

# **National Mooring Network**

## **Plankton products**

### **Product Specification**

---

#### **Document Change Control**

<b>Version Number</b>	<b>Date of Issue</b>	<b>Brief Description of Change</b>
Version 0.2	24 Feb 2021	Draft version for review and comment
Version 0.3	19 Jul 2021	Draft version with larval fish for review and comment
Version 0.4	18/08/2021	Draft updated to mention other stations (not just NRS)
Version 0.5	27-07-2022	Final

## Introduction

### Summary

This dataset comprises the plankton abundance, biomass and biovolume measurements from samples collected at the Integrated Marine Observing Systems (IMOS) National Reference Stations (NRS) and other repeatedly sampled stations. For information on using the data from the NRS please refer to Eriksen et al., 2019. It is advised that anyone using this data should read this methodology or contact the project contact person to understand the nuances of each sampling and analysis chain. Furthermore, this dataset includes larval fish data collected from various locations around Australia, including five NRS (North Stradbroke, Port Hacking, Maria Island, Kangaroo Island and Rottnest Island), between 2014 and 2021, adding information to the biological, ecological and oceanic measurements already sampled at these long-term monitoring sites. The SOTS phytoplankton data is also in this dataset, these data are collected and analysed as per Eriksen et al., 2018.

The data are made available in binned products for ease of use. As taxonomy changes over time and the ability of analysts improves with training more and more species can be identified over time. The products account for these changes and for real absences from the samples. A 0 abundance means that the species was looked for and not seen, -999 means that the species was not looked for in that sample. The products are therefore applicable for use in time series analysis.

There are a total of 14 products to choose from. Phytoplankton data are available as either abundance or biovolume, and each of these at 4 levels of binning:

- Raw product (includes all data)
- Species product (includes all data where species have been identified)
- Genus product (includes all data binned to genus level)
- Higher taxonomic group (includes all data binned to functional groups).

Zooplankton abundance products are available at 5 levels:

- Raw product (includes all data)
- Copepods product (all data where species have been identified, Copepods only)
- Non-Copepods product (all data where species have been identified, non-Copepods only)
- Genus product (includes all data binned to genus level)
- Higher taxonomic group (includes all data binned to functional groups).

The fish larvae product is available as raw count product, where fish larvae are classified by taxon.

### History of changes

Plankton abundances were originally provided on the AODN Portal only as raw data, in flat table form (with each species recorded in a separate row), without any adjustments for taxonomy changes or real absences.

The first version of the "analysis-ready" products were first published on 04/04/2019. However, they were removed again in November 2020 because technical issues prevented them from being regularly updated.

## Product Specifications

### General information

<b>Product name</b>	Phytoplankton Abundance - Higher Taxonomic Groups Phytoplankton Abundance - Genus Data Phytoplankton Abundance - Species Data Phytoplankton Abundance - Raw Data Phytoplankton Biovolume - Higher Taxonomic Groups Phytoplankton Biovolume - Genus Data Phytoplankton Biovolume - Species Data Phytoplankton Biovolume - Raw Data Zooplankton Abundance - Higher Taxonomic Groups Zooplankton Abundance - Genus Data Zooplankton Abundance - Copepods Data Zooplankton Abundance - Non-Copepods Data Zooplankton Abundance - Raw Data Larval Fish Count - Raw Data
<b>Geographic coverage</b>	Australian Coastal region, 9 National Reference Stations for phytoplankton and zooplankton and 5 National Reference Stations for fish larvae
<b>Temporal coverage</b>	phytoplankton and zooplankton: 2008 -- present fish larvae: 2014 - 2021
<b>Temporal resolution</b>	Quarterly to monthly <sup>1</sup>
<b>Update frequency</b>	Weekly
<b>Delivery mechanism</b>	AODN Portal OGC WFS
<b>Format</b>	CSV

### Details of product contents

Name	Description	Units/format	Data type
<b>PHYTOPLANKTON, ZOOPLANKTON AND LARVAL FISH PRODUCTS</b>			
Project	Name of the project the sample was collected for		string
StationName	Name of the station where sample was collected		string
Latitude	Nominal latitude (North) of the station where sample was collected	Degrees	double precision
Longitude	Nominal longitude (East) of the station where sample was collected	Degrees	double precision
TripCode	Unique code for the sampling trip format of station code and date for NRS. Station code and build id and	<station_code>YYYYMMD D	string

