ocean Carbon and Acidification Data System (OCADS). Suggestions for using these external databases can be found below.

Seawater surface pCO_2 - Several databases can be useful to use as quality control checks for OOI mooring data, and a user should be aware that there may be considerable overlap in the data archived in the various databases suggested above. Many seagoing vessels, including both scientific research vessels and other ships-of-opportunity have sensors installed to measure surface underway seawater conditions that can be compared with the surface datasets from the Pro-Oceanus Air-Sea sensor package, and possibly those sensors installed on the near-surface instrument frames. Raw datasets from research vessels can be found in the R2R repository, but these datasets may not be quality controlled. Surface fCO_2 data from vessels with underway CO_2 sensors is usually submitted to the SOCAT database where it is eventually made available in regular data releases after data quality control and review (completed by a working group following the SOCAT QC Cookbook). SOCAT datasets are released every few years incorporating new data that has been submitted since the last release (the current version was released in 2022). Both individual cruise tracks and gridded data products at various timescales are available. The LDEO database product also contains quality controlled surface underway pCO_2 data but additionally has gridded climatological air-sea CO_2 flux estimates that a user may use to identify outliers in reported CO_2 flux products.

Air pCO_2 - External data available from the MBL and CCGG databases can be used to help quality control the pCO_2 Air dataset from the Pro-Oceanus Air-Sea sensor on the surface moorings. CCCG data includes samples collected at NOAA-based observatories and volunteer-collected sites across the globe of which may be found in relatively close proximity to OOI array locations. MBL meridional reference data is a combined and continuous data product from a number of NOAA observatories.

If co-located data are available, we recommend the following procedure to compare external datasets to OOI mooring data (see example in Figure 4.7):

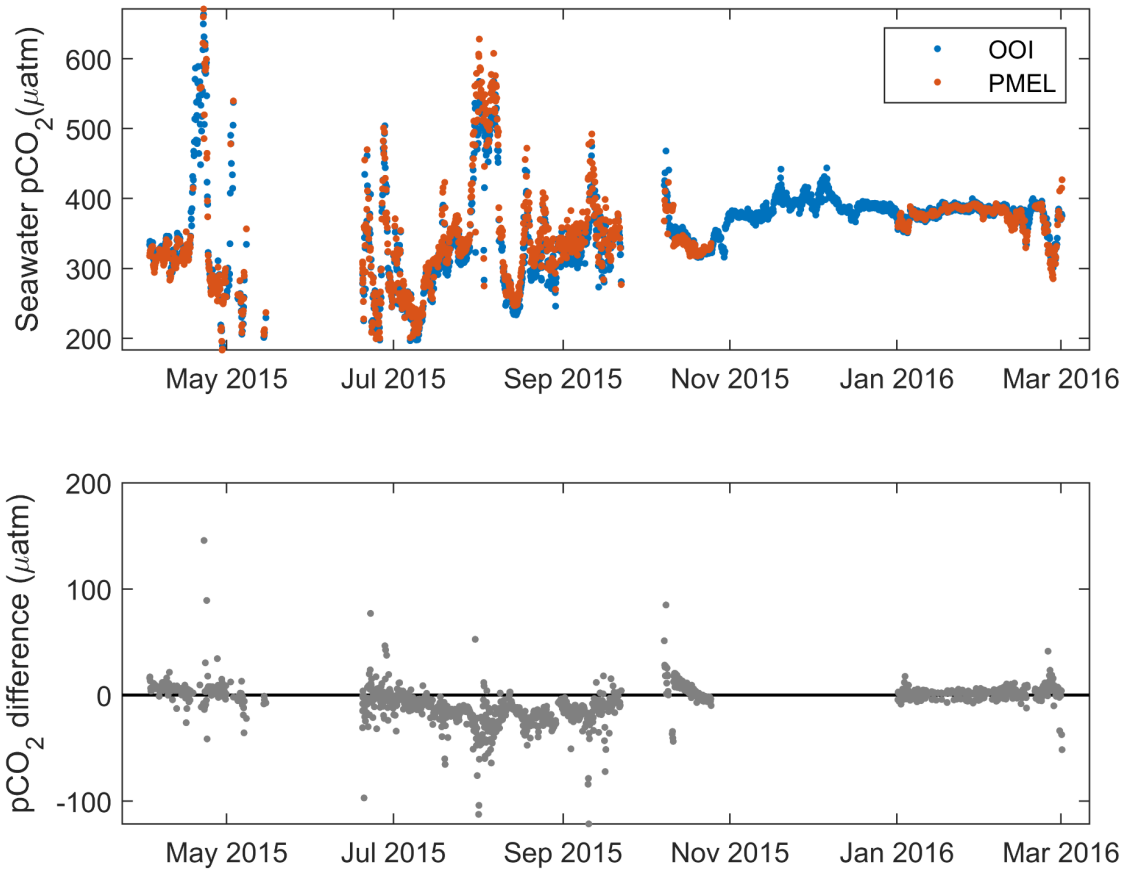


Figure 4.7. Example of comparison with external data. Top: Surface seawater pCO_2 from April 2015 to March 2016 at the OOI Coastal Endurance Oregon Shelf Surface Mooring Surface Buoy (44.64°N, 124.30°W; blue) compared with nearby PMEL NH10 mooring (44.6°N, 124.3°W; red). Bottom: Differences between the pCO_2 measured by the OOI mooring and the PMEL mooring

➔ The user should download or query the external database for co-located observations in both time and space. It is up to the user to decide the appropriate search grid and matching distance for their needs.

➔ Compute and plot differences between the observations or gridded datasets and the OOI mooring data, and compute mean and standard deviation of the residuals per deployment and for the whole time series. Offsets in specific deployments may indicate sensor issues that might not be noticed when looking at individual deployments or the full time series. *Note that an observed offset from gridded or combined data products is acceptable - a user should look for consistency in that offset over the full timeseries.*

➔ If the offset between mooring data and gridded products in any specific deployment shows a significant difference compared to the overall time series offset, an absolute correction to the Pro-Oceanus pCO_2 Air and pCO_2 seawater (from the same instrument) may be applied. This correction, at most, should bring the offset in line with the rest of the time series. This test is normally applied after other QC tests have been performed.

A user should note the magnitude of this correction in their metadata and which deployments this correction was applied to.

➔ If co-located observations in both time and space are available from external discrete datasets that are not gridded, mooring data can be directly compared and corrected if offsets are observed.

4.6 Worked Example

We provide here an example of how to apply the end user processing workflow described in Section 4.5 and summarized in the flowchart in Figure 4.4. This example prepares a seawater pCO_2 product for scientific use from Level 1 (L1) seawater pCO_2 data (PCO2WAT-L1, processed by OOI as shown in Figure 4.1b) from a Sunburst SAMI-CO2 instrument from the Coastal Pioneer Central Surface Mooring Multi-Function-Node (Figure 4.8).

This example illustrates the importance of using discrete sample carbonate chemistry measurements from turn-around cruises to validate and, if needed, correct the final seawater pCO_2 data from the sensor. Turn-around cruises provide discrete sample dissolved inorganic carbon, total alkalinity, and pH measurements corresponding to the location of the SAMI-CO2 at both the beginning and end of the deployment, enabling calculation of *in situ* seawater pCO_2 for comparison with the sensor data.

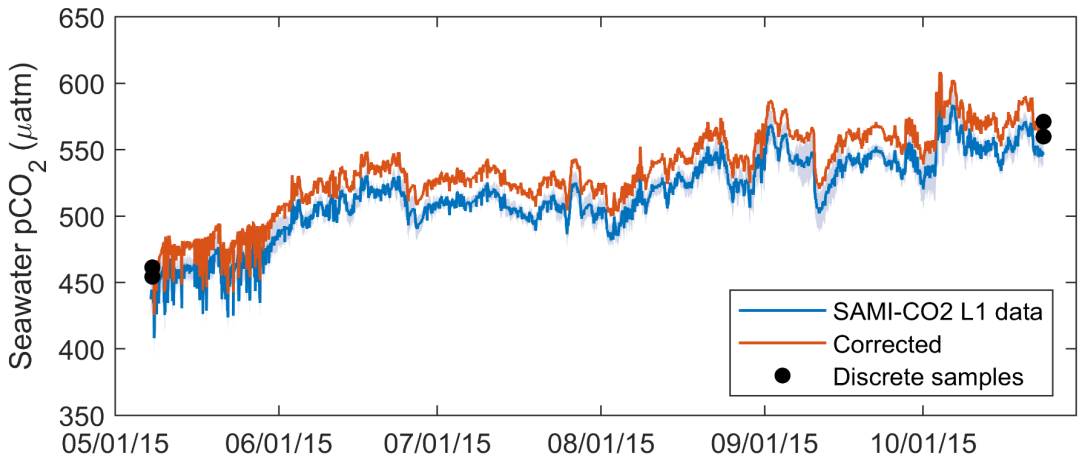


Figure 4.8. Carbonate chemistry worked example, preparing seawater pCO_2 data from a SAMI-CO2 instrument for analysis. Partial pressure of CO_2 in the seawater (data - blue; 1 week running standard deviation - light blue) from the Coastal Pioneer Central Surface Mooring Multi-Function-Node Sunburst SAMI-CO2 near the bottom (approximate depth of 133 m). Bottle samples (black) were collected at the beginning and end of the deployment and measured for DIC, TA, and pH; pCO_2 values are calculated using CO2SYS carbon-system parameterization software using DIC and pH as input parameters.

Pseudo-Code

The pseudo-code provided below provides each step in the data processing pipeline for this worked example, with steps organized following the sequence given in Figure 4.4 and the text in Section 4.5. This pseudo-code is intended to support end users in developing their own data processing sequence following these recommended steps using any programming language or OOI data access method of their choice. The Python notebook used to implement this example, including annotations to assist users in following the script, is provided as a supplementary resource, but is intended solely as a reference and not as a template for end user data processing code.

Assemble sensor data

```
Load and review available seawater pCO2 data;
Load and review OOI HITL Data Annotations;

In this example, cruise data are loaded below
```

Evaluate and clean sensor data

```
Plot time series of seawater pCO2 data;
     Variable analyzed is User-ready output from Figure 4.1

Plot OOI quality control flags;
Filter data to remove data points considered "bad" or "suspect";
     For this example all data flagged suspect data are removed, but
     note that user discretion is needed to evaluate quality flags, and
     for further sensor-specific quality control
```

Load and evaluate cruise discrete sample data

```
Load carbonate chemistry bottle measurements from turn-around
cruises;
Filter carbonate chemistry measurements using quality control flags;
```

Calculate carbon system parameters

```
Calculate conservative physical properties of seawater needed for
carbonate chemistry calculations;

Calculate seawater pCO2 from discrete sample measurements of Total
Alkalinity and Dissolved Inorganic Carbon using CO2SYS;
     In cases with >2 measured parameters, internal consistency among
     the bottle measurements can provide an additional QC check
```

Validate and correct sensor based on cruise data

Filter for discrete samples collected near the pCO_2 sensor;
Plot sensor data together with aligned turn-around cruise bottle data;

Compare sensor and discrete sample pCO_2 from beginning and end of deployment to determine if offset and/or drift corrections are warranted;

Apply offset and drift corrections

Calculate offset and drift (slope) for linear fit sensor correction based on times when discrete samples were collected on turn-around cruises;

T1 = time of turn-around cruise at beginning of sensor deployment
T2 = time of turn-around cruise from end of sensor deployment

$Offset_{T1}$ = Discrete sample value, T1 (µatm) - sensor value, T1 (µatm)
$Offset_{T2}$ = Discrete sample value, T2 (µatm) - sensor value, T2 (µatm)

$Slope = (Offset_{T2} - Offset_{T1}) / (T2 - T1)$

Apply corrections to sensor-measured seawater pCO_2 data:
$pCO2_{L1}$ = time series of uncorrected Level 1 seawater pCO_2 data (µatm)
$pCO2_{CORR}$ = corrected seawater pCO_2 data (µatm)
Tn = time point between T1 and T2

$pCO2_{CORR} = pCO2_{L1} + (Slope * Tn + Offset_{T1})$

Plot analysis-ready pCO_2 data

Additional Examples and Resources

As a companion to the SAMI-CO2 seawater pCO_2 worked example above, we also provide additional notebooks with examples of the processing workflows for OOI carbon system data. These include annotations to assist users in following the scripts and data processing workflows. As with the notebook accompanying the example above, these are intended as a

reference and starting point for users desiring additional examples of how to implement the workflows described in Section 4.5 and Figure 4.4, but are intended solely as a reference and not as a template for end user data processing code.

Carbon system Python notebooks:

- **Bottle_Data**: This notebook provides an outline and example of working with and processing OOI discrete bottle data. It reviews how to load bottle data, parse the data quality flags for initial quality control, and calculate carbon system parameters using CO2SYS to yield carbonate chemistry parameters for direct comparison with OOI *in situ* sensor data as well as assessment of internal consistency among the measured discrete sample carbonate chemistry parameters.
- **Downloading_Data**: This notebook highlights two methods for programmatically accessing and downloading data for OOI carbon system. The first method utilizes OOI's API to perform machine-to-machine (M2M) queries for data from the OOI THREDDS data server. The second method requests data from OOI's Data Explorer ERDDAP server.
- **PCO2A**: This notebook provides an example of working with seawater pCO_2 data from a Pro-Oceanus CO2 Pro-Atmosphere instrument. It utilizes data from the PCO2A dataset deployed on the Global Irminger Apex Surface Mooring (see Figure A.2). The example steps through the process for loading data, plotting data, applying annotations, initial data quality control, and validation against bottle discrete sample data.
- **PHSEN**: This notebook provides an example of working with pH data from a Sunburst SAMI-pH sensor. It utilizes data from the PHSEN dataset from the Global Irminger Flanking Mooring A at 30 m depth (see Figure A.2). The example steps through loading data, plotting data, applying annotations, quality control, and validation against bottle pH.
- **PCO2W:** This notebook provides an example of working with pCO_2 data from a Sunburst SAMI-CO2 sensor. It utilizes data from the PCO2W dataset from the Coastal Pioneer Central Surface Mooring at an approximate depth of 133m (see Figure A.3). The example steps through loading data, plotting data, applying annotations, quality control, and validation against pCO_2 values calculated from discrete bottle carbon system measurements, and is further described above in the worked example in Figure 4.7 and accompanying pseudo-code.

Carbon system MATLAB notebook:

- This MATLAB live script provides an example of working with seawater pCO_2 data from a Pro-Oceanus CO2 Pro-Atmosphere instrument. It utilizes data from the PCO2A dataset deployed on the Coastal Pioneer Array Inshore Surface Mooring (see Figure A.3). The example steps through plotting data, applying annotations, initial data quality control, and validation and correction against bottle discrete sample data.

Chapter 5: Bio-Optical Measurements

5.1 Introduction to the OOI bio-optical sensors

The term ‘bio-optical’ refers to methods and associated sensors that use light to infer information about biota in the ocean. Two common methods discussed herein are fluorometry and optical backscatter. Fluorescence is the excitation of organic molecules by ‘short’ wavelength light that subsequently stimulates measurable emission of longer-wavelength light. Fluorescence sensors are used to infer concentrations of *chlorophyll-a* and fluorescent dissolved organic matter (FDOM) in seawater. Backscatter sensors measure backscattered light at 700 nm, which provides information about particles (living and dead) in the water. These two sensors/methods taken together comprise an ensemble of active optical *in situ* observations relevant to ocean biology. Additionally, both FDOM and backscatter measurements are used for making corrections in estimating nitrate concentration (see Chapter 3: Nitrate).

