IMOS NETCDF CONVENTIONS

Conventions and Reference Tables

Version 1.4.2
June 30th, 2021

info@aodn.org.au

Australia's Integrated Marine Observing System (IMOS) is enabled by the National Collaborative Research Infrastructure Strategy (NCRIS). It is operated by a consortium of institutions as an unincorporated joint venture, with the University of Tasmania as Lead Agent.

Table of contents

Index of Tables

Index of Reference Tables

PREFACE to version 1.4.2

Due to the addition of a DOI to this document, we have created a new version of this document.

A suggested citation has been added to the document which includes the DOI information.

Citation

Integrated Marine Observing System (2021). IMOS NetCDF Conventions Version 1.4.2. Integrated Marine Observing System. DOI: 10.26198/ytxm-vm43 (https://doi.org/10.26198/ytxm-vm43).

1 - OVERVIEW

1.1 - About IMOS

IMOS is a distributed set of equipment and data-information services which collectively contribute to meeting the needs of marine climate research in Australia. The observing system provides data in the open oceans around Australia out to a few thousand kilometres as well as the coastal oceans. The IMOS Office coordinates the deployment of a wide range of equipment and assembles the data through 11 facilities distributed around the country. The data are made available to researchers through the Australian Ocean Data Network (AODN) located at the University of Tasmania. The IMOS infrastructure also contributes to Australia's role in international programs of ocean observing.

Australia's Integrated Marine Observing System (IMOS) is enabled by the National Collaborative Research Infrastructure Strategy (NCRIS). It is operated by a consortium of institutions as an unincorporated joint venture, with the University of Tasmania as Lead Agent.

1.2 - About AODN

The AODN will provide a single integrative framework for data and information management that will allow discovery and access of the data by scientists, managers and the public.

Its activity can be summarised as follows:

- host, manage and archive data produced by the other IMOS facilities.
- provide the standards, protocols and systems to integrate the data and related information into a number of conformal frameworks, and will provide the tools to access and utilise the data.
- For some kinds of data, provide data products as web services and web features for processing, integration and visualisation of data.
- Where possible, will integrate data from sources outside IMOS into IMOS data products and export IMOS data to international programs.

1.3 - About this document

The main purpose of this document is to specify the format of the files that are used to distribute IMOS data, and to document the standards used therein. This includes naming conventions, or taxonomy, as well as metadata content.

The IMOS NetCDF Conventions document was originally based on the one prescribed by the OceanSITES User's Manual, version 1.1. As both documents have evolved since, there are now significant differences between them but we will try to reduce this gap in the future. The OceanSITES program is the global network of open-ocean sustained time series reference stations that have been implemented by an international partnership of researchers.

The IMOS NetCDF Conventions document also draws on documents that have been produced for the IMOS project (see References).

2 - IMOS DATA FORMAT

2.1 - Network Common Data Form (NetCDF)

NetCDF is one of many file formats available for storing marine data. It is a binary file format that is self-describing and portable among other features. The netCDF software libraries and documentation are available online from Unidata. Many netCDF manipulation and display software utilities are also available online.

The IMOS NetCDF Conventions have been written to be used when writing data files in netCDF classic format (version 3.6). Unidata released a netCDF-4 format in 2008. This format is more flexible than the classic format and offers additional features such as compression, groups, compound types and variable length arrays. NetCDF library versions 4.0 and above are now widely adopted and backward compatible with the netCDF classic format so whenever relevant and possible, AODN encourages data providers to produce netCDF-4 files. Common Data Language (CDL) is a human readable text notation that is used to describe the netCDF objects. The netCDF utility *ncdump* can be used to convert netCDF binary file to CDL text. The netCDF utility *ncgen* creates a netCDF binary file from a well-formed CDL text file.

A CDL example which describes IMOS temperature data collected from a mooring is shown in APPENDIX 1: example netcdf file.

2.2 - CF conventions

IMOS follows the netCDF Climate and Forecast (CF) Metadata Conventions v1.6 (Eaton et al 2011). CF conventions require conforming datasets to contain sufficient metadata that they are self-describing, in the sense that each variable in the file has an associated description of what it represents, including physical units if appropriate, and that each value can be located in space and time. The CF conventions supply a standard vocabulary and some metadata conventions. Sometimes it was desirable to incorporate attributes or concepts from other conventions to the IMOS NetCDF Conventions. We have made clear notations in this document where the IMOS NetCDF Conventions extends the CF ones.

2.3 - IMOS conventions

IMOS NetCDF Conventions require that:

- Units are compliant with CF/COARDS/Udunits.
- The time parameter is encoded as recommended by COARDS and CF.
- Parameters are given standard names from the CF table when possible.
- Where time is specified as an attribute, the ISO8601 standard is used.
- File names are created following the IMOS File Naming Convention document.

See also Unidata netCDF best practices.

3 - NETCDF FILE STRUCTURE

3.1 - Feature type templates

3.1.1 -Definition

Most of the observation data collected by IMOS are discrete sampling geometry datasets, that is to say data sets “characterized by a dimensionality that is lower than that of the space-time region that is sampled; discrete sampling geometries are typically “paths” through space-time” (see CF1.6 new chapter 9). We can identify distinct types of discrete sampling geometry like point, time series, profile or trajectory, so that for each of these feature type we can adopt a consistent representation or template in a netCDF file structure.

The US National Oceanographic Data Center (NODC) has designed netCDF templates for a wide variety of feature types. For each of them there is usually an orthogonal and an incomplete template available. IMOS recommends using the orthogonal template in which variables of a dataset must contain identical coordinate values along an axis. When the variables of a dataset contain different coordinate values along an axis then they should be separated into distinct files.

3.1.2 -Examples of CDL templates per feature type

The examples below illustrate the structure of a file for each feature type, using a single data variable as an example. A file can have multiple data variables as long as they share a single set of coordinates. For clarity, most attributes and the data values have been omitted.