<sup>1</sup> Most NRS sites are sampled about once a month. The Kangaroo Island and Darwin NRS sites (and previously Esperance and Ningaloo) are sampled about four times a year.

	sample no for SOTS		
SampleTime_UTC	UTC date and time of sample collection	YYYY-MM-DD HH:MM:SS	timestamp
SampleTime_Local	Local date and time of sample collection	YYYY-MM-DD HH:MM:SS	string
Year_Local	Year of sample collection	YYYY	integer
Month_Local	Month of sample collection	MM	integer
Day_Local	Day of month of sample collection (local time)	DD	integer
Time_Local24hr	Local 24hr time of sample collection	HH:MM	string
SampleDepth_m	Depth of sample collection, or "integrated", or "water column", or depth range (Larval Fish)	metres (m)	string
[taxonomic level name] (e.g., "Noctiluca scintillans")	Taxonomic level abundance (plankton) or biovolume (phytoplankton) or count (fish)	Cells L <sup>-1</sup> (phytoplankton abundance) or µm <sup>3</sup> L <sup>-1</sup> (phytoplankton biovolume) taxa m <sup>-3</sup> (zooplankton) or count (fish)	floating point
<b>PHYTOPLANKTON AND ZOOPLANKTON PRODUCT ONLY</b>			
StationCode	IMOS site code for station		string
CTDSST_degC	Temperature of the water column at 10m depth, from matching CTD cast	Degrees Celsius (°C)	floating point
CTDChlaSurf_mgm3	Artificial chlorophyll concentration at 10m depth, computed from bio-optical sensor on matching CTD cast	milligrams per metre cube (mg m <sup>-3</sup> )	floating point
CTDSalinity_psu	Derived salinity of the water column at 10m depth, from matching CTD cast	Practical Salinity Unit (PSU)	floating point
<b>PHYTOPLANKTON PRODUCT ONLY</b>			
Method	Light microscopy or SEM, plus preservation methods		string
<b>ZOOPLANKTON PRODUCT ONLY</b>			
Biomass_mgm3	Zooplankton dry weight	mg m <sup>-3</sup>	floating point
AshFreeBiomass_mgm3	Organic matter after furnacing	mg m <sup>-3</sup>	Floating point
<b>LARVAL FISH PRODUCT ONLY</b>			

Temperature_degC	Underway measured temperature or CTD measurement for NRS stations	Degrees Celsius (°C)	floating point
Salinity_psu	Underway measured salinity or CTD measurements for NRS stations	Practical Salinity Unit (PSU)	floating point
Volume_m3	Volume of water collected to sample larval fishes	metres cubed (m <sup>3</sup> )	real
Vessel	Name of the vessel that performed the larval fish sampling		string
TowType	Type of fishing net technique used to collect larval fish samples		string
GearMesh_um	Mesh size of the fishing net used to collect larval fish samples	micrometres (µm)	string
Bathymetry_m	Sea floor depth in the larval fish sampling location	metres (m)	real
QC_flag	Quality control for larval fish sampling	values from 0 to 9	real

## Product delivery

### Format

The products are available in comma-separated values (CSV) format. All four variants have the same format, differing only in the columns corresponding to taxonomic groups.

### AODN Portal

Dataset collections:

- **IMOS Phytoplankton Abundance and Biovolume (reference stations)**
- **IMOS Zooplankton Abundance and Biomass (reference stations)**
- **IMOS Larval Fish Sub-Facility - Database of Marine Larval Fish Assemblages in Australian temperate and subtropical waters (1983 - ongoing)**

The specific product is selected under the *Download as...* drop-down menu on Step 3.