Measurement principle of *chlorophyll-a* fluorometry

All plant life contains *chlorophyll-a*, a photosynthetic pigment. *In vivo chlorophyll-a* fluorescence is a widely used method to estimate *chlorophyll-a* concentration, although the accurate determination of *chlorophyll-a* concentration from fluorescence is not trivial and requires several assumptions and a number of procedural steps to improve the quality of the end data product (Cullen, 1982; Falkowski and Kiefer, 1985). The primary assumption is that the fluorescence intensity of *chlorophyll-a* is directly proportional to its concentration.

1. A volume of water is illuminated with a blue-green excitation light source
2. This excitation energy is absorbed by *chlorophyll-a* in phytoplankton photosystem II
3. Subsequently the pigment molecules emit lower-energy red light
4. Emitted light is detected by a sensor and recorded in terms of an intensity value
5. This detected signal is standardized in a calibration process (“factory calibration”) that depends in part on metadata supplied by the manufacturer

Measurement principle of FDOM fluorometry

Fluorescent Dissolved Organic Matter (FDOM) observation works on the same principle as *chlorophyll*-a, except FDOM absorbs ultraviolet light and fluoresces blue light. FDOM is a proxy estimate for Colored Dissolved Organic Matter (CDOM), although only a subset of all CDOM is fluorescent. In open ocean marine systems, this type of fluorescence signal is primarily from organic molecules associated with phytoplankton, while in coastal regions a significant fraction of CDOM can be supplied by terrestrial sources.

The nature of CDOM probed by fluorometry is a function of both the excitation and emission wavelength that is chosen. Near-ultraviolet light (360 nm) is used for excitation, and emission is measured in the blue wavelength (460 nm). As the bio-optical instrument used for fluorometry by the OOI program (see details below) has only a single channel to measure excitation and

emission wavelengths, this sensor cannot provide information on the composition of the CDOM present, which requires information on the spectral shape of CDOM absorption.

Measurement principle of the optical backscatter sensor

Rather than measuring a fluorescence response, optical backscatter sensors (here abbreviated 'bb700') both transmit and receive red light (700 nm). Optical Backscatter (red wavelengths) data are used to provide estimates of turbidity and suspended solids in seawater that scatter photons of red light (wavelengths of light that fall between roughly 630 and 740nm) in the backward direction (see Section 5.3 and Figure 5.1 below for information on the OOI data backscatter data products). Turbidity commonly describes water clarity and is a gross assessment of light attenuation factors by suspended solids. As an optical method, backscatter measures the relative quantity of suspended solids but not their composition.

Table 5.1: Model and internal-to-OOI instrument class-series (six letter reference indicator) for the Sea-Bird/WETLabs fluorometers operated by OOI, and the platforms on which they are deployed. For the class-series, the class FLORD represents the 2-wavelength fluorometer ('D' indicates DUO), measuring chlorophyll-a fluorescence (CJHLAFLO) and optical backscattering (FLUBSCT), and the class FLORT represents the 3-wavelength fluorometer ('T' indicates TRIPLE), which also measures Fluorescent Dissolved Organic Matter, FDOM (CDOMFLO[3]). See Table 5.2 for further information about each of these parameters. The last letter of the class-series is an internal classification by OOI and represents the series (determined by specifications related to the sampling rate, deployment duration, deployment depth, etc.). For a summary of fluorometers deployed across each of the OOI arrays, see Table 1.1. For deployment locations within each array, see Figures A.1-A.6.

Model	OOI Class-Series	Platforms
Sea-Bird/ WETLabs ECO-FLBBCD (triplet)	FLORTD	Fixed depths on all array moorings
	FLORTJ/ FLORTK	Endurance & Pioneer Profiler moorings
	FLORTM/ FLORTO	Gliders
	FLORTN	Pioneer AUVs
Sea-Bird/ WETLabs ECO-FLbb (duplet)	FLORDG	Global Apex surface moorings (≥ 40 m)
	FLORDL	Global subsurface Profiler moorings
	FLORDM	Global array Gliders
Sea-Bird/ WETLabs ECO-FLNTU (duplet)	FLNTU	Endurance and Regional Cabled Deep Profiler moorings
Sea-Bird/ WETLabs ECO-FL (FDOM)	FLCDR	Endurance and Regional Cabled Deep Profiler moorings

[3] FDOM is a proxy measurement of Colored/chromophoric dissolved organic matter (CDOM) in this sensor leading to occasional conflation of acronyms in OOI metadata files and in the literature.

OOI bio-optical instrumentation

There are two instruments used for fluorometry in the OOI program, both from the Sea-Bird Environmental Characterization Optics (ECO) series (Table 5.1). The first fluorometric instrument is a duo configuration that measures *chlorophyll-a* and backscatter. The second is a triplet version that also measures FDOM. WETLabs originally designed and manufactured the fluorometer used by the OOI, before the company was acquired by Sea-Bird in 2010 (and together with Satlantic - the maker of the OOI nitrate sensors, which was acquired by Sea-Bird in 2011). Sea-Bird was subsequently rebranded as Sea-Bird Scientific. The WETLabs legacy name is maintained by Sea-Bird as part of the fluorometer model reference. Details of each of the parameters measured by these instruments is provided in Table 5.2.

Additional bio-optical sensors used in OOI

In addition to the fluorometers and backscattering sensors discussed above, the OOI program includes several other bio-optical measurements. Discussion of these instruments and associated data processing are outside the scope of this document; however, we list them here for the reference of the reader:

- Photosynthetically Active Radiation (PARAD) sensors, which integrate downwelling light within a wavelength band relevant to photosynthesis: 400-700 nm
- Downwelling Spectral Irradiance (SPKIR) sensors, measuring downwelling irradiance at 7 channels
- Spectrophotometer measurements of hyperspectral absorption and attenuation (OPTAA sensors; WetLabs (now Sea-Bird) AC-S instrument, 4 Hz, ~80 spectral channels across the visible wavelengths 400-700 nm).

Table 5.2. ECO fluorometers and backscattering sensors (User manual, 2019)

Parameter	Wavelength	Range, Sensitivity	Output variables and units
Chlorophyll-a (Chl) (CHLAFLO)	470/695 nm (EX/EM)	0–30, 0.015 µg L^{-1} 0–50, 0.025µg L^{-1}	Raw counts and estimated chlorophyll-a concentration (µg L^{-1})
Scattering (FLUBSCT)	700 nm	0–3, 0.002 m^{-1} 0–5, 0.003 m^{-1}	Raw counts, volume scattering function (Beta, m^{-1} sr^{-1}) and total optical backscattering (m^{-1}) corrected for temperature and salinity using data from a co-located CTD is what is reported by OOI
Fluorescent Dissolved Organic Matter (FDOM) (CDOMFLO^{+})	370/460 nm (EX/EM)	0–375, 0.184 ppb	Raw counts and fluorometric FDOM concentration (ppb)

$^{+}$FDOM is a proxy measurement of Colored/chromophoric dissolved organic matter (CDOM) in this sensor leading to occasional conflation of acronyms in OOI metadata files and in the literature.

5.2 OOI standard practices for bio-optical sensor deployment and calibration

5.2.1 Calibration information

Factory calibration information, i.e., using scale factors and dark counts supplied by the instrument maker for calibration of the raw sensor output, is used by OOI for fluorometers; secondary calibrations, i.e., using discrete *chlorophyll-a* measurements, are not performed by OOI and are left for the end user (see Section 5.5). A detailed application note from the vendor describes the calculation of calibration coefficients. The original calibration sheets are provided by the manufacturer and are accessible on OOI's Alfresco server:

```
OOI > Instrument & Platform Documents > Calibration and Repair >
Coastal-Global Arrays >{Instrument ID}
```

5.2.2 Sensor turnaround information

Sensors are mounted on the platform and integrated with the data logger system, confirming the data and power to/from the sensor are functioning as expected. Depending on where the sensor is mounted (e.g., fixed depth sensor in the photic zone), it may be taped with 3 mm polypropylene black pipe tape. In areas where biofouling is a large issue, such as the photic zone (~upper 40 m) off the coasts of Oregon and Washington, an additional wrap of copper foil tape is added around the sensor as well as a copper plate with wipers on the face of the sensor to help prevent biofouling from developing on the side and growing down into the light field of the sensor. Fluorometers on gliders or profilers and those on the Pioneer Array (New England Shelf) have not required additional biofouling measures beyond the factory-installed wiper and copper plates. This practice will be re-examined when the Pioneer Array is relocated to the southern Mid-Atlantic Bight in 2024. The biofouling treatment is not documented in the metadata or calibration documentation, except in pre- and post-deployment photos. AUV's do not have biofouling issues due to their short deployment duration.

Following mooring recovery, all sensors are returned to the vendor for factory service and calibration. OOI intends to factory calibrate AUV fluorometers annually, though this is not always achieved in practice. Sensors returned from the vendor go through an Instrument Receiving (IR) process (collecting vendor supplied documentation, entering receipt of the sensor into the database, etc.) followed by a Quality Conformance Test (QCT) to verify the sensor is working as expected. The results of the QCTs are available OOI's Alfresco server:

```
OOI > Instrument & Platform Documents > Test Documents >
Instruments > Coastal-Global Arrays > {Instrument ID}
```

This test is limited, as it is run in the laboratory and so cannot be used to comprehensively test *in situ* conditions. Updated sensor calibrations that were generated post recovery are not applied by OOI to previously gathered data; they are only used for the next deployment. However, these post-recovery calibrations can be obtained by contacting OOI.

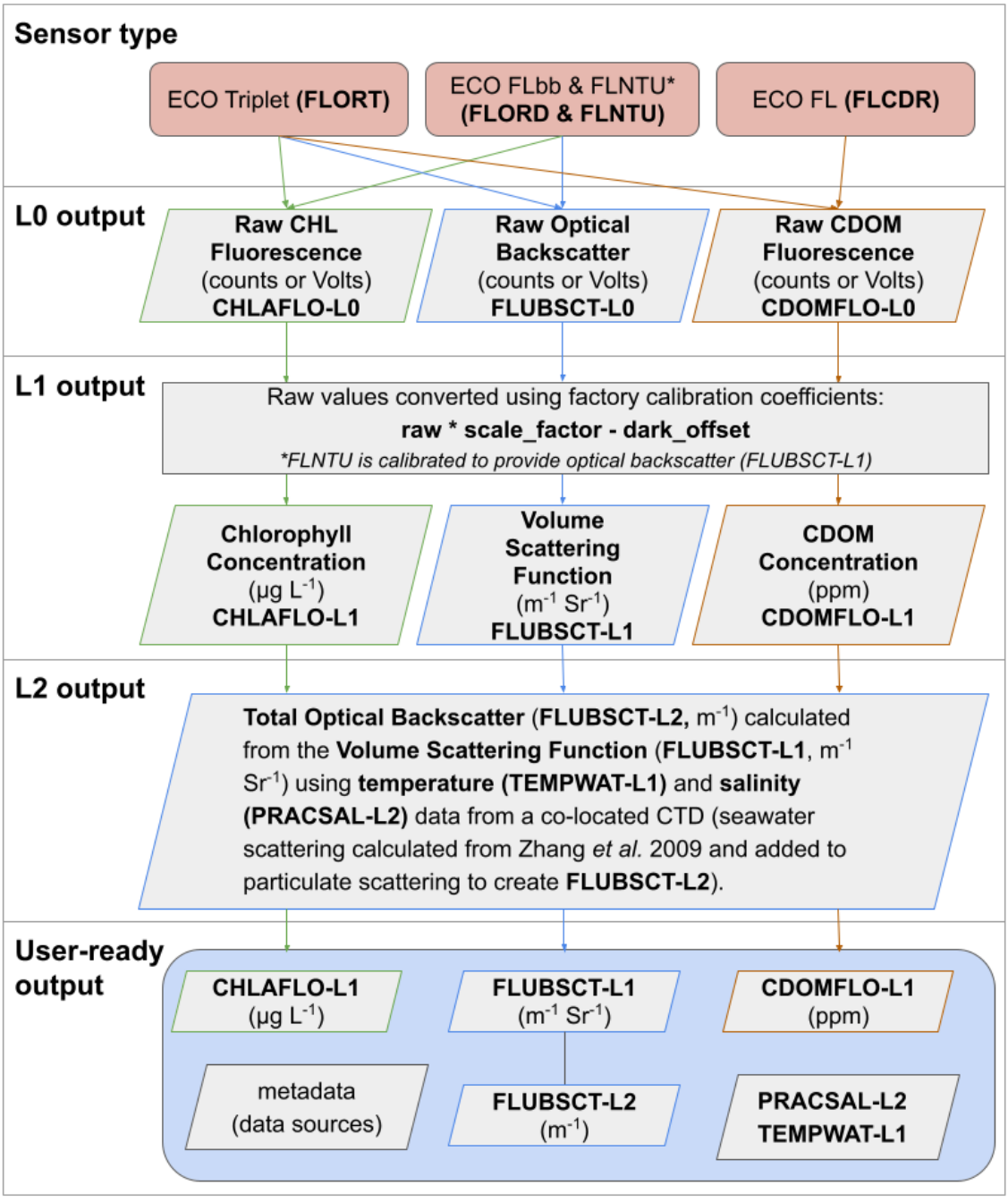


Figure 5.1. Summary of the internal-to-OOI processing workflow for data from Sea-Bird ECO Triplet (FLORT), FLbb (FLORD), FLNTU, and FL (FLCDR) bio-optical sensors. The user-ready bio-optical data products provided by OOI require further processing by the end user to evaluate, clean, and apply corrections to the data. The user-ready output from OOI's internal data processing in this figure is the starting point sensor data for the recommended end user data processing outlined in Section 5.5 and summarized in the flowchart in Figure 5.2

5.3 Internal to OOI Data Processing Workflow

OOI data management defines three different levels of data associated with fluorometer information. Level 0 (L0) are the raw data counts, Level 1 (L1) are values of the variables of interest following application of the factory calibrations, and Level 2 (L2) data are secondary variables produced using ancillary data (Figure 5.1). OOI provides detailed documentation on the computations used to calculate the OOI L1 Fluorometric Chlorophyll-a Concentration (CHLAFLO) data product (WetLabs two and three channel fluorometers; sensor codes = FLORD, FLORT; OOI Document Control number 1341-00530); the OOI L1 Fluorometric CDOM concentration (CDOMFLO) data product (WetLabs three wavelength instrument, sensor code = FLORT; OOI Document Control Number 1341-00550); and the OOI L1 and L2 Optical Backscatter (FLUBSCT) data products (WETLabs two and three channel fluorometers, sensor codes = FLORD, FLORT; OOI Document Control Number 1341-00540).

With respect to backscatter, it is further worth mentioning that OOI reports “total optical backscattering”, which contains backscattering by seawater. Some other programs, for example the BGC-Argo program, report particle backscattering (bbp, m^{-1}), from which the seawater contribution has been removed. Therefore, if the user wishes to compare OOI data to other backscatter data, this needs to be taken into consideration and the seawater contribution either subtracted from the OOI data, or added to the bbp data.