3.1.2.1 - Time series

```
netcdf IMOS_timeseries_ single_dataset {
dimensions:
        TIME = 22345 ;                  // Number of time steps in the time series
variables:
        double TIME(TIME) ;             // Coordinate variable T
        double LATITUDE ;               // Scalar coordinate variable Y
        double LONGITUDE ;              // Scalar coordinate variable X
        float NOMINAL_DEPTH ;           // Scalar coordinate variable Z
```

```
        float TEMP(TIME) ;                  // geophysical measurement variable
                TEMP:coordinates = "TIME LATITUDE LONGITUDE NOMINAL_DEPTH" ;
                TEMP:ancillary_variables = "TEMP_quality_control" ;
        byte TEMP_quality_control(TIME) ;       // geophysical measurement ancillary variable
                TEMP_quality_control: variable_attribute = "variable attribute value" ;
// global attributes:
                :featureType = “timeSeries” ;
}
```

3.1.2.2 - Profile

```
netcdf IMOS_profile_single_dataset {
dimensions:
        DEPTH = 102 ;                   // Number of vertical cells in the profile
variables:
        float DEPTH(DEPTH) ;            // Coordinate variable Z
        double TIME ;                   // Scalar coordinate variable T
        double LATITUDE ;               // Scalar coordinate variable Y
        double LONGITUDE ;              // Scalar coordinate variable X
        float TEMP(DEPTH) ;             // geophysical measurement variable
                TEMP:coordinates = "TIME LATITUDE LONGITUDE DEPTH" ;
                TEMP:ancillary_variables = "TEMP_quality_control" ;
        byte TEMP_quality_control(DEPTH) ;    // geophysical measurement ancillary variable
// global attributes:
                :featureType = “profile” ;
}
```

3.1.2.1 - Time series profile

```
netcdf IMOS_time_series_profile_single_dataset {
dimensions:
        TIME = 250 ;                    // Number of time steps in the time series
        DEPTH = 40 ;                    // Number of vertical cells in the profile
variables:
        double TIME(TIME) ;             // Coordinate variable T
        float DEPTH(DEPTH) ;            // Coordinate variable Z
        double LATITUDE ;               // Scalar coordinate variable Y
        double LONGITUDE ;              // Scalar coordinate variable X
        float CSPD(TIME, DEPTH) ;       // geophysical measurement variable
```

```
                CSPD:coordinates = "TIME LATITUDE LONGITUDE DEPTH" ;
                CSPD:ancillary_variables = " CSPD_quality_control" ;
        byte CSPD_quality_control(TIME, DEPTH) ;      // geophysical measurement ancillary variable
// global attributes:
                :featureType = “timeSeriesProfile” ;
}
```

3.1.2.2 - Trajectory

```
netcdf IMOS_trajectory_ single_dataset {
dimensions:
        TIME = 22345 ;                          // Number of time steps in the trajectory
variables:
        double TIME(TIME) ;                     // Coordinate variable T
        double LATITUDE(TIME) ;                 // Auxiliary coordinate variable Y
        double LONGITUDE(TIME) ;                // Auxiliary coordinate variable X
        float DEPTH(TIME) ;                     // Auxiliary coordinate variable Z
        float TEMP(TIME) ;                      // geophysical measurement variable
                TEMP:coordinates = "TIME LATITUDE LONGITUDE NOMINAL_DEPTH" ;
                TEMP:ancillary_variables = "TEMP_quality_control" ;
        byte TEMP_quality_control(TIME) ;       // geophysical measurement ancillary variable
// global attributes:
                :featureType = “trajectory” ;
}
```

3.2 - Global Attributes

3.2.1 -Definition

The global attribute section of a netCDF file contains metadata that describes the overall contents of the file and allows for data discovery. All fields should be human-readable and can be of either ‘character’ or ‘numeric’ type. IMOS recommends that all listed attributes be used and contain meaningful information unless there are technical reasons rendering this impossible (for example, information not available for historical data). Files must at least contain the attributes listed as “mandatory”. Please contact AODN if this is proving difficult.

Global attributes can be thought of as conveying five kinds of information:

- What: What are the data in the dataset
- Where: The spatial coverage of the data
- When: The temporal coverage of the data
- Who: Who produced the data
- How: How were the data produced and made available

3.2.2 -Core global attributes

The following table lists all the core global attributes are **mandatory** from the point of view of either the CF or IMOS conventions.

The "Type" values are **S** for string, **N** for numeric (byte, short, long, integer, float or double), **D** for the type of the data variable and **Q** for the type of the corresponding quality control data variable.

Table 1 List of core mandatory global attributes for IMOS netCDF files

Name	Type	Example	Definition
What			
project (Non-CF attribute)	S	project = "Integrated Marine Observing System (IMOS)"	The scientific project that produced the data. For data produced under the IMOS project, the field must be filled as shown in the example.
Conventions	S	Conventions = "CF-1.6,IMOS-1.4"	Name of the format convention used by the dataset. Possibility to include two different conventions if necessary. For data produced under the IMOS project, the field must be filled as shown in the example.

