IMOS Integrated Marine Observing System

Australian Ocean Data Network (AODN)

WATER SAMPLING DATA WORKFLOW

(including data from National Reference Stations, Larval Fish, Microbial Coastal Sampling and Southern Ocean Time Series)

Version 5.0

11th July 2024

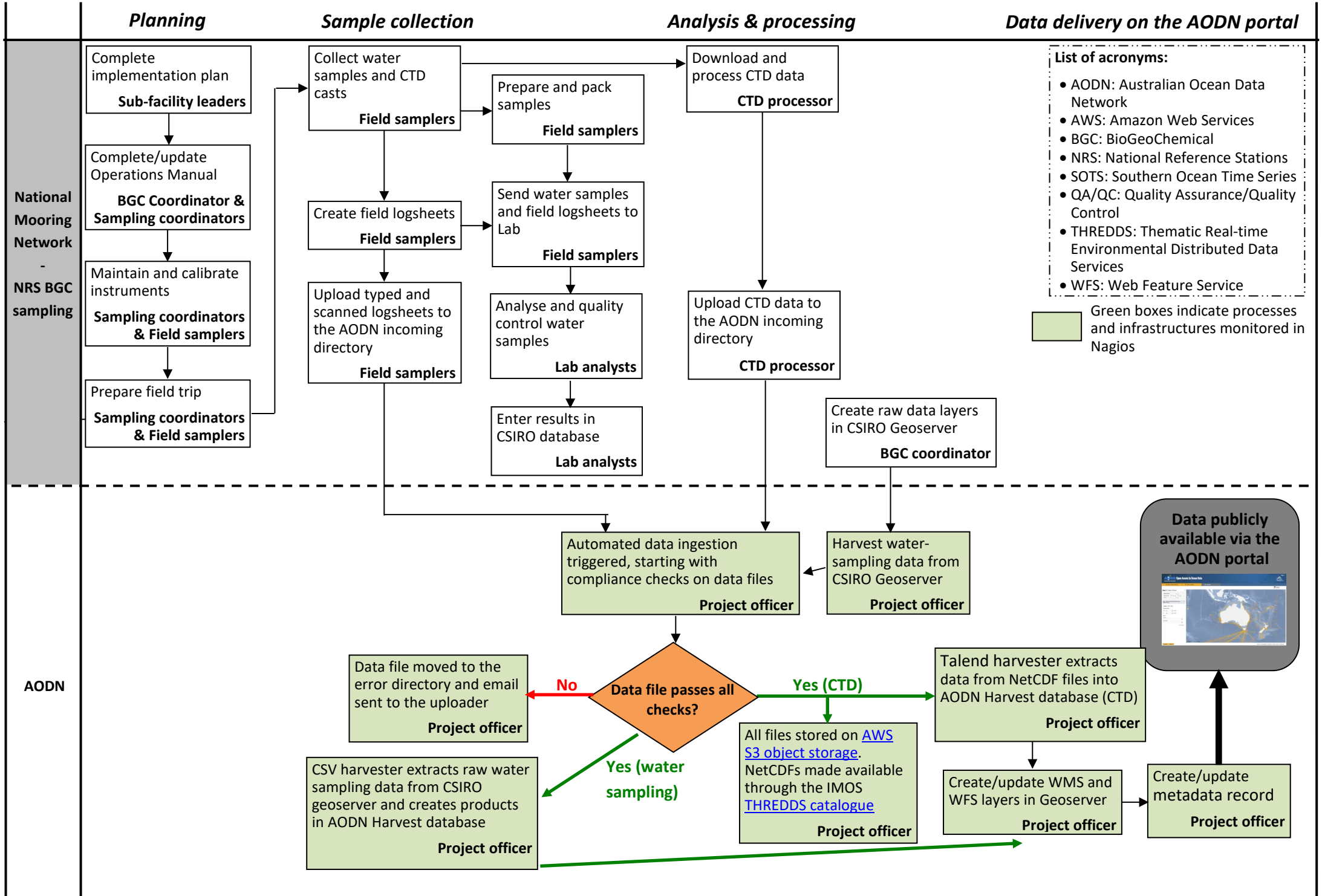
Data Workflows

The AODN, in managing the data for IMOS, has developed workflows for each IMOS sub-facility to describe the flow of IMOS data from planning through data collection to data delivery and public data access.