5.4 Common Data Quality Issues

Common to all bio-optical sensors mentioned here, data issues can occur that relate to sensor functioning and/or outside interference. Specific issues relating to sensor functioning can include offsets in the dark counts, drift over the course of the deployment and biofouling.

5.4.1 Dark counts

In situ dark measurements are often different from factory calibration. Dark measurement differences with respect to the factory calibration show up as offsets. They are particularly apparent when the offset results in negative engineering values at low signal levels. This issue can be addressed by shifting the lowest observed values to the lowest detectable value from either the sensor or the discrete bottle sample, and then adding that offset to all of the data. This is described in more detail in Section 5.5.4 below.

5.4.2 Sensor drift

Drift over the course of deployment as the light source ages. Drift in light output from LEDs significant to the measurements is generally expected to be small and is indistinguishable from biofouling from biofilm development on the face of the instrument. Diagnosing drift in fluorometer data can be done using:

- Engineering data as an indicator: Recorded voltage over time can change as components age and change resistance.
- Dark signal at depth / at night over deployment.
- Comparisons with discrete bottle samples measured by high performance liquid chromatography (HPLC)

5.4.3 Biofouling

Biofouling impacts the accuracy of fluorescence measurements by obscuring the sensor. Users can check the cruise data notes and photos for the post-deployment inspection, noting observed biofouling or lack thereof placed in the HITL metadata. Sequences of decreasing maximum chlorophyll signal are indicative of heterotrophic biofilm growth, particularly in data sets showing diel cycles. Increasing signal levels, especially if they demonstrate no diel cycle and they exponentially approach the maximum output from the instrument, are indicative of growth of autotrophic films on the face of the instrument or more generally on the mooring or vehicle.

Biofouling impact on the backscattering sensor is generally seen as a linear or exponential increase in signal over time. The impact of biofouling on the backscattering sensor is generally more pronounced compared to the chlorophyll sensor. If biofouling is suspected for the backscattering data then the chlorophyll sensor should also be evaluated.
Unlike other biogeochemical sensors, sensor delay has not been documented in fluorometers, as the detected fluorescence is a fairly instantaneous response to the emitted light and should accurately represent the phytoplankton present in the targeted water volume at the measurement time.

5.4.4 Issues specific to *chlorophyll-a* fluorescence

Here we systematically outline frequently encountered issues concerning the *chlorophyll-a* fluorescence sensor, some of which are sensor-issues, while the bulk of issues relates to phytoplankton physiology, which affects the fluorescence measurement. The issues related to phytoplankton physiology, are discussed in more detail in the Section 5.5.5 below. The following issues result in difficulties using *chlorophyll-a* fluorescence to estimate *chlorophyll-a* concentration and/or phytoplankton biomass:

- Cells synthesize more or less *chlorophyll-a* as a function of light and nutrient availability
- Species composition and other factors (nutrient status, light acclimation status) affect the fluorescence emitted per unit of *chlorophyll-a*, with variations in the factor relating factory-calibrated fluorescence-derived *chlorophyll-a* and HPLC-derived (discrete sampling) *chlorophyll-a* ranging between 1 and 8 in the global oceans (Roesler et al., 2017).
- Depression of fluorescence due to non-photochemical quenching (NPQ) under super-saturating irradiance levels, e.g., fluorescence, will be lower during the middle of the day than at night for the same chlorophyll-a concentration.

The distinction between sensor functioning and physiological issues is useful, as different end users may have different requirements. For example, a user interested in the “best possible” estimate of *chlorophyll-a* will want to make all the possible adjustments that are available for an improved estimate of *chlorophyll-a*. Someone interested in cross-referencing the data with other fluorometer data (e.g., from BGC-Argo floats) may be satisfied with an NPQ-correction, followed by a standardized application of the “Roesler factor”, i.e., a universal division by two to adjust the fluorescence-derived *chlorophyll-a* estimate for the global bias observed for WETLabs ECO fluorometers (Roesler et al., 2017).

Deriving *chlorophyll-a* from fluorescence is not trivial, hence data from the *chlorophyll-a* fluorescence sensor should be treated with extra caution, as outlined below. Users of L1 data should note that the units for this sensor are already in concentration for *chlorophyll-a*. This factory-calibrated *chlorophyll-a* estimate can be off by up to factor 8 when compared to high performance liquid chromatography (HPLC)-derived *chlorophyll-a* estimates from discrete samples (Roesler et al., 2017).

How different kinds of issues can be approached for a given sensor depends in part on where the sensor is mounted. For example, if the sensor is profiling, then the *in situ* dark can be estimated from deep data, i.e., from strata of the water column where the *chlorophyll-a* concentration is expected to be nil. If the sensor is mounted at a fixed depth, this correction is not available and the resulting increased uncertainty in the *chlorophyll-a* estimates needs to be taken into consideration.

5.4.5 Issues specific to backscatter

On profiling instruments such as gliders, the presence of bubbles in the upper water column can affect measurement quality during the downcast as soon as the glider leaves the surface. This effect is seen in the backscattering sensor maxing out at shallower depths, followed by an exponential decay to normal backscattering values. The magnitude of the effect tends to be related to the sea state but does not result in predictable affected depth layers. It is recommended that where possible, users only use backscatter data from upcasts (for simplicity), or remove the affected data from the profile.

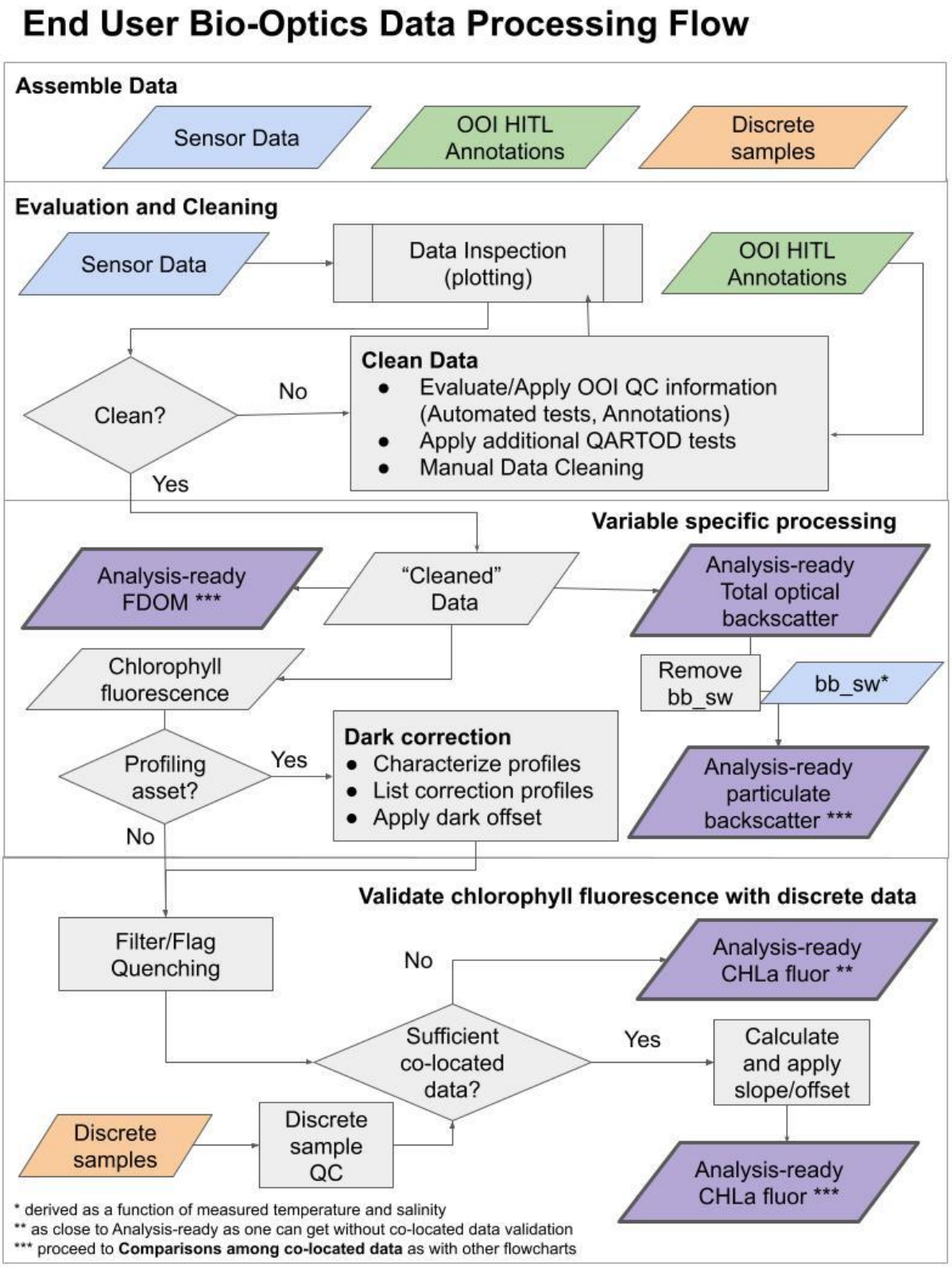


Figure 5.2. Summary of the recommended end-user processing workflow for bio-optical data. These steps will allow the user to evaluate, clean, and apply corrections to the data provided by OOI (see Figure 5.1) in order to process it into analysis-ready *chlorophyll-a*, backscatter and FDOM data products.

5.5 Recommendations for End-User Data Processing

The recommended steps for end-user processing of bio-optical data are summarized in Figure 5.2. These steps will allow end-users to properly evaluate, clean and apply additional corrections to the data. These steps are essential to prepare the OOI-provided bio-optical *chlorophyll-a*, backscatter and FDOM data for scientific applications and analyses, especially those involving quantitative interpretation.

Chapter 1 of this document provides an overview of QA/QC procedures recommended for all OOI biogeochemical sensors (Section 1.5), and provides a high-level walk-through and context for each step in the recommended end user bio-optical data processing. We recommend that users intending to work with OOI bio-optical data use the flowchart in Figure 5.2 and instructions in Section 1.5 as a starting point and reference for each data processing step. Here, we walk through each of the steps outlined in Figure 5.2, which synthesize the approaches end users can take to correct OOI data products for the most common bio-optical sensor-specific behaviors and data quality issues described in Section 5.4.

In the following subsections we discuss the general end-user data processing steps in terms of bio-optical data, but highlight any steps or procedures that are specific to any of the specific data products, e.g., *chlorophyll-a* concentration, backscatter, FDOM.

5.5.1 Assemble Data

Users will need to assemble bio-optical sensor data, accompanying OOI HITL annotations, and corresponding turn-around cruise data to use in preparing their final analysis-ready bio-optical data. See Section 1.5.1 for details of each of these components.

5.5.2 Evaluation and cleaning & sensor-specific quality control

The initial step in recommended end user OOI bio-optical sensor data processing is to prepare a “cleaned” dataset, applying both automated and human-in-the-loop (HITL) QA/QC to evaluate the data and identify points that may need to be filtered or removed. Section 1.5.2 summarizes the recommended steps for OOI BGC sensor end user data evaluation and cleaning. Here we provide additional context on the application of these steps specifically for OOI bio-optical sensor data:

1. Evaluate and apply OOI-provided HITL Data Annotations. Annotations by the OOI data team identify many platform-wide issues users may need to be aware of (e.g., power failure-caused data gaps, etc.), as well as bio-optical sensor-specific issues.

2 & 3. Apply QA/QC based on published QARTOD recommendations. As of December 2022, automated data quality flags for *chlorophyll-a* fluorescence (FLORD, FLORT) are in development within OOI (see Table 1.2), but are not yet fully implemented. Users are encouraged to check for and apply OOI’s automated data quality flags once they become available. If OOI implementation of these automated tests is still not available, end users are

strongly encouraged to implement their own version of the tests listed in Table 5.3 and/or to manually inspect data to determine whether any of these issues are present in their dataset. OOI is not currently prioritizing the development of automated data quality flags for FDOM (CDOMFLO) or backscatter (FLUBSCT) data products, and so users are urged, when possible, to implement their own QARTOD tests for these data products. We provide below some brief instructions on the implementation of appropriate QARTOD tests for bio-optical data, summarized in Table 5.3.

Table 5.3: Summary of QARTOD test recommendations for OOI bio-optical data.

	Chlorophyll-a (FLORD, FLORT)	**Total Optical Backscatter (FLUBSCT)**	**FDOM (CDOMFLO)**
Flat line test	fixed-depth and profiling sensor data; use L0 data	fixed-depth and profiling sensor data; use L0 data	fixed-depth and profiling sensor data; use L0 data
Gross range test	fixed-depth and profiling sensor data; use L1 data	fixed-depth and profiling sensor data; use L1 data	fixed-depth and profiling sensor data; use L1 data
Climatology test	fixed-depth and profiling sensor data; use L1 or higher	fixed-depth and profiling sensor data; use L1 or higher	fixed-depth and profiling sensor data; use L1 or higher
Spike test	Fixed-depth data only; L0 or L1 data OK	Fixed-depth data only; L0 or L1 data OK	Fixed-depth data only; L0 or L1 data OK

Flat line test: The flat line test should be applied to raw (count) data rather than calibrated data. The test is expected to catch issues with sensor performance. It is applicable to all data types (fluorescence and backscatter) and should be used on fixed-depth as well as profiling sensor data. It can be run on the calibrated L1 or uncalibrated L0 data.

The flat line test on the burst level is implemented to catch instrument problems, i.e., if a sensor is stuck on a single value for the duration of a burst. Data should be flagged as bad if the variability in data during a burst is equal to zero.

The flat line test can further be applied to data that has been pooled, for example by taking the median of each measurement burst. Alternatively, it can be applied to the burst level data but with a different window of comparison, i.e., rather than comparing the data across a single burst (see above), the test can also be applied over a given time or depth interval where natural variability in the measurements is expected.

- ➔ **For fixed-depth sensors**, this test can be implemented on data gathered over the course of a day. The expectation is that some variability should be observed over 24 hours if the sensor were performing correctly, so if the sensor is stuck on a single value

over the course of the day, the data should be flagged as suspicious or bad. If there is a data gap such that the first and last data point in a data window are more than 26 hours apart, the test is not performed and the data are flagged as "probably good".

➔ **For profiling sensors**, this test can be implemented on a "per profile" basis, again with the expectation that some variability should be observed in a profile. If the sensor is stuck on a single value over the whole profile, the data should be flagged as suspicious or bad.

Gross Range test: The gross range test checks whether data are within a reasonable range and is applicable to all sensors and regardless of where/how they are mounted. For *chlorophyll-a*, all data <-0.2 or >100 mg m^{-3} are flagged as "probably bad". The assumption is that the sensor is performing correctly in these cases, but biofouling of some sort (e.g., gooseneck barnacles in the field of view) is affecting the signal. For backscatter data, the gross range test is performed on the calculated particle backscattering coefficient (b_{bp}) or total backscattering coefficient (b_b; see Figure 5.1 and Section 5.3 for context on how OOI produces these variables). All b_{bp} <0 or >0.01 m^{-1} are flagged as "suspect or of high interest", and all total optical backscattering <0 or >1 m^{-1} are flagged as "suspect or of high interest". Since scatterers such as barnacles have a stronger effect on the backscatter signal than on fluorescence, but are also expected to confound the fluorescence signal, fluorescence data (both for *chlorophyll-a* and FDOM) should also be flagged where the backscatter signal failed this test. For FDOM, no automated flagging is performed and users will need to use their own judgment to devise limits for this test.