	standard_name_vocabulary	S	standard_name_vocabulary = "NetCDF Climate and Forecast (CF) Metadata Convention Standard Name Table Version 29"	Table number used for CF standard names.
	title	S	title = “Radar data from Tannum Sands station, Queensland”	Short description of the dataset
	institution	S	institution = “ACORN”	Name of the institute or facility where the original data was produced
Non-CF attribute	date_created	S	date_created = “2008-11-23T08:35:00Z”	The date on which the file was created. See chapter **Error! Reference source not found.** on time format below.
Non-CF attribute	abstract	S	abstract = “NSW-IMOS Port Hacking 100m Mooring Water quality meters (WQM's) at one or more depths collect bursting data. Data from the bursts have been cleaned and averaged to create data products. This file is one such product.”	A paragraph describing the dataset: type of data contained in the dataset, how the data was created, the creator of the dataset, the project for which the data was created, the geospatial coverage of the data, the temporal coverage of the data. In some instances the abstract may be autogenerated from other netCDF fields. Please discuss this with AODN staff if you think autogeneration will be appropriate for your data.
Non-CF attribute	naming_authority	S	naming_authority = “IMOS”	This will always be “IMOS”
Non-CF attribute	**Where**			

	geospatial_lat_min / geospatial_lat_max	N	geospatial_lat_min = 59.8 geospatial_lat_max = 59.8	The southernmost / northernmost latitude covered by the data set, a value between -90 and 90 decimal degrees North unless otherwise specified in geospatial_lat_units.
	geospatial_lon_min / geospatial_lon_max	N	geospatial_lon_min =-41.2 geospatial_lon_max=-41.2	The westernmost / easternmost longitude, a value between -180 and 180 decimal degrees East if not specified.
	geospatial_vertical_min / geospatial_vertical_max	N	geospatial_vertical_min = 10.0 geospatial_vertical_max = 2000	Minimum / maximum depth for data included in dataset, in metres and referenced from the sea surface if not specified.
	geospatial_vertical_positive	S	geospatial_vertical_positive = “down”	Direction towards which depth is positive. Possible values are either “up” or “down”.
Non-CF attribute	**When**			
	time_coverage_start / time_coverage_end	S	time_coverage_start = “2008-11-23T08:35:00Z” time_coverage_end = “2009-01-06T15:47:00Z”	Start / final date of the data in UTC. See chapter **Error! Reference source not found.** on time format.
Non-CF attribute	**Who**			
	data_centre	S	data_centre = “Australian Ocean Data Network (AODN)”	Data centre in charge of the data management or party who distributed the resource
	data_centre_email	S	data_centre_email = “info@aodn.org.au”	Data Centre contact e-mail address

	author	S	author = “Doe, John”	Name of the person responsible for the creation of the dataset. Convention is last name and then first name separated by a comma.
	principal_investigator	S	principal_investigator = “Doe, John”	Name of the principal investigator in charge of the platform. Convention is last name and then first name separated by a comma.
Non-CF attribute	**How**			
	citation	S	citation = Integrated Marine Observing System. 2008, "Australian Acoustic Tagging and Monitoring System (AATAMS) data", http://imos.org.au/emii_aatams.html, accessed 20 Dec 2008.	The citation to be used in publications using the dataset should follow the format: “IMOS. [year-of-data-download], [Title], [Data access URL], accessed [date-of-access]”.
	acknowledgement	S	acknowledgement = “Any users of IMOS data are required to clearly acknowledge the source of the material in the format: "Data was sourced from Australia’s Integrated Marine Observing System (IMOS) - IMOS is enabled by the National Collaborative Research Infrastructure Strategy (NCRIS).””	Information about how to acknowledge the source of the material. For data produced under the IMOS project, the field must be filled as shown in the example. If relevant, also credit other organisations involved in collection of this particular data stream.

disclaimer	S	disclaimer = “Data, products and services from IMOS are provided "as is" without any warranty as to fitness for a particular purpose.”	Statement limiting the liability of the data provider. For data produced under the IMOS project, the field must be filled as shown in the example.
license	S	license = http://creativecommons.org/licenses/by/4.0/	Describe the restrictions to data access and distribution. For data produced under the IMOS project, the field must be filled as shown in the example.

3.2.3 -Optional global attributes

Table 2 below lists some of the **optional** global attributes than can be used by any facility to describe an IMOS netCDF dataset.

For an example of facility-oriented optional global attributes, see APPENDIX 1: example netcdf file.

The “Type” values are **S** for string, **N** for numeric (byte, short, long, integer, float or double), **D** for the type of the data variable and **Q** for the type of the corresponding quality control data variable.

Table 2 List of optional global attributes for IMOS netCDF files

Name	Type	Example	Definition
What			
featureType	S	featureType = “timeSeries”	Specifies the type of discrete sampling geometry to which the data in the file belongs, and implies that all data variables in the file contain collections of features of that type. Possible CF featureType values are “point”, “timeSeries”, “profile”, “trajectory”, “timeSeriesProfile”

				or “trajectoryProfile”.
Non-CF attribute	date_modified	S	date_modified = “2008-12-23T20:35:00Z”	The date on which the file was modified. See chapter 3.2.4 - on time format. If this attribute is used for the first time or modified, a new entry needs to be added to the “history” attribute.
	history	S	history = “2014-03-27T23:46:36Z - timeOffsetPP: TIME dimension and time_coverage_start/end global attributes have been applied the following offset : -10 hours.”	Provides an audit trail for modifications to the original data. It should contain a separate line for each modification, with each line beginning with a timestamp and including user name, modification name and modification arguments.
	comment	S	comment =“Geospatial vertical min information has been computed using the Gibbs-SeaWater toolbox (TEOS-10) v3.02 from latitude and relative pressure measurements”	Miscellaneous information about the data or methods used to produce it. Any free-format text is appropriate
	source	S	source = “Radar Observation”	Method of production of the original data
Non-CF attribute	instrument	S	instrument = “WETLABS WQM”	Make and model of the instruments from which the data has been collected.
Non-CF attribute	instrument_serial_number	S	instrument_serial_number = “5124”	Serial number of the instrument which has produced the dataset.
	references	S	references =	Published or web-based references that describe the