The primary goals of this workflow are to:

- Improve data flow and data handoff, making tracking of data status easy and preventing data loss
- Identify and delimit precisely the responsibilities of each person involved
- Improve communication at the interface between IMOS facilities and the AODN
- Improve transparency for end users by providing more details to populate metadata records (*i.e.* limitations and processing methods applied to datasets)
- Assist in reporting planned deployments against actual deployments and data delivery

The workflow is available on the next page of this document. Additional information (*i.e.* timeline, input, output, step description) for each operation step is available in the 'Supporting Information' section. The role and contact details of people involved in the workflow are summarised in a table and suggested potential improvements are listed at the end of the document.


Planning
Sample collection
Analysis & processing
Data delivery on the AODN portal
National Mooring Network - NRS BGC sampling
AODN
Complete implementation plan
Sub-facility leaders
Complete/update Operations Manual
BGC Coordinator & Sampling coordinators
Maintain and calibrate instruments
Sampling coordinators & Field samplers
Prepare field trip
Sampling coordinators & Field samplers
Collect water samples and CTD casts
Field samplers
Prepare and pack samples
Field samplers
Create field logsheets
Field samplers
Send water samples and field logsheets to Lab
Field samplers
Upload typed and scanned logsheets to the AODN incoming directory
Field samplers
Analyse and quality control water samples
Lab analysts
Enter results in CSIRO database
Lab analysts
Download and process CTD data
CTD processor
Upload CTD data to the AODN incoming directory
CTD processor
Create raw data layers in CSIRO Geoserver
BGC coordinator
List of acronyms:
AODN: Australian Ocean Data Network
AWS: Amazon Web Services
BGC: BioGeoChemical
NRS: National Reference Stations
SOTS: Southern Ocean Time Series
QA/QC: Quality Assurance/Quality Control
THREDDS: Thematic Real-time Environmental Distributed Data Services
WFS: Web Feature Service
Green boxes indicate processes and infrastructures monitored in Nagios
Automated data ingestion triggered, starting with compliance checks on data files
Project officer
Harvest water-sampling data from CSIRO Geoserver
Project officer
Data publicly available via the AODN portal
Data file passes all checks?
No
Data file moved to the error directory and email sent to the uploader
Project officer
Yes (CTD)
Yes (water sampling)
CSV harvester extracts raw water sampling data from CSIRO geoserver and creates products in AODN Harvest database
Project officer
All files stored on AWS S3 object storage. NetCDFs made available through the IMOS THREDDS catalogue
Project officer
Talend harvester extracts data from NetCDF files into AODN Harvest database (CTD)
Project officer
Create/update WMS and WFS layers in Geoserver
Project officer
Create/update metadata record
Project officer