Climatology test: The climatology test is similar to the Gross Range test, but the bracket of allowable values is based on local knowledge of the area where a mooring or glider is deployed, and so will be smaller. For *chlorophyll-a* data it is also advised to do this test on data that has had physiological adjustments done (see Section 5.5.5 on physiological adjustments below), so that the "best *chlorophyll-a* estimate" is used. All data outside a range of values appropriate for the region are flagged as "suspect or of high interest".

Spike test: The spike test is implemented to identify spikes in data from fixed depth platforms, presumably caused by interference from bio-fouling or fish. It is currently not standard practice to flag spikes in profiling data as these spikes are considered valuable data, useful in identifying larger particles or aggregates (e.g., of phytoplankton or detrital matter; Briggs et al., 2011; 2020). In data from moored sensors, such particles could cause spikes in single measurements in a burst. However, assuming that a burst does not last longer than a few seconds and given the small field of view of the sensor, it is highly unlikely that a particle remains in the field of view for the duration of a burst. Therefore, if burst data have been pooled, for example by using the median, then any remaining spikes are not expected to be due to particles passing through the field of view of the sensor and thus are not considered data of interest, but rather data anomalies that should be removed.

Due to the reasons outlined above, the spike test is recommended to be run on pooled (i.e., not burst-level) data. The spike test can be implemented on any time window deemed appropriate, determined by the native time resolution of the data set. In the below example of an implementation, we assume that data were acquired hourly.

1. A running median and standard deviation with a window size representative of 1 day is calculated (e.g., 23 or 25 data points for hourly data) using only data not previously flagged as bad or probably bad. The window should be centered on the individual data point of interest (e.g. 11 or 12 measurements on either side of the individual point for hourly data).
2. The difference between each individual data point and the running median for the respective point is calculated.
3. The test is failed if the absolute value of the difference is greater than 3 times the standard deviation over the time window.

4. Manually inspect data to identify and address commonly-known issues. Bio-optical data users should check for evidence of biofouled sensors, and in particular for *chlorophyll-a* data for diel variability due to photo-physiological variability (Section 5.4.4). While some examples of biofouling will be flagged by OOI-provided Data Annotations, variability due to physiological processes will not, and users should carefully inspect all data in preparation for their own analysis. At this stage, “cleaned” FDOM and total optical backscatter data will be ready for comparison with co-located data and analysis, but *chlorophyll-a* and particulate backscatter data products will require further processing as outlined in the steps below.

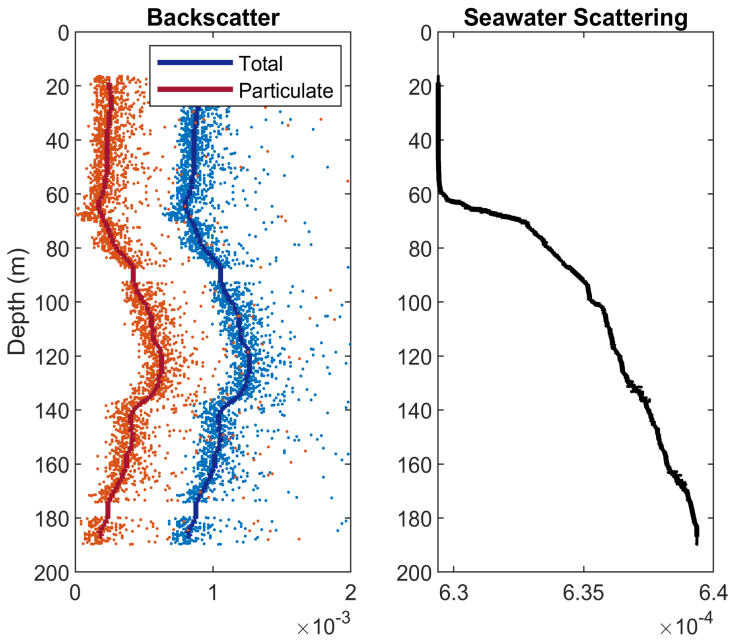


Figure 5.3 Application of the seawater backscatter correction to derive particulate backscatter from OOI-provided total optical backscatter. An example profile of total optical backscattering (blue, FLUBSCT-L2) and particulate backscattering (red), are plotted in the left (dots are all data points, lines are the 200-point running median). The seawater scattering coefficient (black) is plotted on the right. These data were collected on March 10, 2021 by the Regional Cabled Array Oregon Slope Base Shallow Profiler Mooring.

5.5.3 Particulate backscatter correction

OOI provides data on total backscatter (FLUBSCT), that has not had the backscatter due to seawater removed. In order to derive particulate backscatter data, the user will need to subtract the seawater backscatter contribution from the total optical backscatter (see example in Figure 5.3).

5.5.4 Dark correction for chlorophyll-a

The dark reading of the fluorometer sensor *in situ* is frequently slightly different from the dark measured in the factory. Therefore, if the data allow such a correction, the *chlorophyll-a* fluorescence signal is corrected with an *in situ* dark value. At the moment, **this correction is only possible on data from profiling sensors**, not those mounted at a fixed depth. The correction can be done on either L0 or L1 data, with implementation on L1 data slightly less complicated as there is no need to go through an updated calibration process.

For fluorometers on profiling platforms (including gliders), an *in situ* dark can be estimated from deep profiles. Preferably, a dark estimate should only be attempted on profiles that reach deeper than 950 m, however shallower profiles can also be considered if they're well below the mixed layer and the user is confident that *chlorophyll-a* at the deepest depths measured is effectively zero. The dark value on any suitable profile is the minimum *chlorophyll-a* fluorescence measured on that profile (at any depth; it does not have to be below 950 m). The *in situ* dark value for a deployment can be estimated based on a number of profiles and then assumed to be constant. One recommendation is to take the median of the first five darks that are estimated for a sensor deployment (i.e., from the first five profiles that were sufficiently deep) and then apply that dark value for the whole time series of the deployment. This is the latest approach recommended for *chlorophyll-a* data from BGC-Argo floats (the documentation is in the process of being updated). Then, as a QC measure, one can check the dark over time and see how much it drifts. On a well-functioning sensor, the dark value should not drift significantly, so any significant drift is a flag for a potential problem, e.g., due to biofouling.

Once a dark has been estimated for a deployment, it should be subtracted from all *chlorophyll-a* fluorescence data. In principle, the backscatter and FDOM measurements may also suffer from a discrepancy between the factory and *in situ* darks, but there is no straightforward way to correct for that as neither backscatter nor FDOM can be assumed to be zero at depth.

5.5.5 Physiological adjustments for *chlorophyll-a*

Two types of physiological adjustments may be required for data collected by the fluorescence sensor:

- For daytime data, a non-photochemical quenching (NPQ) correction may need to be applied
- In order to improve the absolute *chlorophyll-a* estimate, the sensor may need to be compared against other data sources (*in situ* samples, satellites, independent *chlorophyll-a* estimates based on light attenuation, e.g, from PAR sensors). If this isn't

possible and the sensor is an ECOpuck, then the “Roesler factor” should be applied, i.e., a division by 2 to adjust the *chlorophyll-a* estimate for the global bias observed for ECOpuck fluorometers (Roesler et al., 2017).

NPQ correction

For fixed-depth mooring data, it is recommended that only night-time data should be used. Alternatively one can make use of the backscatter signal, for example one can estimate the ratio between *chlorophyll-a* fluorescence and backscatter at night, then apply that ratio to the daytime backscatter data to estimate *chlorophyll-a* fluorescence (Thomalla et al., 2018). That way, one gets an around-the-clock *chlorophyll-a* estimate that is not affected by NPQ.

For profile data, an NPQ correction has been described by Xing et al. (2012, 2018), and is widely used to correct BGC-Argo data, with some modifications (Schmechtig et al., 2018) that can be applied to day and night time data. However, non photochemical quenching correction relies on several assumptions and knowledge of the dynamics of a given environment, namely the structure of the phytoplankton distribution in the water column (e.g., existence of deep chlorophyll maxima). While the NPQ correction is beyond the scope of this document, we provide some references above and encourage the user to read them to decide the best correction method for a given dataset.

5.5.6 Validate *chlorophyll-a* fluorescence data with discrete samples

Discrete samples (collected in Niskin bottles on the CTD rosette) are collected during OOI turn-around cruises in proximity to OOI *in situ* sensors. From these samples, comparable measurements can be made that are independent analogs of OOI *in situ* sensor data. Guidelines for how close discrete measurements need to be compared to fluorometry sensors varies by oceanographic region. In the case of *chlorophyll-a*, concentrations are determined from discrete water samples using High Performance Liquid Chromatography (HPLC) laboratory analysis (Van Heukelem et al., 2001). The extent to which HPLC *chlorophyll-a* concentrations can be used for validation of fluorometry depends on the proximity of the measurements in space and time. With the low temporal resolution of cruises, calibration curves (over time) are not possible in relation to other sources of variability. However, discrete sample measurements can be used to estimate sensor drift over time, as described below.

Getting to “true” *chlorophyll-a*

Even though *chlorophyll-a* fluorescence stems from *chlorophyll-a*, it is not necessarily an accurate indicator of *chlorophyll-a in vivo* because conversion of fluorescence measurement to *chlorophyll-a* concentration can vary by about factor 10 in nature due to species composition, light acclimation and nutrient status of the resident phytoplankton (Roesler et al., 2017). The widely used Wetlabs ECOPuck instrument was (and is) calibrated against a standard that refers back to a “lab rat” phytoplankton culture of *Thalassiosira weissflogii*. Globally, this leads to an average over-estimation of *chlorophyll-a* concentration by a factor of 2, with strong regional variability (Roesler et al., 2017). To improve a given *chlorophyll-a* estimate based on fluorescence measurements, it is thus advised - for open-ocean deployments, not necessarily

near the coasts - to divide the *chlorophyll-a* concentration that the sensor calibration produces by factor 2. This is standard practice for BGC-Argo data (Schmechtig et al., 2018). Note that additional studies for a given region may have led to a regional estimate of the “Roesler factor” for a given mooring. Additionally, data from other sensors or from discrete samples can be used to calibrate/cross-check fluorescence-derived *chlorophyll-a* data. These include:

- ➔ A co-located ac-s (hyperspectral absorption and attenuation), from which the absorption line-height at 676 nm can be derived, which can be used to estimate *chlorophyll-a* concentration (Boss et al., 2013; Roesler & Barnard, 2013). When possible, it is recommended to calculate the relationship between line-height and *chlorophyll-a* from HPLC for a given deployment location, as the coefficients of this relationship can change in different water types.
- ➔ Profile data from irradiance sensors, as they allow estimation of downwelling light attenuation (K_d), based on which *chlorophyll-a* can be estimated (Morel et al., 2007; Xing et al., 2011)
- ➔ Satellite data, which can provide an independent comparison via match-ups of estimated *chlorophyll-a* concentration from remote-sensing reflectance in time and space against which data from a fluorometer can be compared.
- ➔ Other co-located assets that carry well-calibrated fluorometers (e.g., gliders, floats, profilers)

5.6 Worked Example

Here we provide an example of how to apply the end-user processing workflow described in Section 5.5 and summarized in Figure 5.2 to *chlorophyll-a* fluorescence data. This example walks through the processing of Level 1 (L1) fluorometer data (L1 CHLAFLO, processed by OOI as shown in Figure 5.1) from a Shallow Profiler on the OOI Regional Cabled Continental Margin Array located at the ‘Oregon Slope Base’ site collected during June 2018 (see Figure A.5 for Regional Cabled Continental Margin Array map schematic).

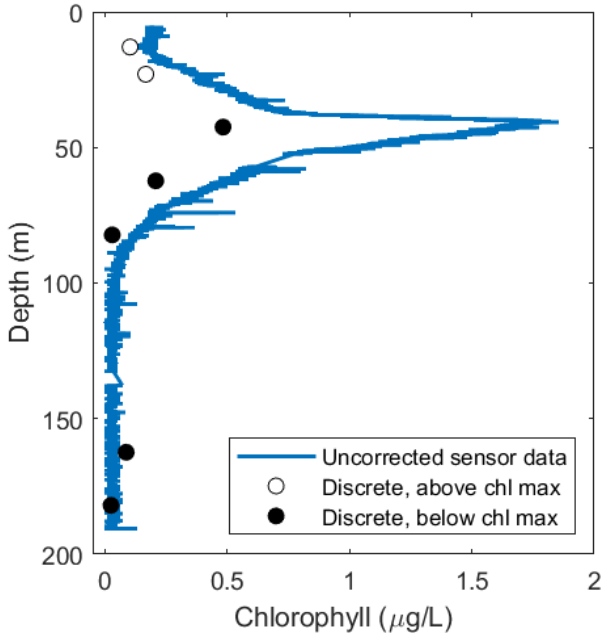


Figure 5.4. Comparison between *chlorophyll-a* fluorescence and *chlorophyll-a* concentration from bottle samples. Shallow Profiler L1 fluorometer (blue) versus lab-analyzed HPLC *chlorophyll-a* from discrete bottle samples data (filled and open circles). Observations were co-located and simultaneous at noon on June 26th 2018.

The L1 CHLAFLO *chlorophyll-a* fluorescence data is compared to the concurrently collected *in situ chlorophyll-a* data from discrete samples analyzed by HPLC (Figure 5.4). Although the shape of the *chlorophyll-a* profile is largely consistent between the two data sets, their comparison reveals a significant offset between the sensor data and the discrete sample data. The offset is apparent at all depths, with the sensor overestimating *chlorophyll-a* concentrations by a factor of 2 or more in the deep chlorophyll maximum at around 40m depth.

The next step is to apply the dark correction to the *chlorophyll-a* fluorescence sensor data. The method for estimating the instrument dark value is outlined in detail in Section 5.5.4. Here we estimated the instrument dark by taking the minimum sensor *chlorophyll-a* value for the profile, and then subtracting it from all of the *chlorophyll-a* sensor data for the profile (Figure 5.5).

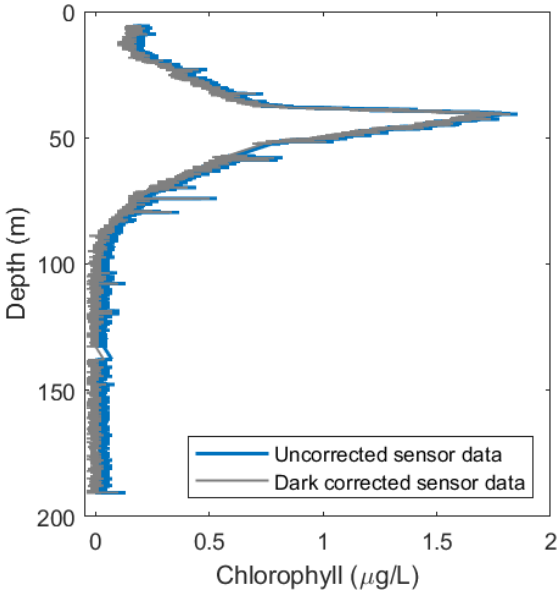


Figure 5.5. Application of the dark correction to profiling *chlorophyll-a* fluorescence sensor data. In this example, the instrument dark value was estimated based on the minimum *chlorophyll-a* fluorescence value for the profile, and then subtracted from the entire profile..