			"http://www.imos.org.au"	data or the methods used to produce the data. Include a reference to IMOS and a project-specific reference if appropriate. Multiple references should be separated with a semicolon ";".
Non-CF attribute	site_code	S	site_code = "NRSMAI"	Unique site code within IMOS project. A site refers to a nominal location around which (within a certain radius) repeated measurements are performed to produce a time-series dataset.
Non-CF attribute	site	S	site = "Maria Island National Reference Station, TAS"	Site description.
Non-CF attribute	platform_code	S	platform_code = "NRSMAI-SubSurface"	Unique platform code within IMOS project. The platform codes are listed in Reference Table 3 of the File Naming Convention document. A platform refers to the actual structure to which the instrument or sensor is attached.
Non-CF attribute	platform	S	platform = "Subsurface mooring at Maria Island National Reference Station, TAS site"	Platform description. They are listed in Reference Table 3 of the File Naming Convention document.
Non-CF attribute	cdm_data_type	S	cdm_data_type = "Station"	The "cdm_data_type" attribute gives the Unidata CDM (Common Data Model) data type used by THREDDS. E.g. "Point", "Trajectory", "Station",

				“Radial”, “Grid”, “Swath”.
Non-CF attribute	keywords	S	keywords = “Oceans > Ocean Circulation > Ocean Currents , Oceans > Ocean Waves > Significant Wave Height ,…”	A comma separated list of keywords coming from the keywords_vocabulary.
Non-CF attribute	keywords_vocabulary	S	keywords_vocabulary = “NASA/GCMD Earth Science Keywords”	Identifies the controlled keyword vocabulary used to specify the values within the attribute "keywords".
Non-CF attribute	metadata	S	metadata = “ http://.....”	URL to the metadata record corresponding to the netCDF file. AODN is considering adding this information to the data files as part of AODN data processing. Facilities would not be required to complete these fields when submitting data.
Non-CF attribute	sensorML	S	sensorML = “ http://.....”	Link to the sensorML record corresponding to the netCDF file. AODN is considering adding this information to the data files as part of AODN data processing. Facilities would not be required to complete these fields when submitting data.
	institution_address	S	institution_address = “Oceanography laboratory, School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052”	Address of the institute or facility where the original data was produced.

	institution_postal_address	S	institution_postal_address = “University of New South Wales, Sydney NSW 2052”	Postal address of the institute or facility where the original data was produced.
Non-CF attribute	file version	S	file_version = “Level 0 - Raw data”	Information about the file version of the file. Three levels are possible at the moment: - Level 0 - Raw data - Level 1 - Quality Controlled data - Level 2 - Derived product
Non-CF attribute	file_version_quality_control	S	file_version_quality_control = “Data in this file has not been quality controlled”	Description of the level of Quality Control applied to the data. See section **Error! Reference source not found.**for more information and examples.
Non-CF attribute	quality_control_log	S	quality_control_log = "imosImpossibleDateQC(dateMin=01/01/2007, dateMax=15/07/2014) did not fail on any TIME sample.\nimosImpossibleLocationSetQC(distanceKmPlusMinusThreshold=2.5) did not fail on any LATITUDE sample.\nimosImpossibleLocationSetQC(distanceKmPlusMinusThreshold=2.5) did not fail on any LONGITUDE sample....”	QC procedures applied with their specific parameters and results.
Non-CF attribute	**Where**			
	geospatial_lat_units	S	geospatial_lat_units = “degrees_north”	Units used for geospatial_lat_min/max

				attributes.
	geospatial_lon_units	S	geospatial_lon_units = “degrees_east”	Units used for geospatial_lon_min/max attributes.
	geospatial_vertical_units	S	geospatial_vertical_units = “metres”	Units used for geospatial_vertical_min/max attributes.
Non-CF attribute	**When**			
	local_time_zone	N	local_time_zone = 10	Local time zone at the location of the dataset. See section 3.2.4 - on time format. If local time does not fall into one zone for the full dataset, do not use this attribute.
Non-CF attribute	**Who**			
	author_email	S	author_email = “info@aodn.org.au”	NetCDF file author contact e-mail address
	principal_investigator_email	S	principal_investigator_email = “john.doe@utas.edu.au”	Principal Investigator e-mail address
	institution_references	S	institution_references = “http://imos.org.au/facilities/aodn/”	References that describe the data provider institution, the place to find all information on the dataset (web-based, i.e. give URLs). Multiple references should be separated with a semicolon “;”.
	How			
	lineage	S	lineage = “Data were processed according to standard ANMN-NRS procedures before	Information about how the data has been produced and processed, modified.

		conversion to netCDF. See link for details : http://help.aodn.org.au/help/sites/help.aodn.org.au/files/ANMN%20CTD%20Processing%20Procedures.pdf"	

3.2.4 -Date and Time formats

All time values in IMOS netCDF files will be recorded as Coordinated Universal Time (UTC, for most purposes equivalent to Greenwich Mean Time).

Whenever time information is given in the global attributes, it will be formatted as a string according to the ISO 8601 standard (International Organization for Standardization, 2009): "YYYY-MM-DDThh:mm:ssZ" (i.e. year – month – day "T" hour : minute : second "Z") The "Z" indicates the time zone is (zero offset from) UTC. If higher resolution than seconds is needed, any number of decimal digits (".s") for the seconds is acceptable: "YYYY-MM-DDThh:mm:ss.sZ". Examples of the time format are shown below.

2005-10-24T08:00:00Z

2008-01-01T22:50:02.03Z

The representation of time as a numerical variable is described under the TIME variable chapter 3.4.1.1 -.

The global attribute 'local_time_zone' gives the time zone at the location where the measurements were collected. This allows the UTC times in the file to be converted into local time, which is important in considering many biological processes and phenomena that depend on the diurnal cycle.

The local time zone will be recorded as an offset in hours from UTC. For example Australian Eastern Standard Time is represented as 10, meaning 10 hours ahead of UTC.

Offsets from local to UTC time for Australian time zones can be found online.