Supporting information

Phase	Operation step	Timeline	Input	Output	Step description	Step operator
Planning	**Complete implementation plan**		• IMOS funding • Node Science plans	Implementation plan		**Sub-facility leaders**
	Complete/update Operations Manual		Implementation plan	Operations Manual completed/updated	Operations are described in the *NRS Biogeochemical Operations Manual*, which describes all practical aspects of the program, including safety, equipment maintenance, field sampling protocol, post-processing, CTD pre-run checks & data processing, analytical methods and data handling.	**BGC Coordinator & Sampling coordinators**
	Maintain and calibrate instruments		Operations Handbooks	Instruments operational	Send to CSIRO calibration lab when necessary.	**ANMN Sampling coordinators & Field samplers**
	Prepare field trip	**Half a day**	• Operations Handbooks • Operational instruments	Field trip organised	• 0.5 day for lab analysts to get gear ready and organise freighting. • Liaison between field samplers and lab analysts as the latter supply equipment for each field trip.	**ANMN Sampling coordinators & Field samplers**
Sample collection	**Collect water samples and CTD casts**	**1 day**	• Operations Handbook • Field trip organised	• Water samples • CTD casts • Field logsheet	Visit the NRS site, complete all sampling tasks as per Operations Handbooks and fill out field logsheet. Additional data: Larval Fish – Discontinued in June 2022. Selected NRS sampling trips. Went to SIMS for processing. Microbial Coastal Sampling – Nutrients and Flow Cytometry are collected and processed as per the NRS samples (CSIRO and SARDI). SOTS – Phytoplankton samples collected from the SOTS mooring after retrieval.	**ANMN Field samplers**
	Prepare and pack samples		• Operations Handbook • Water samples	Samples ready for shipping	Filtering, preserving and packing samples following procedures described in Operations Handbook	**ANMN Field samplers**

	Create field logsheets	**1 hour**	Field logsheet	Scanned and typed field logsheets	Scan the handwritten field logsheet and transfer all information into a typed version. Save both as PDF.	**ANMN Field samplers**
	Upload field logsheets to the AODN incoming directory		Scanned and typed field logsheets	Scanned and typed field logsheets in AODN incoming directory	Upload logsheets as early as possible so that AODN can keep track of when field trips have taken place at each station. Field sheet information is also uploaded into the CSIRO database to provide the metadata for the samples	**ANMN Field samplers**
Sample collection	**Send water samples and field logsheets to Lab**	**Every three or four months**	• Samples ready for shipping • Field logsheets (electronic form)	Samples and logsheets at CSIRO Labs (Hobart, Floreat or Cleveland)	2 out of 3 zooplankton samples (count and biomass) apart from MAI and NSI and PHB phytoplankton samples are sent to CSIRO Dutton Park. All other samples are sent to CSIRO Hobart. Nutrients samples from WA are sent to CSIRO IOMCRA . Notify BGC coordinator and AODN Project officer when samples are shipped. Field logsheets are sent in electronic format.	**ANMN Field samplers**
Analysis & processing	**Analyse and quality control water samples**	**Three days for analysis and processing**	Samples and logsheets at CSIRO Labs (Hobart, Floreat or Cleveland)	Quality-controlled data		**ANMN Lab analysts**
	Enter analysis results into database at CSIRO		Quality-controlled data	New analysis results in database	Water sampling data are entered into a database at CSIRO. Record a “missing sample” entry if no sample is available for a given field trip.	**ANMN Lab analysts**
	Download and process CTD data		CTD casts	CTD profiles in text (.cnv) and NetCDF formats, preview plots	Process raw data as per *Standardised Profiling CTD Data Post-Processing Procedures* document to create .cnv files. Then use the **latest stable version** of the IMOS Matlab Toolbox to apply further quality control, create NetCDF files (FV00 & FV001) and preview plots (PNG format).	**ANMN CTD processor**
	Upload CTD data to the AODN incoming directory		Processed CTD data (.cnv, NetCDF and plots)	CTD files in AODN incoming directory	Upload all .cnv, NetCDF (FV00 & FV01) and preview plots (PNG)	**ANMN CTD processor**