Once the dark correction has been applied, users should assess what data is likely quenched, and flag those data (e.g., daytime data shallower than the *chlorophyll-a* max depth). In this worked example, the data above the chlorophyll maximum is flagged as potentially quenched data as the cast was taken during the day. We then apply the slope correction based on the co-located discrete sample data. Remember that the slope correction should be determined using only data that is likely not impacted by quenching, i.e., using the quenching filter to determine the regression slope (Figure 5.6). The slope is estimated by applying a linear regression with the intercept forced through zero in order to maintain zero and non-negative *chlorophyll-a* concentrations in the corrected *chlorophyll-a* sensor data at depth. Then, apply the calculated slope to the entire dataset, including the likely quenched data. If the user decides to correct the data for NPQ, it should be done now, only after the *in situ* calibration has been done (i.e., the slope has been applied to the sensor data). This NPQ correction is beyond the scope of this document but some references for existing methods are provided in section 5.5.5.

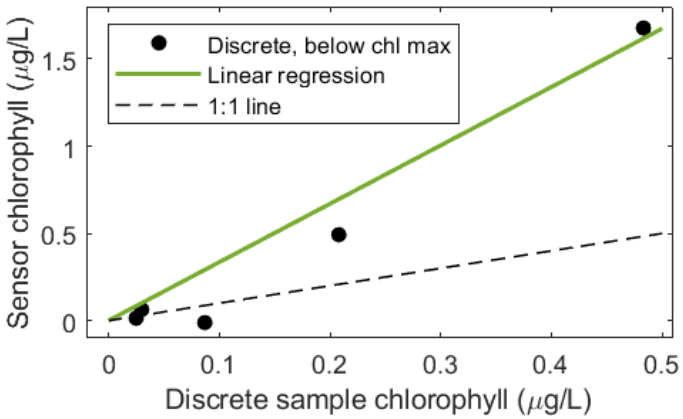


Figure 5.6. Linear regression of the bottle *chlorophyll-a* and sensor *chlorophyll-a* data using match-ups based on depth. A quenching filter has been applied (i.e., only data below the chlorophyll maximum depth are used) to determine the linear regression relationship.

In Figure 5.7, we show the sensor data after it has had all of the above corrections applied. This example clearly shows the extent to which the original uncorrected L1 *chlorophyll-a* data downloaded from OOI overestimated *chlorophyll-a* concentrations throughout the top 100m of the profile. This dataset is now analysis-ready.

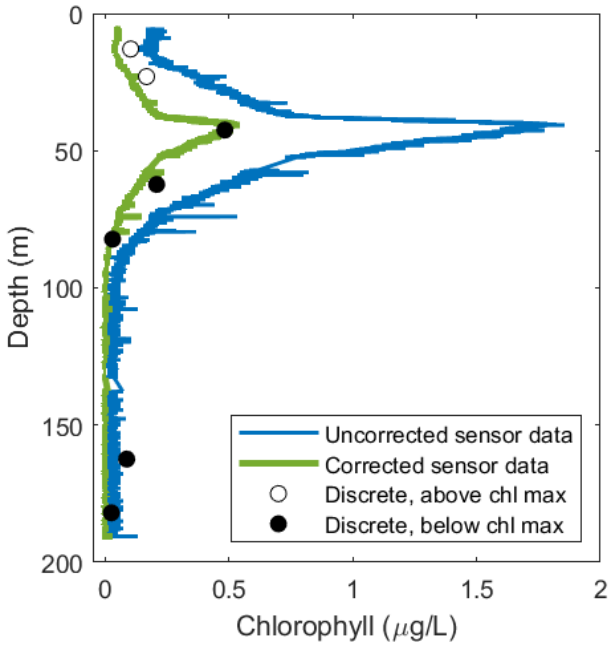


Figure 5.7. Corrected chlorophyll-a sensor data plotted along with the original uncorrected *chlorophyll-a* sensor data and the discrete bottle sample *chlorophyll-a* data. The corrected data was produced by applying each of the steps described in this worked example.

Pseudo-Code

The pseudo-code provided below provides each step in the data processing pipeline for this worked example, with steps organized following the sequence given in Figure 5.2 (End User Data Bio-Optics Data Processing Flow) and the text in Sections 5.4 and 5.5. This pseudo-code is intended to support end users in developing their own data processing sequence following these recommended steps using any programming language or OOI data access method of their choice.

Assemble Data

Review available FLORT/FLORD/FLNTU/FLCDR data on Data Explorer;
Review OOI HITL annotations for optical sensors;

Download sensor data of interest including time, depth, asset type and location;

Download co-occurring turn-around cruise discrete chlorophyll-a data from HPLC analyses;

Evaluation and cleaning

Plot the sensor data;
Note and evaluate discontinuities and any other obvious issues;

Apply OOI QC flags to available sensor and discrete sample data;
Apply OOI Human In The Loop (HITL) annotations to available sensor and discrete sample data;

Apply QARTOD flags to available data;

Apply Manual Cleaning to available data (Except for Chlorophyll-a Quenching, which is applied later): Refer to Common Data Quality Issues in Chapter 5: Bio-Optical Measurements

Filter or discard data, as appropriate.

Sensor data are now “clean”.

FDOM

Analysis Ready FDOM: No additional steps are needed

Total Optical Backscatter

Analysis Ready Total optical backscattering: (bb, at 700 nm) this is an OOI L2 product and no additional steps are needed

Particulate Backscatter

Calculate seawater backscattering (bb_sw) as a function of matching temperature, salinity (following Zhang et al 2009);

Subtract the seawater backscattering (bb_sw) from total optical backscattering (bb) to derive Analysis Ready Particulate Backscatter using the equation below

bb_p = bb - bb_sw

Reminder: bb_p is not an OOI data product.

Chlorophyll-a: dark correction

Reminder: the dark correction is only applicable to chlorophyll-a fluorescence on profiling sensors

Requirements for Chlorophyll-a: Sensor is profiling and there are sufficient profiles (at least 5) of required depth below the productive layer:

Calculate the instrument specific dark value (inst_dark):

inst_dark = median[min(CHL) from first 5 profiles];

Apply dark correction to the filtered dataset:

CHL_dark_corrected = CHL - inst_dark;

Chlorophyll-a: quenching flags

Apply a filter for (daytime) quenched Chlorophyll-a data:

Simplest method:

daytime = hours between sunrise and sunset;
Be sure to use asset local time.

Second, more complex method:

use PAR data from a collocated light meter (surface mooring) to identify daytime.
daytime = hours when PAR > 0;

Identify depth of daytime Chlorophyll-a max for per asset and per day.

Flag daytime Chlorophyll-a data acquired at depth less than the Chlorophyll-a max depth.

Note: these Chlorophyll-a data are biased by non-photochemical quenching and must be filtered from the dataset before determining the slope calculation to be applied later to the Analysis Ready Chlorophyll-a Fluorescence, and then returned to the final dataset. These data biased by photoquenching should NOT BE REMOVED from the

overall dataset. They contain information on phytoplankton physiology.

Chlorophyll-a: slope correction

Note: Mooring data with fixed depth fluorometers will have fewer comparative sample points than profilers and gliders, and may fall within the quenching depth filter, preventing any matchups.

Plot by depth: Discrete HPLC chlorophyll-a data against contemporaneous OOI chlorophyll-a fluorescence sensor data.

If any, FIND matchups between quenching-filtered sensor data (Chlorophyll-a) and QC'd OOI discrete sample data, then calculate slope/offset corrections based on fit. (note best quality is achieved with >3 matchups). When the dynamic range in chlorophyll from the bottled samples is not regularly spaced (e.g. lots of points at low concentrations and one point at high), it is recommended to calculate the regression in log-log space. So calculate the linear regression between log10(chlorophyll fluorescence) and log10(chlorophyll concentration - from bottles).

Apply slope corrections to the entire dataset. In this case, it is preferable to calculate the slope based on a zero intercept:

```
CHL_adj = m * CHL + 0;

where m is the slope of the regression line, and the
intercept is forced to 0.
```

CHL_adj is the analysis ready chlorophyll-a fluorescence dataset. IF appropriate, users may now apply any NPQ correction method to remove and correct quenching effects on daytime data (not described here).

Plot analysis-ready chlorophyll-a fluorescence data

Appendix

A1. OOI websites with key information

Observation and Sampling Plan

This document contains details on the OOI sampling approach, including the arrays, platforms and sensors.

Instruments

Information about each instrument in the OOI array can be found here. This includes makes and models, codes for accessing data, and data product sheets (DPS) that detail instrument metadata, theory of operation, and implementation. DPS are often helpful in the QA/QC process.

OOI Cruise Data and Array Documentation (Alfresco server)

Data from OOI cruises is distributed via the Alfresco repository. Discrete bottle samples, underway measurements and CTD casts may all be helpful for QA/QC. This repository also includes vendor documentation on delivery, recalibration, and refurbishment of instruments, OOI's internal design documents and pre- and post-deployment testing documentation for all instruments and platforms.

OOI Discourse Server

OOI supports an online Discourse discussion forum for users of OOI data to share science, data tips, and answer questions about how to most effectively use OOI data in research and in the classroom. The News and HelpDesk channels are a great place to find updates on OOI data platforms and to see questions posted by other users. There is also a Data Tools channel where OOI staff and users can share tools that they have developed to work with OOI data streams.

How to use and cite OOI data

A2. Terminology

OOI Glossary

Definitions used within the OOI program for infrastructure, types of platforms and nodes, data, and general terminology.

OOI-specific Acronyms used in this document

General Acronyms used in this document

A3. Additional external-to-OOI resources on QA/QC and BGC sensors

IOOS QARTOD

The United States Integrated Ocean Observing System® (U.S. IOOS®) Quality Assurance / Quality Control of Real-Time Oceanographic Data (QARTOD) Project. Details on the quality control steps for evaluating real-time ocean data. QARTOD tests are in the process of being implemented by OOI, but additional tests may be recommended for the user.

Ocean Best Practices Repository

A repository of published methods and best practices documents hosted by the International Oceanographic Data and Information Exchange (IODE). An open-access repository containing community accepted practices towards collection, processing and quality control of oceanographic data. This searchable resource contains recommended best practices for both ship-borne and autonomous ocean data collection, including by biogeochemical sensors.

Online training course on biogeochemical sensors

IOCCP & BONUS INTEGRAL Training Course on "Instrumenting our oceans for better observation: a training course on a suite of biogeochemical sensors"

OceanGliders community

OceanGliders is the glider component of the integrated Global Ocean Observing System. This repository facilitates sharing and discussing Best Practices for gliders, including processing and management of data from biogeochemical sensors.

A4. Code toolboxes

OOI Official GitHub

The OOI provides public access to Matlab, Python, R, and Julia code examples and packages used for all data processing. Many of these repositories were referenced directly within the Best

Practices & User Guide, but users are encouraged to peruse the full set of these publicly-available repositories.

OOI Data Explorations GitHub Repository

Within the OOI official GitHub, the OOI Data Explorations repository contains Matlab, Python, R and limited Julia code that users can use to access OOI data through the M2M system. The Python code also has tools for accessing data from the OOI Gold Copy THREDDS catalog and code that applies additional processing such as additional data QC checks to certain data sets, and code that shows some of the QARTOD workflows.

OOI Data Labs GitHub Repository

This repository provides a series of short Jupyter Notebooks to get people started with coding in Python using OOI data. These notebooks are very accessible and were developed with undergraduates in mind. They include requesting OOI data, merging datasets, basic statistics and plotting. Different folders were developed for different workshops/purposes. An explainer of the repository can be found here. Unless otherwise noted, all examples were developed by Sage Lichtenwalner, at the Rutgers University Center for Ocean Observing Leadership.

Biogeochemical Argo Tools

Biogeochemical Argo is a global network of autonomous profiling floats, which uses many of the same BGC sensors as OOI. BGC Argo has developed excellent code libraries in multiple programming languages (Python, R, and MATLAB) for data access, visualization, and quality control that OOI BGC sensor data users may find helpful.

A5. GitHub repository for Worked Examples

All code and calculations used in processing the Worked Examples in the individual chapters of this Best Practices & User Guide are provided as a supplementary resource in this GitHub repository. The repository includes separate folders with the Worked Example calculations for each chapter, which parallel the pseudo-code included in the chapters. This data processing code is intended solely as a reference and not as a template for end user data processing.

Since our intention in this document is to support end users who may choose to access OOI data using a range of approaches and complete their end user data processing in any coding language or calculation approach, the Worked Examples across Chapters 2-5 illustrate a variety of possible approaches:

- The Carbonate Chemistry and Bio-Optics Worked Examples (Sections 4.6 and 5.6) are implemented in Python Jupyter notebooks
- The Dissolved Oxygen Worked Example (Section 2.6) is implemented in MATLAB
- The Nitrate Worked Example (Section 3.6) is implemented as spreadsheet calculations in Microsoft Excel workbooks.

Although this document is not intended as a guide for how to access OOI data, we also include in this repository two supporting examples of current methods for accessing OOI data, as applied to OOI BGC datasets:

- The Nitrate Data Download Example provides a step-by-step walkthrough of how the dataset used for the Nitrate Worked Example (Section 3.6) was downloaded from OOI's Data Explorer
- The Carbon Data Download Jupyter Notebook illustrates two methods for programmatically accessing and downloading OOI data, as applied in the Carbon Worked Example (Section 4.6). The first method utilizes OOI's API to perform machine-to-machine (M2M) queries for data from the OOI THREDDS data server, which serves the same datasets which can be accessed via OOI's Data Portal. The second method requests data from OOI's Data Explorer ERDDAP server.

A6. OOI for Teaching

The OOI Ocean Data Labs Project

The Ocean Data Labs Project has developed numerous resources for classroom and research use. The target audience is geoscience professors of undergraduates but many resources may also be relevant for high school and graduate students. Resources include 15 minute guided data explorations with online widgets, "data nuggets" which are cleaned up and interesting data chunks ready for use, Python code to get students started in research, individual lesson plans using OOI data, and a complete lab manual ready for implementation in Introductory Oceanography classes. The team also runs a blog, webinars and workshops on how to use OOI data in the classroom.

Project EDDIE

Project EDDIE (Environmental Data-Driven Inquiry and Exploration) is an NSF funded project to bring data into the classroom. They have developed a series of teaching models, some of which include OOI data such as this one investigating bomb cyclones.

A7. Schematic drawings of OOI Arrays

The following figures (A.1-A.6) provide schematic drawings of the OOI arrays, with captions identifying the locations within each array of all BGC sensors covered in this Best Practices & User Guide. See Table 1.1 for a synthesis of all BGC sensors included in the scope of this document and numbers deployed across all OOI arrays.