The global attribute 'local_time_zone' should be used only when all data points in a dataset are from the same time zone. If time zone changes during the dataset (e.g. moving point measurements on a ship of opportunity) then do not use this global attribute. Instead, advanced users may choose to create a user-defined variable called 'local_time_zone'. If local time zone is not defined in the dataset, it will be calculated by data users from the latitude, longitude and UTC time variables.

3.2.5 -Processing level

Two attributes <file_version> and <file_version_quality_control> respectively describe the title and definition of level of processing and Quality Control (QC) found in the dataset.

Three possibilities are available for these two attributes:

Global attribute <file_version>	Global attribute <file_version_quality_control>
Level 0 - Raw data	Raw data is defined as unprocessed data and data products that have not undergone quality control. The data may be in engineering or physical units, time and location details can be in relative units and values can be pre-calibration measurements.
Level 1 – Quality Controlled data	Quality controlled data have been through quality assurance procedures such as automated routines and sensor calibration and/or a level of visual inspection and flag of obvious errors. The data are in physical units using standard SI metric units with calibration and other pre-processing routines applied, all time and location values are in absolute coordinates to comply with standards and datum. Data includes flags for each measurement to indicate the estimated quality of the measurement. Metadata exists for the data or for the higher level dataset that the data belongs to. This is the standard IMOS data level and is what should be made available to AODN and to the IMOS community.
Level 2 – Derived Products	Derived products require scientific and technical interpretation. Normally these will be defined by the community that collects or utilises the data.

3.2.6 -Global Attributes added by the data provider

The global attributes listed in the table in section 3.2.2 - are most important to define a dataset as clearly as possible. However, this list will not in all cases be exhaustive and AODN requests that other meaningful global attributes be used where necessary.

It is possible to add global attributes to meet specific facility needs. New attributes will need to be self defined, including a description and an example of how it is used. User-defined global attributes should be added to the existing list (appropriate appendix section) in the next version of the IMOS NetCDF Conventions document.

An example of a 'facility defined' global attribute and supporting information is below:

This example can be used by the ACORN facility to illustrate the kind of data and the type of radar used to produce the data. The prefix "ssr" added to each attribute means "sea surface radar".

ssr_Data_Type = "Range_Time_Series"

ssr_Radar = "Helzel/WERA"

3.3 - Dimensions

NetCDF file dimensions provide information on the size of the data variables. An unlimited dimension will make extension (new data added to an existing file) and aggregation (distinct files piled together into a single one) possible later if needed along this dimension. A dimension of size 1 can also be replaced by a scalar variable. The example in Table 3 allows for measurements at an unlimited number of time steps, at five different static depths.

Requirements are described further in the section 3.4.1 -.

Table 3 An example list of Dimensions for an IMOS netCDF file

Name	Example	Comment
TIME	TIME = unlimited	Number of time steps
DEPTH	DEPTH = 5	Number of depth levels

3.4 - Variables

NetCDF variables include data measured by instruments, parameters derived from the primary measurements and coordinate variables, which may be nominal values such as values for depth for instruments that do not directly record depth. Variables are also used to store quality flags associated to a measurement, and other metadata that is not applicable to all the data in a file. Ways to represent each type of variable in an IMOS netCDF file are described in the sub-sections below. Defined variable names are listed in chapter 5.1 -. Each variable has a specific set of attributes, some of which are mandatory.

In the tables below for variable attributes, the “Type” values are **S** for string, **N** for numeric (byte, short, long, integer, float or double), **D** for the type of the data variable and **Q** for the type of the quality control data variable. **Mandatory attributes** are marked **in bold**.

3.4.1 -Spatio-temporal variables

The spatio-temporal variables locate the data in time and space. For this purpose, they have an “axis” attribute to indicate that they represent either latitude (axis = “X”), longitude (“Y”), depth/height (“Z”) or time (“T”). Any measured parameter can be located in space and time by listing the relevant spatio-temporal variables in its coordinates attribute. The use of a common set of spatial and temporal units and measures is the basic requirement to be able to integrate the various data collected by the IMOS project.

These variables can be coordinate variables[1] (representing an independent dimension of the data in the file), auxiliary coordinate variables[2] (dependent on one or more dimensions in the file) or scalar coordinate variables[3] depending on the CF feature type template chosen. For more details on ways to associate data variables with their spatio-temporal coordinates, see chapters 5 and 9 of the CF conventions document.

3.4.1.1 - Time

In a netCDF file, time is represented numerically as an interval (e.g. number of days or hours) from a reference time. For IMOS data, AODN recommends using a decimal number of days since the ARGO reference time of 1st of January 1950 at 00:00:00 UTC. In these conditions, using single precision would only achieve a 30s resolution so we recommend that any time measurement is stored in double precision using the type Double.

The following table presents the attributes used to describe the time variable.

Table 4 List of attributes defining the TIME variable for IMOS netCDF files

[1] It is a one-dimensional variable with the same name as its dimension [e.g., TIME(TIME)], and it is defined as a numeric data type with values that are ordered monotonically. Missing values are not allowed in coordinate variables.

[2] Any netCDF variable that contains coordinate data, but is not a coordinate variable (see above). Unlike coordinate variables, there is no relationship between the name of an auxiliary coordinate variable and the name(s) of its dimension(s).

[3] A scalar variable that contains coordinate data. Functionally equivalent to either a size one coordinate variable or a size one auxiliary coordinate variable.

Attributes	Type	Example	Comment
standard_name	S	standard_name="time"	A CF standard name that references a description of a variable's content in the standard name table
long_name	S	long_name = "time"	A descriptive name that indicates a variable's content.
units	S	units = "days since 1950-01-01 00:00:00 UTC"	This is AODN's recommended unit and reference time for IMOS netCDF files. Please contact AODN if you require help with this conversion.
axis	S	axis = "T"	Identifies time axes
valid_min	D	valid_min = 0.0	Smallest valid value of a variable. Should be of the same type as the variable type.
valid_max	D	valid_max = 90000.0	Largest valid value of a variable. Should be of the same type as the variable type.
_FillValue	D	_FillValue = 999999.0	A value used to represent missing or undefined data[4]. Cannot be used for coordinate variables. Required if there could be missing values in the data. Should be of the same type as the variable type.
calendar	S	calendar = "gregorian"	Calendar used for encoding time axes. See Eaton et. al. 2009 for clarification.