The collections can be subset by

- Geographic bounding box
- Temporal range
- Station name

### AODN Geoserver

The products can be accessed via a Web Feature Service (WFS) query to the [AODN Geoserver](#). The relevant layers are:

- imos:bgc\_phytoplankton\_abundance\_raw\_data
- imos:bgc\_phytoplankton\_abundance\_species\_data
- imos:bgc\_phytoplankton\_abundance\_genus\_data
- imos:bgc\_phytoplankton\_abundance\_htg\_data
- imos:bgc\_phytoplankton\_biovolume\_raw\_data
- imos:bgc\_phytoplankton\_biovolume\_species\_data
- imos:bgc\_phytoplankton\_biovolume\_genus\_data
- imos:bgc\_phytoplankton\_biovolume\_htg\_data
- imos:bgc\_zooplankton\_abundance\_raw\_data
- imos:bgc\_zooplankton\_abundance\_copepods\_data
- imos:bgc\_zooplankton\_abundance\_non\_copepods\_data
- imos:bgc\_zooplankton\_abundance\_genus\_data
- imos:bgc\_zooplankton\_abundance\_htg\_data
- imos:bgc\_larval\_fish\_count\_raw\_data

The Web Map Service (WMS) layers used for preview in Step 2 on the AODN Portal are

- imos:bgc\_phytoplankton\_map
- imos:bgc\_zooplankton\_map
- imos:bgc\_larval\_fish\_map

## Data Lineage

### Provenance

Water samples are collected off small vessels at the IMOS National Reference Stations. The depth of the sample varies at each station. The sampling methods are fully described in the [IMOS NRS Biogeochemical Operations Manual](#) (Davies & Sommerville 2020). The analysis and quality control (QC) procedures (performed at CSIRO) are described in Eriksen et al., 2019. The SOTS data comes from the Southern Ocean Time Series project. The location is a fixed mooring with samples taken at various depths at regular intervals, methodologies and quality control procedures are described in Eriksen et al., 2020.

### Harvest to AODN

The phytoplankton products are based on raw data harvested from the following WFS layers in the CSIRO Geoserver:

- **bgc\_trip** - metadata for sampling trips and samples taken (+ zooplankton biomass, CTD surface temperature & chlorophyll)
- **bgc\_phyto\_raw** - observed abundances, biovolumes, and taxonomic details of individual taxa identified in each sample. This will include the SOTS data.
- **bgc\_phyto\_changelog** - details when specific taxa were first looked for and first observed
- **bgc\_zoop\_raw** - observed abundances and taxonomic details of individual taxa identified in each sample
- **bgc\_zoop\_changelog** - details when specific taxa were first looked for and first observed
- **bgc\_lfish\_countraw** - observed counts and taxonomic details of individual taxa identified in each sample
- **bgc\_lfish\_samples** - metadata related to larval fish, including volume of water sampled, vessel name, tow type, gear depth, gear mesh, bathymetry Each of these layers is downloaded into a CSV file, then processed by the AODN data ingestion pipeline.

- **phytoinfo** - metadata related to phytoplankton, it will be stored in S3 at IMOS/BGC\_DB/harvested\_from\_CSIRO to be publicly available.
- **zoopinfo** - metadata related to phytoplankton, it will be stored in S3 at IMOS/BGC\_DB/harvested\_from\_CSIRO to be publicly available.

## Processing at AODN

Harvest all input data into the Harvest database

Create **Raw Data** product:

1. Join sample metadata onto raw data on Sample code
2. Pivot unique values of TaxonName to separate columns, with values set from Cells\_L (nulls go to 0).

Create **Higher Taxonomic Groups** product:

1. Group the raw data by Sample and TaxonGroup, summing up the Cells\_L values
2. Remove rows with TaxonGroup in ('Other', 'Coccolithophore', 'Diatom', 'Protozoa')
3. Join sample metadata
4. Set StartDate to 2012-08-07
5. Make a dummy line of data for samples that have nothing recorded against them (unlikely at this level, but could happen) and replace nulls in Cells\_L with zero
6. Pivot unique values of TaxonGroup to columns, with values set from Cells\_L