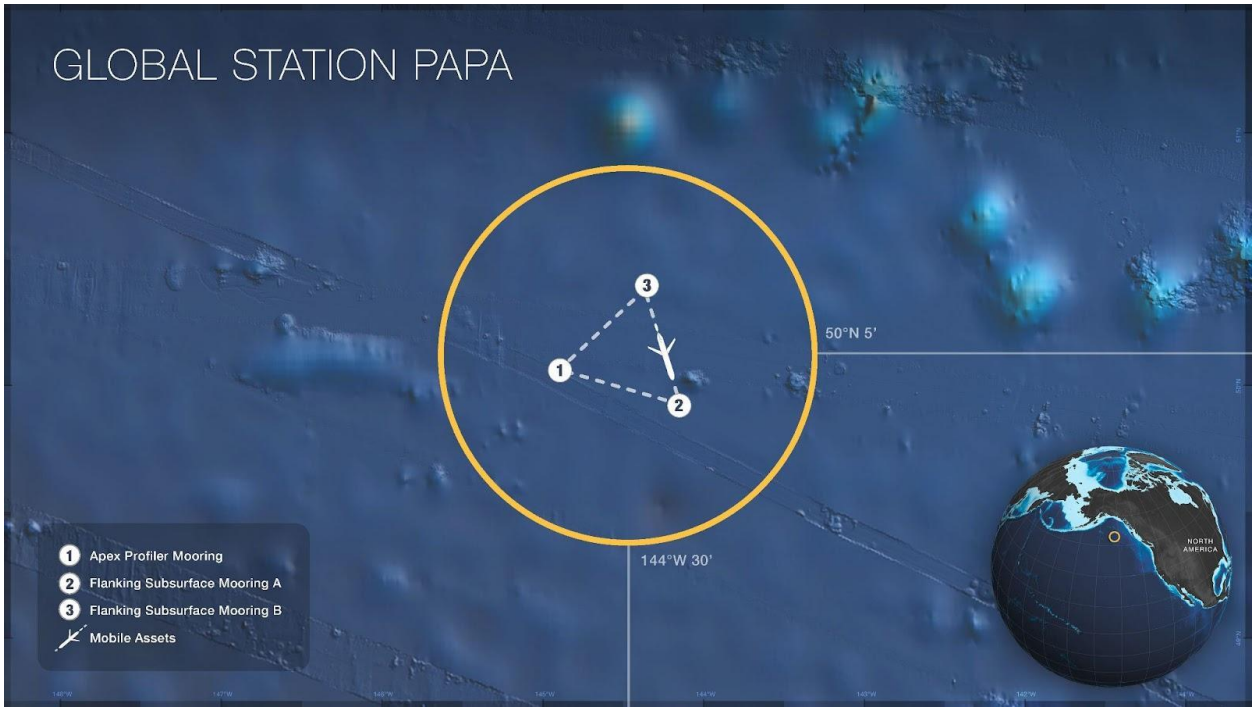


Figure A.1. Schematic drawing of the OOI Global Station Papa Array. All platforms include BGC sensors:

(1) Apex Profiler Mooring: Oxygen and Fluorescence/Chlorophyll-a on shallow (≈150-2100m) and deep (2150-4100m) profilers.

(2) and (3) Flanking Moorings A & B: Oxygen, Fluorescence/Chlorophyll-a, and pH at 30 m.

Mobile Assets: Oxygen and Fluorescence/Chlorophyll-a on Open Ocean Gliders (0-1000 m); Oxygen, Fluorescence/Chlorophyll-a, and Nitrate on Global Profiling Gliders (0-200 m).

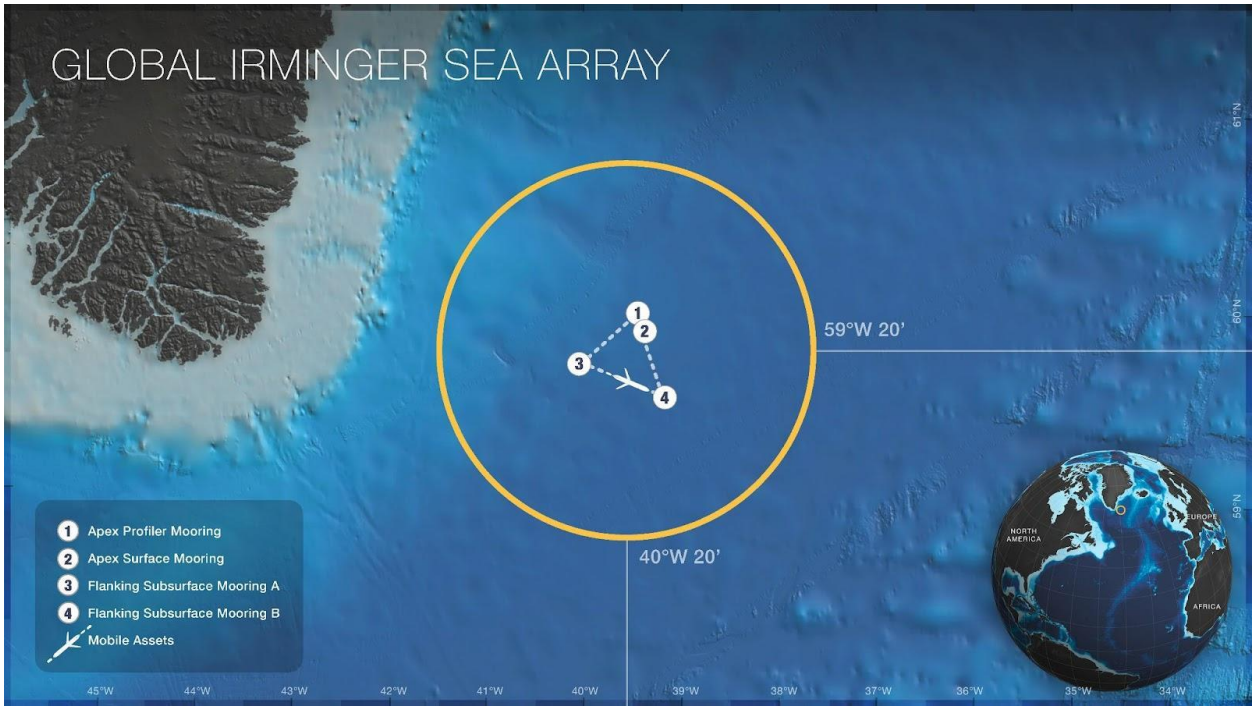


Figure A.2. Schematic drawing of the [OOI Global Irminger Sea Array]. Also applicable to the [Argentine Basin Array] (active 2015-2018) and [Southern Ocean Array] (active 2015-2020). All platforms include BGC sensors:

(1) Apex Profiler Mooring: Oxygen and Fluorescence/Chlorophyll-a on the shallow (≈ 150-2100m/2600 m at Irminger) and deep (≈2150-4100m, not at Irminger) profilers.

(2) Apex Surface Mooring: pCO_2 and Oxygen at surface, 12, 40, 80, 130 m; pH at 20 & 100 m; Fluorescence/Chlorophyll and Nitrate at surface & 12 m.

(3) & (4) Flanking Moorings A & B: Oxygen, Fluorescence/Chlorophyll, and pH at 30 m.

Mobile Assets: Oxygen and Fluorescence/Chlorophyll on Open Ocean Gliders (0-1000 m); Oxygen, Fluorescence/Chlorophyll, and Nitrate on Global Profiling Gliders (0-200 m).

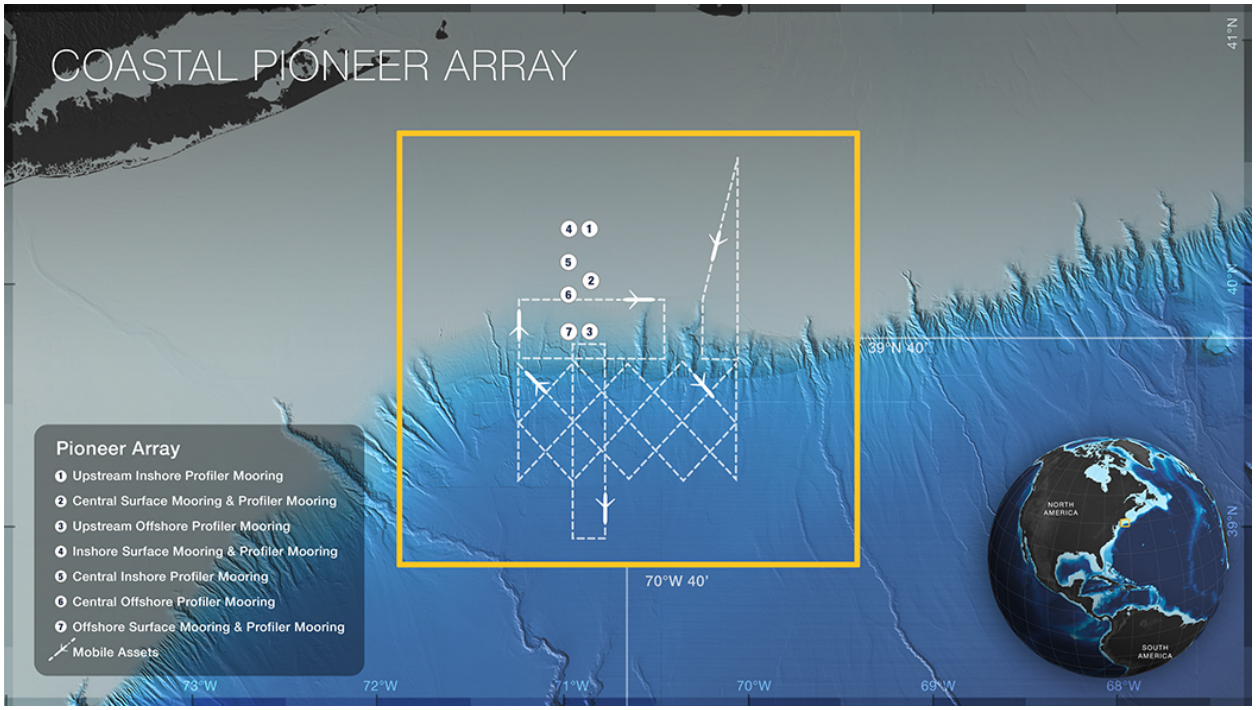


Figure A.3. Schematic drawing of the [OOI Coastal Pioneer Array]*. All platforms include BGC sensors:

Inshore (1), Central (2), & Offshore (3) Surface Moorings: Surface buoy - pCO_2; Near-surface instrument frame (NSIF) at 7 m - pH, Oxygen, Fluorescence/Chlorophyll, Nitrate; Multi-function node (MFN) at the bottom - pH, pCO_2, Oxygen.

Inshore (4), Central Inshore (5), Central Offshore (6), & Offshore (7) Profiler Moorings: Oxygen and Fluorescence/Chlorophyll (25 m to 20 m above bottom).

Mobile Assets: Coastal Gliders (0 - 200 m, 0 - 1000 m or 10 m above bottom) - Oxygen and Fluorescence/Chlorophyll; Coastal Profiling Gliders (0 - 200 m or 10 m above bottom) and AUVs - Oxygen, Fluorescence/Chlorophyll, Nitrate.

*New England Shelf location, deployed 2016-2022; the Coastal Pioneer Array will move to the Southern Mid-Atlantic Bight with a new array configuration beginning in 2024

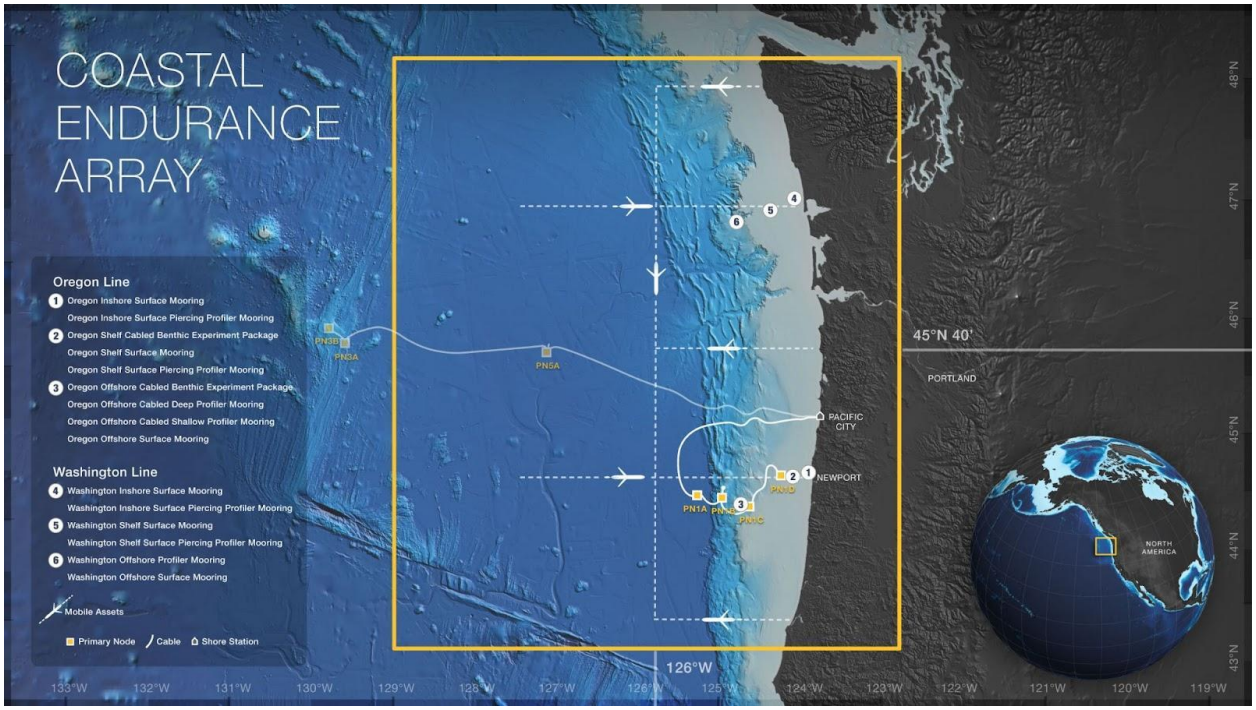


Figure A.4. Schematic drawing of the [OOI Coastal Endurance Array]. All platforms include BGC sensors:

Inshore Surface Moorings (1, 4): Surface Buoy - Fluorescence/Chlorophyll; Near-surface instrument frame (NSIF) at 7 m - pCO_2, pH, Oxygen, Fluorescence/Chlorophyll, Nitrate; Multi-function node (MFN) at the bottom - pH, pCO_2, Oxygen

Coastal Surface Moorings (2, 3, 5, 6): Surface Buoy - pCO_2; Near-surface instrument frame (NSIF) at 7 m - pH, Oxygen, Fluorescence/Chlorophyll, Nitrate; Washington Line (5, 6) Multi-function node (MFN) at the bottom or Oregon Line (2, 3) co-located Cabled Benthic Experiment Package - pH, pCO_2, Oxygen

Coastal Surface-Piercing Profiler Moorings (1, 2, 4, 5): Oxygen, Fluorescence/Chlorophyll, Nitrate (surface to bottom)

Cabled Profiler Moorings (3): Shallow (20 to 200 m) - pCO_2, pH, Oxygen, Nitrate, Fluorescence/Chlorophyll; Deep (175 to 500 m) - Oxygen, Fluorescence/Chlorophyll; Fixed at 200 m - pH, pCO_2, Oxygen

Wire-Following Profiler Mooring (6) - Oxygen, Fluorescence/Chlorophyll (15 to 540 m)

Mobile Assets: Coastal Gliders (0 - 200 m, 0 - 1000 m or 10 m above bottom) - Oxygen and Fluorescence/Chlorophyll.