[4] We ask that NaN "not a number" value shouldn't be used. Instead use a decimal value e.g. _FillValue of 999999.0.

	comment	S	comment = "……."	Miscellaneous information about the data or methods used to produce it.
Non-CF attribute	uncertainty	N	uncertainty = 0.00001	Overall measurement uncertainty. Choose appropriate value. See section 3.4.4.2 -.
Non-CF attribute	local_time_zone	N	local_time_zone = 10	Local time zone. See chapter 3.2.4 -. If local time does not fall into one zone for the full dataset, do not use this attribute.

See APPENDIX 1: example netcdf file.

3.4.1.2 - Horizontal coordinates

Horizontal coordinates (Latitude and Longitude) in an IMOS netCDF file will be specified in decimal degrees relative to the WGS84 coordinate reference system, with locations south of the equator having negative Latitude and locations west of zero degrees of Longitude having negative Longitude.

Using single precision would only achieve at worst an 11m resolution while modern GPS can provide <1m resolution. We recommend that any measurement of geographic coordinates is stored in double precision using the type Double.

The table below presents the attributes used to describe the variables LATITUDE and LONGITUDE.

Table 5 List of attributes that define the LATITUDE and LONGITUDE variables for IMOS netCDF files

Attributes	Type	Example	Comment
standard_name	S	standard_name = "latitude" or standard_name = "longitude"	A CF standard name that references a description of a variable's content in the standard name table

long_name	S	long_name = “latitude” or long_name = “longitude”	A descriptive name that indicates a variable’s content. This name is the IMOS long name given in chapter 5.1 -.
units	S	units = “degrees_north” (LATITUDE) or units = “degrees_east” (LONGITUDE)	Latitude unit: degrees north ; southern latitudes are negative Longitude unit: degrees east ; western longitudes are negative
axis	S	axis = “ Y ” (LATITUDE) or axis = “ X ” (LONGITUDE)	Identifies type of spatial axis
valid_min	D	valid_min = -90.0 (LATITUDE) valid_min = -180.0 (LONGITUDE)	Smallest valid value of a variable. Should be of the same type as the variable type.
valid_max	D	valid_max = 90.0 (LATITUDE) valid_min = 180.0 (LONGITUDE)	Largest valid value of a variable. Should be of the same type as the variable type.
_FillValue	D	_FillValue = 999999.0	A value used to represent missing or undefined data[5]. Cannot be used for coordinate variables. Required if there could be missing values in the data. Should be of the same type as the variable type.

[5] We ask that NaN “not a number” value shouldn’t be used. Instead use a decimal value e.g. _FillValue of 999999.0.

	comment	S	comment = " ……."	Miscellaneous information about the data or methods used to produce it.
Non-CF attribute	uncertainty	N	uncertainty = 0.001	Overall measurement uncertainty. Choose appropriate value. See section 3.4.4.2 -.
Non-CF attribute	**reference_datum**	S	reference_datum = "WGS84 coordinate reference system"	Text description of the geographic reference datum for the variable

See APPENDIX 1: example netcdf file.

3.4.1.3 - Vertical coordinates

For any vertical coordinate, variable attributes must define a reference point or datum (attribute "reference_datum"), the direction of increasing values (attribute "positive", up or down) and the unit of measure (attribute "units"). Four datums are recognised:

-Mean Sea Level (MSL),

-sea surface,

-sea bottom,

-sensor.

The time and location of the instrument will need to be included to allow for tide corrections. When necessary, vertical measurements should be converted into one of the acceptable datums.

For example: depth may be recorded as metres below MSL, positive down.

Depth and height should be measured in metres or other SI units. Pressure measurements should be recorded in a separate parameter. It is not permissible to label a pressure measurement as "depth" or "height".

The following table presents the attributes used to describe any depth or height variable.

Table 6 List of attributes that define any depth or height variable for IMOS netCDF files

Attributes	Type	Example	Comment

standard_name	S	standard_name = “depth” or “height”	A CF standard name that references a description of a variable’s content in the standard name table
long_name	S	long_name = “depth” or “height”	A descriptive name that indicates a variable’s content. This name is the IMOS long name given in chapter 5.1 -
units	S	units = “ m ”	Usually “m” for metres.
axis	S	axis = “ Z ”	Identifies vertical axis
positive	S	positive = “down ” (depth) or “up” (height)	Direction of increasing vertical coordinate value.
valid_min	D	valid_min = 0.0	Smallest valid value of the variable. Should be of the same type as the variable type.
valid_max	D	valid_max = 12000.0	Largest valid value of the variable. Should be of the same type as the variable type.
_FillValue	D	_FillValue = -99999.0	A value used to represent missing or undefined data[6]. Cannot be used if DEPTH/HEIGHT is a dimension. Required if there could be missing values in the data. Should be of the same type as the variable type.

[6] We ask that NaN “not a number” value shouldn’t be used. Instead use a decimal value e.g. _FillValue of 999999.0.

	comment	S	comment = “depthPP: Depth computed using the SeaWater toolbox from latitude and relative pressure measurements…”	Miscellaneous information about the data or methods used to produce it.
Non-CF attribute	uncertainty	N	uncertainty = 0.05	Overall measurement uncertainty. Choose appropriate value. See section 3.4.4.2 -.
Non-CF attribute	**reference_datum**	S	reference_datum = “Mean Sea Level (MSL)”	Text description of the reference datum for the variable

See APPENDIX 1: example netcdf file.