	Create raw data layers in CSIRO Geoserver		New analysis results in database	Raw data layers in CSIRO Geoserver	Create/update Geoserver layers (and underlying database views) so that AODN can periodically harvest all the raw water-sampling (WS) data without direct access to the CSIRO database.	**BGC coordinator**
	Harvest water-sampling data from CSIRO Geoserver	**Automatic (when files uploaded)**	Raw data layers in CSIRO Geoserver	Raw water sampling (WS) data in AODN incoming directory	A weekly scheduled process, downloads raw data from CSIRO Geoserver, into CSV files, and pushes them into the AODN ingestion pipeline.	**AODN Project officer**
	Automated data ingestion triggered, starting with compliance checks on data files		CTD files, raw WS data and logsheets available in the AODN incoming directory	Valid files ready to be published, non-compliant files moved to error directory	Handling of incoming files is automatically triggered when a file arrives in the AODN incoming directory. Checks are applied to verify that each file • is a correct data product for that upload location; • is compliant with the CF and IMOS conventions (NetCDF files); or • Is valid raw WS data file according to schemas agreed with CSIRO Any file that fails any of these checks is moved to an error directory and the provider is notified by email. The steps below are only performed for files that pass all checks.	**AODN Project officer**
	All files stored on AWS S3 object storage. NetCDFs made available through the IMOS THREDDS catalogue	**Automatic (when files uploaded**	Valid files	Files stored on S3, NetCDF files accessible via IMOS THREDDS catalogue.	All logsheets (scanned & typed) and CTD files (.cnv, FV00 & FV01 NetCDF) are stored and accessible via S3. NetCDF files are also accessible via THREDDS.	**AODN Project officer**

Analysis & processing	**Talend harvester extracts data and creates/populates AODN Harvest database (CTD)**		Valid NetCDF files	Metadata and data in database tables for access via WMS/WFS	Harvester extracts the metadata required to make the files accessible via the AODN Portal and stores them in the database. For some file types the data (measurements) are also extracted.	**AODN Project officer**
	CSV harvester extracts raw water sampling data from CSIRO Geoserver and creates products in AODN Harvest database		Valid raw WS data in CSV files	Raw WS data and data products in AODN database	Harvester loads all raw data into AODN Harvest database and creates database views for the derived products to be served via AODN Geoserver	**AODN Project officer**

Data delivery on the AODN portal	**Create/update WMS and WFS layers in Geoserver**	**1 week (when necessary)**	Populated AODN Harvest database tables	WMS and WFS layers created in Geoserver	• Database tables and/or views are used to create a WMS and WFS layer in Geoserver. • Configure the pop-up window (content.ftl) and filters. • Create style for WMS visualisation.	**AODN Project officer**
	Create/update metadata record		Populated AODN Harvest database tables	Geonetwork record created and configured to support data discovery, visualisation and download via the AODN portal	• Create a metadata record with a new UUID. • Configure the newly created record (*e.g.* abstract, point of contact, parameters, timeframe). • Fill out the distribution section with links to the corresponding Geoserver WMS and WFS layers and other AODN's download services. • Talend harvester automatically updates bounding box.	**AODN Project officer**

Contact details (as at July 2024)