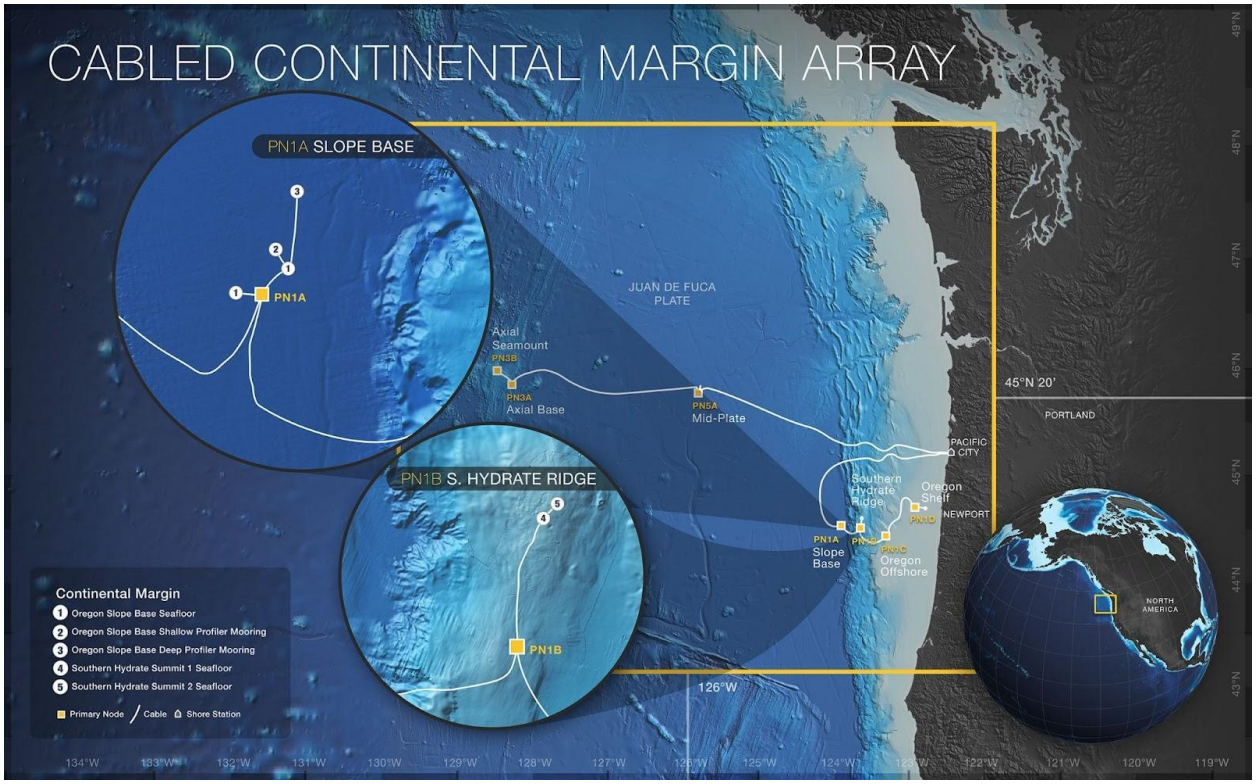


Figure A.5. Schematic drawing of the [OOI Regional Cabled Continental Margin Array]. Only the Oregon Slope Base platforms contain BGC sensors, including:

(1) Oregon Slope Base Seafloor: Oxygen

(2) Oregon Slope Base Shallow Profiler Mooring: pCO_2, pH, Oxygen, Fluorescence/Chlorophyll, Nitrate (5 to 200 m); Fixed at 100 m - pH, Oxygen, Fluorescence/Chlorophyll

(3) Oregon Slope Base Deep Profiler Mooring: Oxygen, Fluorescence/Chlorophyll (200 to 2,905 m)

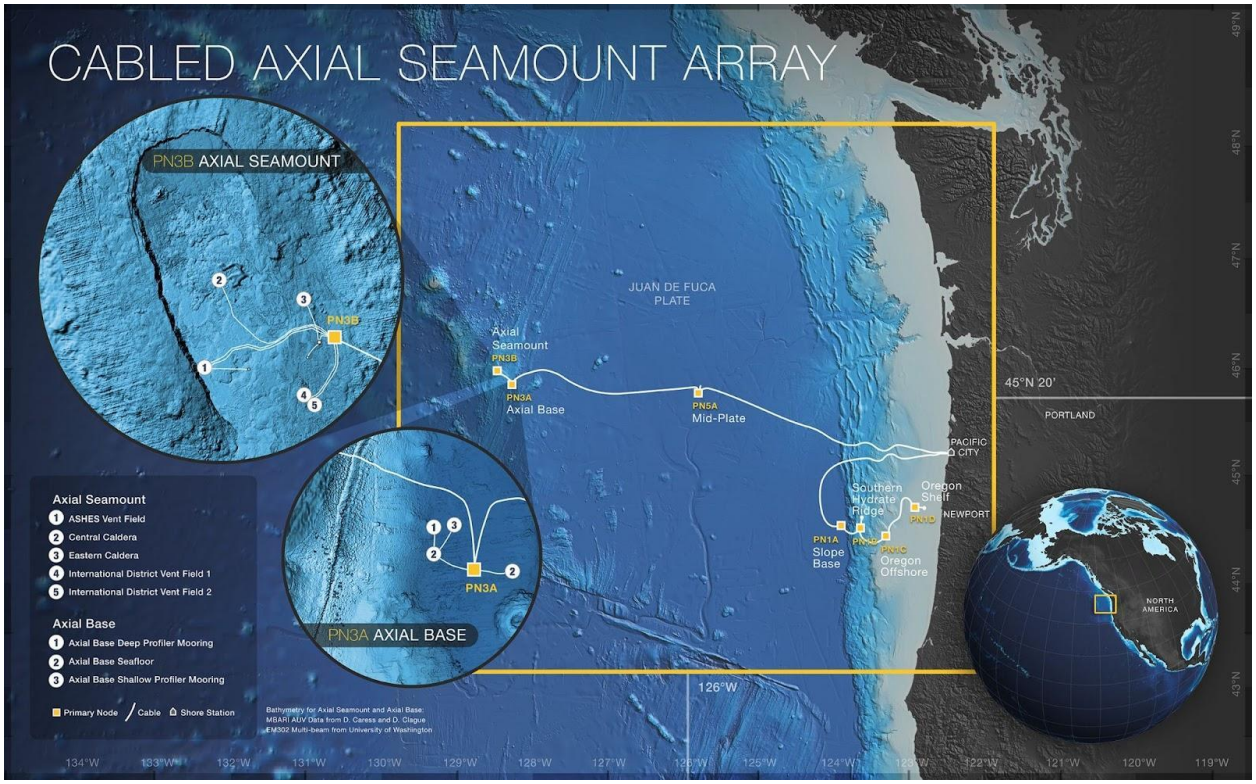


Figure A.6. Schematic drawing of the [OOI Regional Cabled Axial Seamount Array]. Only the Axial Base platforms contain BGC sensors, including:

(1) Axial Base Deep Profiler Mooring: Oxygen, Fluorescence/Chlorophyll (150 to 2,465 m)

(2) Axial Base Seafloor: Oxygen

(3) Axial Base Shallow Profiler Mooring: pCO2, pH, Oxygen, Fluorescence/Chlorophyll, Nitrate (5 to 200 m); Fixed at 200 m - pH, Oxygen, Fluorescence/Chlorophyll

References

Alliance for Coastal Technologies. 2009. Protocols for demonstrating the performance of in situ pCO_2 analyzers. ACT PD09-01. 10 pp http://dx.doi.org/10.25607/OBP-334

Becker, S., Aoyama, M., Woodward, E.M.S., Bakker, K., Coverly, S., Mahaffey, C. and Tanhua, T., 2020. GO-SHIP repeat hydrography nutrient manual: the precise and accurate determination of dissolved inorganic nutrients in seawater, using continuous flow analysis methods. Frontiers in Marine Science, 7, p.581790. https://doi.org/10.3389/fmars.2020.581790

Bittig, H. C., Fiedler, B., Scholz, R., Krahmann, G., & Körtzinger, A. (2014). Time response of oxygen optodes on profiling platforms and its dependence on flow speed and temperature. Limnology and Oceanography: Methods, 12(8), 617-636. https://doi.org/10.4319/lom.2014.12.617

Bittig, H. C., & Körtzinger, A. (2015). Tackling oxygen optode drift: Near-surface and in-air oxygen optode measurements on a float provide an accurate in situ reference. Journal of Atmospheric and Oceanic Technology, 32(8), 1536-1543. https://doi.org/10.1175/JTECH-D-14-00162.1

Bittig, H. C., & Körtzinger, A. (2017). Update on response times, in-air measurements, and in situ drift for oxygen optodes on profiling platforms. Ocean Science, 13(1), 1-11. https://doi.org/10.5194/os-13-1-2017

Bittig, H.C., Körtzinger, A., Neill, C., Van Ooijen, E., Plant, J.N., Hahn, J., Johnson, K.S., Yang, B. and Emerson, S.R., 2018. Oxygen optode sensors: principle, characterization, calibration, and application in the ocean. *Frontiers in Marine Science*, 4, p.429. https://doi.org/10.3389/fmars.2017.00429

Bittig, H. C., Maurer, T. L., Plant, J. N., Schmechtig, C., Wong, A. P., Claustre, H., ... & Xing, X. 2019. A BGC-Argo guide: Planning, deployment, data handling and usage. Frontiers in Marine Science, 6, 502. https://doi.org/10.3389/fmars.2019.00502

Boss, E., Picheral, M., Leeuw, T., Chase, A., Karsenti, E., Gorsky, G., ... & Claustre, H. (2013). The characteristics of particulate absorption, scattering and attenuation coefficients in the surface ocean; Contribution of the Tara Oceans expedition. *Methods in Oceanography*, 7, 52-62. https://doi.org/10.1016/j.mio.2013.11.002

Briggs, N., Perry, M.J., Cetinić, I., Lee, C., D'Asaro, E., Gray, A.M. and Rehm, E., 2011. High-resolution observations of aggregate flux during a sub-polar North Atlantic spring bloom. *Deep Sea Research Part I: Oceanographic Research Papers*, *58*(10), pp.1031-1039. https://doi.org/10.1016/j.dsr.2011.07.007

Briggs, Nathan, Giorgio Dall'Olmo, and Hervé Claustre. 2020. "Major Role of Particle Fragmentation in Regulating Biological Sequestration of CO2 by the Oceans." *Science* 367 (6479): 791–93. https://doi.org/10.1126/science.aay1790.

Broullón, D., Pérez, F. F., Velo, A., Hoppema, M., Olsen, A., Takahashi, T., ... & van Heuven, S. M. (2019). A global monthly climatology of total alkalinity: a neural network approach. Earth System Science Data, 11(3), 1109-1127. https://doi.org/10.5194/essd-11-1109-2019

Bushinsky, S.M., Emerson, S.R., Riser, S.C. and Swift, D.D. (2016), Accurate oxygen measurements on modified Argo floats using in situ air calibrations. Limnol. Oceanogr. Methods, 14: 491-505. https://doi.org/10.1002/lom3.10107

Carter, B. R., Feely, R. A., Williams, N. L., Dickson, A. G., Fong, M. B., & Takeshita, Y. (2018). Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate. *Limnology and Oceanography: Methods*, 16(2), 119-131. https://doi.org/10.1002/lom3.10232

Claustre, H., Johnson, K. S., & Takeshita, Y. (2020). Observing the global ocean with biogeochemical-Argo. Annual review of marine science, 23-48. https://doi.org/10.1146/annurev-marine-010419-010956

Cullen, J. J. (1982). The deep chlorophyll maximum: comparing vertical profiles of chlorophyll a. *Canadian Journal of Fisheries and Aquatic Sciences*, 39(5), 791-803. https://doi.org/10.1139/f82-108

DeGrandpre, Al. D., Hammar, T. R., Smith, S. P., Sayles, F. L., (1995), In situ measurements of seawater pCO_2 , *Limnology and Oceanography*, 40m, doi: 10.4319/lo.1995.40.5.0969.

DeGrandpre, Mike D., Spaulding, Reggie S., Newton, Jenny O., Jaqueth, Emma J., Hamblock, Sarah E., Umansky, Andre A., Harris, Katherine E., (2014), Considerations for the measurement of spectrophotometric pH for ocean acidification and other studies, *Limnology and Oceanography: Methods*, 12, doi: 10.4319/lom.2014.12.830.

Dickson, A. G., Wesolowski, D. J., Palmer, D. A., & Mesmer, R. E. (1990). Dissociation constant of bisulfate ion in aqueous sodium chloride solutions to 250. degree. C. Journal of Physical Chemistry, 94(20), 7978-7985. https://doi.org/10.1021/j100383a042

Falkowski, P., & Kiefer, D. A. (1985). Chlorophyll a fluorescence in phytoplankton: relationship to photosynthesis and biomass. *Journal of Plankton Research*, 7(5), 715-731. https://doi.org/10.1093/plankt/7.5.715

Fassbender, A. J., Alin, S. R., Feely, R. A., Sutton, A. J., Newton, J. A., Byrne, R. H. (2017). Estimating Total Alkalinity in the Washington State Coastal Zone: Complexities and Surprising Utility for Ocean Acidification Research. Estuaries and Coasts, 40: 404-418. doi: 10.1007/s12237-016-0168-z

Finch, M.S., Hydes, D.J., Clayson, C.H., Weigl, B., Dakin, J. and Gwilliam, P., 1998. A low power ultra violet spectrophotometer for measurement of nitrate in seawater: introduction, calibration and initial sea trials. *Analytica Chimica Acta*, *377*(2-3), pp.167-177. https://doi.org/10.1016/S0003-2670(98)00616-3

Gordon, C., Fennel, K., Richards, C., Shay, L. K., and Brewster, J. K. 2020. Can ocean community production and respiration be determined by measuring high-frequency oxygen profiles from autonomous floats?, Biogeosciences, 17, 4119–4134, https://doi.org/10.5194/bg-17-4119-2020

Jiang, Z. P., Hydes, D. J., Hartman, S. E., Hartman, M. C., & Campbell, J. M. 2014. Application and assessment of a membrane-based pCO2 sensor under field and laboratory. L&O methods, 12, doi:10.4319/lom.2014.12.264.