3.4.2 -Data variables

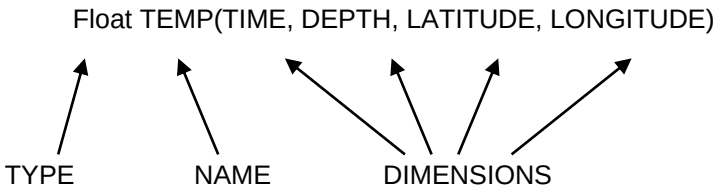
IMOS recommended variable names are listed in chapter 5.1 -. For example, TEMP represents the sea water temperature and PSAL represents the sea water salinity. Variable names should begin with a letter and be composed of letters, digits and underscores.

In some cases, two instruments may measure the same variable (e.g. wind speed measured by two anemometers mounted on the same vessel). In these cases, the second instance of a variable should be identified with the suffix ‘_2’ (e.g. “WSPD” and “WSPD_2”). Several variable attributes can be used to specify the differences, including comments, descriptive attributes (sensor_depth, sensor_height) and long names. Two anemometers might be distinguished in their long names as “wind_speed_starboard” and “wind_speed_port”.

The definition of a variable in a netCDF file has three components: type, name and dimensions.

CF says: “If any or all of the dimensions of a variable have the interpretations of "date or time" (T), "height or depth" (Z), "latitude" (Y), or "longitude" (X) then it is recommended, but do not required, that those dimensions appear in the relative order T , then Z , then Y , then X in the CDL definition corresponding to the file. All other dimensions should, whenever possible, be placed to the left of the spatiotemporal dimensions”.

Example for the sea water temperature output by a 3D model over a time period:



The following table presents the attributes used to describe a data variable.

Table 7 List of data variables attributes for IMOS netCDF files

Attributes	Type	Example	Comment
standard_name	S	standard_name = “sea_surface_temperature”	A CF standard name that references a description of a variable’s content in the CF standard name table.
units	S	units = “Celsius”	Unit of measure
_FillValue	D	_FillValue = 999999.0	A value used to represent missing or undefined data[7]. Required if there could be missing values in the data. Should be of the same type as the variable type.
long_name	S	long_name = “sea_surface_temperature”	A descriptive name that indicates a variable’s content. This name is the IMOS long name given in chapter 5.1 -. Please contact AODN regarding any parameters that don’t appear in the table
coordinates	S	coordinates = "TIME LATITUDE LONGITUDE DEPTH" ;	A blank-separated list of the names of the relevant variables that include spatio-temporal coordinate information.

[7] We ask that NaN “not a number” value shouldn’t be used. Instead use a decimal value e.g. _FillValue of 999999.0.

	valid_min valid_max	D	valid_min = -2.0 valid_max = 40.0	Range of values for valid data. Should be of the same type as the variable type.
	add_offset	N	add_offset = 25.0	If present for a variable, this number is to be added to the data after it is read by an application. If both scale_factor and add_offset attributes are present, the data are first scaled before the offset is added. Should be of the same type as the intended unpacked variable type. See below.
	scale_factor	N	scale_factor = 0.01	If present for a variable, the data are to be multiplied by this factor after the data are read by an application. Should be of the same type as the intended unpacked variable type. A packed variable can be of type integer and yet will be unpacked to a type double if the scale_facor is of type double.
Non-CF attribute	applied_offset	N	applied_offset = -10.1325	If present for a variable, this offset has been added to the original data to correct it. Should be of the same type as the variable type.
Non-CF attribute	original_units	S	original_units = “mg/l”	Original units of a variable’s content in the original instrument file. Dissolved

				oxygen concentration for example can be expressed in many units but umol l^{-1} is IMOS preferred units for dissemination.
Non-CF attribute	original_name	S	original_name = “oxsolMg/L”	Original name of a variable in the original instrument file.
Non-CF attribute	sensor_serial_number	S	sensor_serial_number = “DO142”	Serial number of the sensor which collected the data of this variable.
	comment	S		Miscellaneous information about the this variable or methods used to produce it
	history	S		As for global attribute history, but specific to this variable.
	references	S		References that describe the data or methods used to produce it. Multiple references should be separated with a semicolon.
	ancillary_variables	S	ancillary_variables = “TEMP_quality_control TEMP_uncertainty”	Blank-separated list of variables that contain closely associated data, e.g. the measurement uncertainties of instrument data. See section 3.4.4 -.
Non-CF attribute	sensor_depth	N	sensor_depth = 0	Nominal sensor depth(s) in metres positive down.
Non-CF attribute	sensor_height	N	sensor_height = 2	Nominal sensor height(s) in metres positive up.

Non-CF attribute	observation_type	S	observation_type = “measured”	Type of observation. If for example, the variable is measured or calculated.
Non-CF attribute	uncertainty	N	uncertainty = 0.001	Overall measurement uncertainty, if constant. See section 3.4.4.2 -.
Non-CF attribute	accuracy	N	accuracy = 0.01	Nominal sensor accuracy. See section 3.4.4.2 -.
Non-CF attribute	precision	N	precision = 0.01	Nominal sensor precision. See section 3.4.4.2 -.
Non-CF attribute	resolution	N	resolution = 0.01	Nominal resolution of this data parameter. See section 3.4.4.2 -.
	cell_methods	S	cell_methods = “point”	Records the method used to derive data that represents cell values. See section 5.3 -

See APPENDIX 1: example netcdf file.

3.4.3 -Variable attributes added by data provider

Additional variable attributes can be defined by the data provider (in consultation with AODN). Such attributes will be incorporated into the next version of this document.

3.4.4 -Ancillary variables

When one variable provides metadata about the individual values of another variable it may be desirable to express this association by providing a link between the variables. For example, instrument data may have associated measures of uncertainty or quality control flags. The attribute “ancillary_variable” is used to express these types of relationships.