Create **Genus Data** product

1. From the taxon changelog find the TaxonNames where the genus was affected (i.e., genus in TaxonName  $\neq$  genus in ParentName)
2. From the raw data with genus not affected by changelog (TaxonName not in the list from step 1.):
  - a. Group by Sample and Genus, summing up the Cells\_L values, and dropping rows where Genus is empty
  - b. Join sample metadata
  - c. Replace missing values in Cells\_L with zero
  - d. Group by all sample metadata columns (Station, Lat/Lon, date/time) and Genus, summing up the Cells\_L values
3. For raw data affected by the changelog
  - a. For each Genus observed,
    - i. replace missing StartDate with value from the changelog from parent taxon
    - ii. set default abundance values (NULL for 'not looked for', 0 for 'not found')
  - b. Group by all sample metadata columns (Station, Lat/Lon, date/time) and Genus, summing up the Cells\_L values
  - c. Append to the other data from step 2.

Create **Species Data** product



1. As per genus data product but based on taxon identified to species, i.e., group by genus and species and drop all taxa where species is not a defined species.

## AODN Harvest Database schema

**Schema name:** imos\_bgc\_db

Tables to hold data & metadata harvested from CSIRO:

- bgc\_trip
- bgc\_phyto\_raw
- bgc\_phyto\_changelog
- bgc\_zoop\_raw
- bgc\_zoop\_changelog
- larval\_fish\_raw\_harvest
- larval\_fish\_samples\_harvest

Materialised views for Geoserver layers:

- bgc\_phytoplankton\_abundance\_raw\_data
- bgc\_phytoplankton\_abundance\_species\_data
- bgc\_phytoplankton\_abundance\_genus\_data
- bgc\_phytoplankton\_abundance\_htg\_data
- bgc\_phytoplankton\_biovolume\_raw\_data
- bgc\_phytoplankton\_biovolume\_species\_data
- bgc\_phytoplankton\_biovolume\_genus\_data
- bgc\_phytoplankton\_biovolume\_htg\_data
- bgc\_zooplankton\_abundance\_raw\_data
- bgc\_zooplankton\_abundance\_copepods\_data
- bgc\_zooplankton\_abundance\_noncopepods\_data
- bgc\_zooplankton\_abundance\_genus\_data
- bgc\_zooplankton\_abundance\_htg\_data
- bgc\_larval\_fish\_count\_raw\_data

Views for Geoserver layers:

- bgc\_phytoplankton\_map
- bgc\_zooplankton\_map
- bgc\_larval\_fish\_map

## Glossary and abbreviations

AODN	Australian Ocean Data Network
AWS	Amazon Web Services
BGC	Biogeochemical
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CSV	Comma-separated values
CTD	Conductivity Temperature Depth
IMOS	Integrated Marine Observing System
NRS	National Reference Station
QC	Quality Control
S3	AWS S3 cloud object storage
SOTS	Southern Ocean Time Series
WFS	Web Feature Service
WMS	Web Map Service

## References

Eriksen R, Trull TW, Davies D, Jansen P, Davidson AT, Westwood K and Enden R (2018). Seasonal succession of phytoplankton community structure from autonomous sampling at the Australian Southern ocean time series (SOTS) observatory. *Marine Ecology Progress Series* 589, 13–31. doi: 10.3354/meps12420

Eriksen RS, Davies CH, Bonham P, Coman FE, Edgar S, McEnnulty FR, McLeod D, Miller MJ, Rochester W, Slotwinski A, Tonks ML, Uribe-Palomino J and Richardson AJ (2019) Australia's Long-Term Plankton Observations: The Integrated Marine Observing System National Reference Station Network. *Front. Mar. Sci.* 6:161. DOI: 10.3389/fmars.2019.00161

Eriksen RS, Davies DM, Wynn-Edwards CA, Trull TW (2020) Southern Ocean Time Series (SOTS) Quality Assessment and Control Report Remote Access Sampler Sample Analysis. Phytoplankton analysis. Version 1.0 CSIRO, Australia. DOI: 10.26198/5f115475407b3 (<https://doi.org/10.26198/5f115475407b3>).

Davies, C. and Sommerville, E. (Eds.) (2020), National Reference Stations Biogeochemical Operations Manual. Integrated Marine Observing System. DOI: 10.26198/5c4a56f2a8ae3 (<http://dx.doi.org/10.26198/5c4a56f2a8ae3>)