Jiang, L. Q., Feely, R. A., Wanninkhof, R., Greeley, D., Barbero, L., Alin, S., ... & Townsend, D. W. (2021). Coastal Ocean Data Analysis Product in North America (CODAP-NA)–an internally consistent data product for discrete inorganic carbon, oxygen, and nutrients on the North American ocean margins. *Earth system science data*, 13(6), 2777-2799. https://doi.org/10.5194/essd-13-2777-2021

Johnson, K. and Coletti, L. J. 2002. In situ ultraviolet spectrophotometry for high resolution and long term monitoring of nitrate, bromide and bisulfide in the ocean, Deep-Sea Research I, 49, 1291–130 https://doi.org/10.1016/S0967-0637(02)00020-1

Johnson, K. S., Coletti, L. J., Jannasch, H. W., Sakamoto, C. M., Swift, D. D., & Riser, S. C. (2013). Long-term nitrate measurements in the ocean using the In Situ Ultraviolet Spectrophotometer: sensor integration into the Apex profiling float. *Journal of Atmospheric and Oceanic Technology*, 30(8), 1854-1866. https://doi.org/10.1175/JTECH-D-12-00221.1

Johnson, K. S., Plant, J. N., Coletti, L. J., Jannasch, H. W., Sakamoto, C. M., Riser, S. C., ... & Sarmiento, J. L. (2017). Biogeochemical sensor performance in the SOCCOM profiling float array. Journal of Geophysical Research: Oceans, 122(8), 6416-6436. https://doi.org/10.1002/2017JC012838

Johnson, K. S., Plant, J. N., Riser, S. C., & Gilbert, D. (2015). Air oxygen calibration of oxygen optodes on a profiling float array. Journal of Atmospheric and Oceanic Technology, 32(11), 2160-2172. https://doi.org/10.1175/JTECH-D-15-0101.1

Johnson, K.; Pasqueron De Fommervault, O.; Serr, R.; D'Ortenzio, F.; Schmechtig, C.; Claustre, H. and Poteau Antoine (2018) Processing Bio-Argo nitrate concentration at the DAC Level. Version 1.1, March 3rd 2018. IFREMER for Argo Data Management, 22pp. DOI: http://doi.org/10.13155/46121

Lai, C-Z., DeGrandpre, M.D., Darlington, R.C. (2018). Autonomous Optofluidic Chemical Analyzers for Marine Applications: Insights from the Submersible Autonomous Moored Instruments (SAMI) for pH and pCO_2. Front. Mar. Sci. 4, 438. doi: 10.3389/fmars.2017.00438.

Laruelle, G. G., Landschützer, P., Gruber, N., Tison, J.-L., Delille, B., and Regnier, P. (2017). Global high-resolution monthly pCO2 climatology for the coastal ocean derived from neural network interpolation, Biogeosciences, 14, 4545–4561, https://doi.org/10.5194/bg-14-4545-2017

Lee, K., Kim, T. W., Byrne, R. H., Millero, F. J., Feely, R. A., & Liu, Y. M. (2010). The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans. *Geochimica et Cosmochimica Acta*, 74(6), 1801-1811. https://doi.org/10.1016/j.gca.2009.12.027

López-García,, P., Hull, T., Thomsen, S., Hahn, J., Queste, B.Y, et al (2022) OceanGliders Oxygen SOP, Version 1.0.0. OceanGliders, 55pp. DOI: http://dx.doi.org/10.25607/OBP-1756 . (GitHub Repository, OceanGliders Oxygen SOP. Available: https://oceangliderscommunity.github.io/Oxygen_SOP/sections/authors_SOP_development_process.html)

Lorenzoni, L., M. Telszewski, H. Benway, A. P. Palacz (Eds.), 2017. A user's guide for selected autonomous biogeochemical sensors. An outcome from the 1st IOCCP International Sensors Summer Course. IOCCP Report No. 2/2017, 83 pp. http://dx.doi.org/10.25607/OBP-519

Lueker, T. J., Dickson, A. G., & Keeling, C. D. (2000). Ocean pCO2 calculated from dissolved inorganic carbon, alkalinity, and equations for K1 and K2: validation based on laboratory measurements of CO_2 in gas and seawater at equilibrium. *Marine chemistry*, 70(1-3), 105-119. https://doi.org/10.1016/S0304-4203(00)00022-0

Martz, T.R., Carr, J.J., French, C.R. and M.D. DeGrandpre. (2003). A submersible autonomous sensor for spectrophotometric pH measurements of natural waters, Anal. Chem., 75,1844-1850 https://doi.org/10.1021/ac020568l

Meyer, D., Prien, R. D., Rautmann, L., Pallentin, M., Waniek, J. J., & Schulz-Bull, D. E. (2018). In situ determination of nitrate and hydrogen sulfide in the Baltic Sea using an ultraviolet spectrophotometer. Frontiers in Marine Science, 5, 431. https://doi.org/10.3389/fmars.2018.00431

Morel, A., Claustre, H., Antoine, D. and Gentili, B., 2007. Natural variability of bio-optical properties in Case 1 waters: attenuation and reflectance within the visible and near-UV spectral domains, as observed in South Pacific and Mediterranean waters. *Biogeosciences*, *4*(5), pp.913-925. https://doi.org/10.5194/bg-4-913-2007

Murphy, D. J., Larson, N. G., and Edwards, B. C. (2008). Improvements to the SBE 43 Oxygen Calibration Algorithm. Poster Presentation Ocean Sciences Meeting, Orlando, Florida. https://www.comm-tec.com/Docs/Library/CTD/NewDOEqtnPosterMarch08-4Pages.pdf

Nehir, M., Esposito, M., Begler, C., Frank, C., Zielinski, O., & Achterberg, E. P. (2021). Improved Calibration and Data Processing Procedures of OPUS Optical Sensor for High-Resolution in situ Monitoring of Nitrate in Seawater. Frontiers in Marine Science, 8, Art-Nr. https://doi.org/10.3389/fmars.2021.663800

Nicholson, D., Emerson, S., & Eriksen, C. C., (2008), Net community production in the deep euphotic zone of the subtropical North Pacific gyre from glider surveys, Limnology and Oceanography, 53, https://doi.org/10.4319/lo.2008.53.5_part_2.2226.

Nicholson, D.P. and Feen, M.L., 2017. Air calibration of an oxygen optode on an underwater glider. *Limnology and Oceanography: Methods*, *15*(5), pp.495-502. https://doi.org/10.1002/lom3.10177

Ocean Observatories Initiative Facility Board. 2021. Ocean Observatories Initiative (OOI) Science Plan: Exciting Science Opportunities using OOI Data. https://ooifb.org/reports/ooi-science-plan, DOI: https://doi.org/10.23860/ooi-science-plan-2021-01

Orr, J. C., Epitalon, J. M., Dickson, A. G., & Gattuso, J. P. (2018). Routine uncertainty propagation for the marine carbon dioxide system. *Marine Chemistry*, 207, 84-107. https://doi.org/10.1016/j.marchem.2018.10.006

Orr, J. C., Epitalon, J. M., & Gattuso, J. P. (2015). Comparison of ten packages that compute ocean carbonate chemistry. *Biogeosciences*, 12(5), 1483-1510. https://doi.org/10.5194/bg-12-1483-2015

Owens, W. B., & Millard Jr, R. C. (1985). A new algorithm for CTD oxygen calibration. *Journal of Physical Oceanography*, *15*(5), 621-631. https://doi.org/10.1175/1520-0485(1985)015%3C0621:ANAFCO%3E2.0.CO;2

Palevsky, H.I. and Nicholson, D.P., 2018. The North Atlantic biological pump: Insights from the ocean observatories initiative Irminger Sea Array. *Oceanography*, *31*(1), pp.42-49. https://doi.org/10.5670/oceanog.2018.108

Perez, F. F., & Fraga, F. (1987). Association constant of fluoride and hydrogen ions in seawater. *Marine Chemistry*, 21(2), 161-168. https://doi.org/10.1016/0304-4203(87)90036-3

Roesler, C. S., & Barnard, A. H. (2013). Optical proxy for phytoplankton biomass in the absence of photophysiology: Rethinking the absorption line height. *Methods in Oceanography*, 7, 79-94. https://doi.org/10.1016/j.mio.2013.12.003

Roesler, C., J. Uitz, H. Claustre, E. Boss, X. Xing, E. Organelli, N. Briggs, et al. 2017. “Recommendations for Obtaining Unbiased Chlorophyll Estimates from in Situ Chlorophyll

Fluorometers: A Global Analysis of WET Labs ECO Sensors." *Limnology and Oceanography: Methods* 15 (6): 572–85. https://doi.org/10.1002/lom3.10185.

Sakamoto, Carole M., Kenneth S. Johnson, and Luke J. Coletti. Improved algorithm for the computation of nitrate concentrations in seawater using an in situ ultraviolet spectrophotometer. Limnology and Oceanography: Methods 7.1 (2009): 132-143. https://doi.org/10.3390%2Fs21030965

Sastri, A. R., Christian, J. R., Achterberg, E. P., Atamanchuk, D., Buck, J. J., Bresnahan, P., ... & Tsunoda, T. (2019). Perspectives on in situ sensors for ocean acidification research. Frontiers in Marine Science, 653. https://doi.org/10.3389/fmars.2019.00653

Schallenberg, C.; Jansen, P. and Trull, T.W. (2017) Southern Ocean Time Series (SOTS) Quality Assessment and Control Report Wetlabs FLNTUS instruments Version 2.0. Fluorescence and optical backscatter records, 2009-2016. Hobart, Australia, CSIRO, 37pp. (OA/IR SOTS001). http://dx.doi.org/10.26198/5c4a932ca8ae4

Schar, D.; Atkinson, M.; Johengen, T.; Pinchuk, A.; Purcell, H.; Robertson, C.; Smith, G.J. and Tamburri, M. (2009) Performance Demonstration Statement Sunburst Sensors SAMI-CO2. Solomon, MD, Alliance for Coastal Technologies (ACT), 25pp. (ACTDS10-04). DOI: http://dx.doi.org/10.25607/OBP-344.

Schmechtig, C., Poteau, A., Claustre, H., D'Ortenzio, F., Dall'Olmo, G. and Boss, E., 2018. Processing BGC-Argo particle backscattering at the DAC level. Version 1.4, 07 March 2018. http://dx.doi.org/10.25607/OBP-71

Seidel, M.P., DeGrandpre, M.D., and A.G. Dickson (2008). A sensor for in situ indicator-based measurements of seawater pH. Mar. Chem., 109, 18-28. https://doi.org/10.1016/j.marchem.2007.11.013

Smith, L.M., J.A. Barth, D.S. Kelley, A. Plueddemann, I. Rodero, G.A. Ulses, M.F. Vardaro, and R. Weller. 2018. The Ocean Observatories Initiative. Oceanography 31(1):16–35, https://doi.org/10.5670/oceanog.2018.105.

Snazelle, T.T., 2016, The effect of suspended sediment and color on ultraviolet spectrophotometric nitrate sensors: U.S. Geological Survey Open-File Report, 2016−1014, 10 p., http://dx.doi.org/10.3133/ofr20161014.

SUNA V2 Users Manual, 2018
https://www.seabird.com/asset-get.download.jsa?id=54627862534

Sutton, A.J., Feely, R.A.,et al., (2019). Autonomous Seawater pCO_2 and pH Time Series from 40 Surface Buoys and the Emergence of Anthropogenic Trends. Earth Syst. Sci. Data, 11, 421-439. https://doi.org/10.5194/essd-11-421-2019

Sutton, A. J., Sabine, C. L., Maenner-Jones, S., Lawrence-Slavas, N., Meinig, C., Feely, R. A., Mathis, J. T., Musielewicz, S., Bott, R., McLain, P. D., Fought, H. J., and Kozyr, A. (2014). A high-frequency atmospheric and seawater pCO2 data set from 14 open-ocean sites using a moored autonomous system. Earth Syst. Sci. Data, 6:353-366. doi: 10.5194/essd-6-353-2014.

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Watson, A., Bakker, D. C. E., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Sabine, C., Olafsson, J., Arnarson, T. S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Körtzinger, A., Steinhoff, T., Hoppema, M., de Baar, H. J. W., Wong, C. S., Delille, B. and Bates, N. R. (2009). Climatological mean and decadal changes in surface ocean pCO_2 , and net sea-air CO_2 flux over the global oceans. Deep-Sea Res. II, 554-577. doi: 10.1016/j.dsr2.2008.12.009.

Takeshita, Y., Martz, T.R., Johnson, K.S., Plant, J.N., Gilbert, D., Riser, S.C., Neill, C. and Tilbrook, B., 2013. A climatology‐based quality control procedure for profiling float oxygen data. Journal of Geophysical Research: Oceans, 118(10), pp.5640-5650. https://doi.org/10.1002/jgrc.20399

Thomalla, S. J., W. Moutier, T. J. Ryan-Keogh, L. Gregor, and J. Schütt (2018), An optimized method for correcting fluorescence quenching using optical backscattering on autonomous platforms, Limnology and Oceanography: Methods, 16(2), 132-144, https://doi.org/10.1002/lom3.10234

Toll, R. (ed.) (2012) U.S. IOOS QARTOD Project Plan. Silver Spring, MD, IOOS, 8pp. DOI: http://dx.doi.org/10.25607/OBP-533

Trowbridge, J., Weller, R., Kelley, D., Dever, E., Plueddemann, A., Barth, J.A. and Kawka, O., 2019. The Ocean Observatories Initiative. Frontiers in Marine Science, p.74. https://doi.org/10.3389/fmars.2019.00074

Uchida, H. Johnson, G.,C. and McTaggart, G,. C. (2010) CTD Oxygen Sensor Calibration Procedures. In, The GO-SHIP Repeat Hydrography Manual: A Collection of Expert Reports and Guidelines. Version 1, (eds Hood, E.M., C.L. Sabine, and B.M. Sloyan), 17pp. (IOCCP Report Number 14; ICPO Publication Series Number 134). DOI: https://doi.org/10.25607/OBP-1344

Van Heukelem, L., & Thomas, C. S. (2001). Computer-assisted high-performance liquid chromatography method development with appli- cations to the isolation and analysis of phytoplankton pigments. Journal of Chromatography A, 910(1), 31–49. https://doi.org/10.1016/s0378-4347(00)00603-4

Wolf, M.K., Hamme, R.C., Gilbert, D., Yashayaev, I. and Thierry, V., 2018. Oxygen saturation surrounding deep water formation events in the Labrador Sea from Argo‐O2 data. Global Biogeochemical Cycles, 32(4), pp.635-653. https://doi.org/10.1002/2017GB005829

Xing, X., Morel, A., Claustre, H., Antoine, D., D'Ortenzio, F., Poteau, A., and Mignot, A. (2011), Combined processing and mutual interpretation of radiometry and fluorimetry from autonomous profiling Bio-Argo floats: Chlorophyll a retrieval, J. Geophys. Res., 116, C06020, doi:10.1029/2010JC006899.

Xing, X., Claustre, H., Blain, S., d'Ortenzio, F., Antoine, D., Ras, J. and Guinet, C., 2012. Quenching correction for in vivo chlorophyll fluorescence acquired by autonomous platforms: A case study with instrumented elephant seals in the Kerguelen region (Southern Ocean). *Limnology and Oceanography: Methods*, *10*(7), pp.483-495. https://doi.org/10.4319/lom.2012.10.483

Xing, X., N. Briggs, E. Boss, and H. Claustre (2018), Improved correction for non-photochemical quenching of in situ chlorophyll fluorescence based on a synchronous irradiance profile, Opt Express, 26(19), 24734-24751, https://doi.org/10.1364/OE.26.024734