	Role	Stations/Data streams	Name	Email address	Phone number
National Mooring Network - NRS biological sampling	**Sub-facility leaders**	**QLD and NA** (NRSNIN, NRSDAR, NRSYON)	Simon Spagnol	s.spagnol@aims.gov.au	(08) 6369 4004
		NSW (NRSPHB)	Moninya Roughan	mroughan@unsw.edu.au	(02) 2046 6020
		SA (NRSKAI)	Mark Doubell	mark.doubell@sa.gov.au	(08) 8429 0982
		WA (NRSESP, NRSROT)	Ming Feng	ming.feng@csiro.au	(08) 9333 6512
		NRSMAI & NRSNSI	David Hughes	david.hughes@csiro.au	(03) 6232 5439
		SOTS			
	BGC coordinator		Claire Davies	Claire.Davies@csiro.au	(03) 6232 5273
	Sampling coordinators	NRSDAR, NRSYON	Simon Spagnol	s.spagnol@aims.gov.au	(08) 6369 4004
		NRSPHB	Tim Ingleton	Tim.Ingleton@environment.nsw.gov.au	0428212760
		NRSKAI	Mark Doubell	mark.doubell@sa.gov.au	(08) 8429 0982
		NRSROT	Ryan Crossing	Ryan.crossing@csiro.au	(08) 9333 6513
		NRSNSI	Anthony Richardson	a.richardson1@uq.edu.au	(07) 3833 5958
		NRSVB	Daniel Ierodiaconou	daniel.ierodiaconou@deakin.edu.au	
		NRSMAI	David Hughes	david.hughes@csiro.au	(03) 6232 5439
	Field samplers	NRSDAR	Constanza Buccella	c.buccella@aims.gov.au	
		NRSYON	Ulysse Bove	u.bove@aims.gov.au	
		NRSPHB	Michael Sutherland	Michael.Sutherland@environment.nsw.gov.au	
		NRSKAI	Ian Moody	Ian.Moody@sa.gov.au	
		NRSROT	Ryan Crossing	Ryan.crossing@csiro.au	
		NRSVBM	James Devenport	j.devenport@deakin.edu.au	
		NRSMAI	Julie Janssens	julie.janssens@csiro.au	
		NRSNSI	Frank Coman	Frank.Coman@csiro.au	
	CTD processors	NRSDAR, NRSYON,	Constanza Buccella	c.buccella@aims.gov.au	
		NRSNSI	Felicity McEnnulty	felicity.mcennulty@csiro.au	
		NRSMAI	Julie Janssens	julie.janssens@csiro.au	
		NRSPHB	Neil Doszpot	neil.doszpot@environment.nsw.gov.au	
		NRSKAI / NRSVBM	Hugo Bastos de Oliviera	hugo.oliveria@sa.gov.au	

		NRSNIN, NRSESP, NRSROT	Mark Snell	mark.snell@csiro.au	
National Mooring Network - NRS biological sampling	**Lab analysts**	Carbon	Abe Passmore	abe.passmore@csiro.au	
		Nutrients, salinity, dissolved O_2	Julie Janssens	julie.janssens@csiro.au	
		TSS, pigments	Rasanthi Gunasekera	Rasanthi.gunasekera@csiro.au	
		Flow cytology	Dion Frampton	Dion.frampton@csiro.au	
		Phytoplankton	Claire Davies	Claire.Davies@csiro.au	
		Genetics	Jodie Van de Kamp	Jodie.vandekamp@csiro.au	
		Zooplankton	Frank Coman	frank.coman@csiro.au	
AODN	**Project officer**		Leonardo Laiolo	Leonardo.Laiolo@utas.edu.au	(03) 6226 2242
	Data services team leader		Benedicte Pasquer	Benedicte.Pasquer@utas.edu.au	(03) 6226 1927

Supporting links

AODN portal: http://portal.aodn.org.au

AODN portal help page: https://help.aodn.org.au/

IMOS metadata catalogue: http://catalogue-imos.aodn.org.au/geonetwork/

Data upload instructions: https://help.aodn.org.au/aodn-ftp-upload/ (contact AODN for access)
IMOS Moorings internal discussions: https://github.com/aodn/anmn-internal-discussions (contact AODN for access)
National Reference Stations data on S3 server: http://data.aodn.org.au/?prefix=IMOS/ANMN/NRS/
National Reference Stations THREDDS server: http://thredds.aodn.org.au/thredds/catalog/IMOS/ANMN/NRS/catalog.html

IMOS user code library: https://github.com/aodn/imos-user-code-library
National Mooring Network pipeline scripts on GitHub: https://github.com/aodn/python-aodndata/tree/master/aodndata/moorings and https://github.com/aodn/public-schema
IMOS website: http://www.imos.org.au/

National Mooring Network facility: https://imos.org.au/facilities/nationalmooringnetwork
National Mooring Network documentation (including link to the latest version of the BGC operations manual):
https://imos.org.au/facilities/nationalmooringnetwork/moorings-documentation