The use of ancillary variables in the context of data Quality Control and uncertainty are described in sections 3.4.4.1 - and.3.4.4.2 -.

3.4.4.1 - Quality control (QC)

Quality control involves some sort of assessment of the data to identify data points or even data sets which have errors that limit their use. The basic approach used by IMOS is to keep all of the data but to flag data or data sets that do not meet the quality assessment standards of data collectors / principal investigators. This section describes how to represent this data quality information in an IMOS netCDF file.

3.4.4.1.1 - Quality Control sets used by the IMOS project

The attribute "quality_control_conventions**"** will enable users to define which set of quality control (QC) flags was applied to the dataset. Please see chapter 5.2 -. Additional QC sets may be added in future versions of this document.

3.4.4.1.2 - Definition of quality control variables and attributes

Quality flags for the data in a variable <PARAM> should be described by an ancillary variable named "<PARAM>_quality_control".

The following table lists all the attributes used to define a Quality Control variable. It also includes an example using the quality control set 1 (IMOS standard flags).

Table 8 List of attributes used to define a Quality Control variable

Attributes	Type	Example	Comment
long_name	S	long_name = "quality flag for sea_surface_temperature"	Non-standardised name of variable
standard_name	S	standard_name = "sea_surface_temperature status_flag"	Standardised name (status_flag suffix) of QC variable using the CF convention

quality_control_conventions	S	quality_control_conventions =”IMOS standard flags”	Name of the Quality Control convention used. See Reference table A.
_FillValue	D	_FillValue = 99b	Value used to represent missing QC flags
flag_values	D	flag_values = 0b, 1b, 2b, 3b, 4b, 5b, 6b, 7b, 8b, 9b	List of flag values used.
flag_meanings	S	flag_meanings = "No_QC_performed Good_data Probably_good_data Bad_data_that_are_potentially_correctable Bad_data Value_changed Not_used Not_used Not_used Missing_value"	The meaning of each flag (in the same order as flag_values).
quality_control_global_conventions	S	quality_control_global_conventions = "Argo reference table 2a (see http://www.cmar.csiro.au/argo/dmqc/ user_doc/QC_flags.html), applied to data in position only (between global attributes time_deployment_start and time_deployment_end)"	Convention used to describe the global quality of a variable.
quality_control_global	S	quality_control_global = "A"	Global quality control value. See 5.2.2.2 -

See APPENDIX 1: example netcdf file.

3.4.4.2 - Uncertainty, accuracy, precision and resolution

The term uncertainty is here defined as “the parameter, associated with the result of a measurement that characterises the dispersion of the values that could reasonably be attributed to the measurand” (Underwood, 2008).

The document "IMOS Data Streams and their Uncertainties" (Underwood, 2008) contains a calculation or estimation of the uncertainty for each data stream that is provided by IMOS..

3.4.4.2.1 - Implementation in IMOS

If the overall measurement uncertainty for a variable <PARAM> is reasonably well-known, it must be provided in the attributes. If it is constant is should be provided in the attribute <PARAM>:uncertainty. If not constant it should be provided in a variable of its own, called <PARAM>_uncertainty. The standard name for this variable is formatted as: "<parameter_standard_name> standard_error" (e.g. TEMP_uncertainty:standard_name = "sea_surface_temperature standard_error")

If it is impossible to estimate the overall measurement uncertainty, the attribute <PARAM>:accuracy can be set to the nominal sensor accuracy. Accuracy characterises how close the measurement of a quantity is to the value of that quantity (systematic errors reduce accuracy).

The attributes <PARAM>:precision and <PARAM>:resolution contain the sensor precision and resolution if defined. Precision characterises the level of agreement between repeated measurements of a quantity under unchanged conditions (random errors reduce precision). Resolution is a limit on the precision of a measurement imposed by the amount of information recorded (e.g. number of decimal places, or number of measurements in a time interval).

3.4.4.2.2 - Example

This example is extracted from the "IMOS Data Streams and their Uncertainties" document (Underwood, 2008).

A temperature measurement may be given as "20.12 degrees Celsius with a 95% confidence of 0.01 degree Celsius". To rephrase this, there are 5 chances in one hundred that the real temperature (the measurand) was outside the range 20.12 C +- 0.01 C.

In a netCDF file, the uncertainty on this temperature measurement will appear as:

TEMP:uncertainty = 0.01

4 - IMOS FILE NAMING CONVENTION

NetCDF files will be named according to the IMOS NetCDF File Naming Convention document. See APPENDIX 3: IMOS File Naming Convention.

5 - REFERENCE TABLES

5.1 - IMOS parameter dictionary

The IMOS parameter dictionary seeks to list recommended netCDF variable names, long_name, standard_name and units attributes to be used.

The most up to date list of parameter names, standard names (or long names for non-CF parameters) and units can be found on the IMOS Toolbox GitHub repository. This list is not exhaustive and will continue to expand and evolve. Ultimately, it will constitute an IMOS data parameter dictionary. The standard names used in the IMOS parameter dictionary are taken from the CF standard names.

Non-CF long names and variable names are derived from the following sources:

- a list of parameter names available on the WOCE website using GF3 codes (WOCE, 2009).
- a parameter dictionary available on the OceanSites data format reference manual (OceanSites, 2008).

Non-CF parameters (with no CF standard name) should be described using only the attribute long_name.

All units defined for the parameters must comply with Udunits (Unidata Program Centre of the University Corporation for Atmospheric Research, 2008) as implemented by the CF standard.

5.2 - Quality control flags

Quality control (QC) flags are added to a file by various procedures to indicate the quality of individual data values.

The attribute “quality_control_set” enables the user to define which Quality Control flags set was used in the dataset. Please see section 3.4.4.1 - for definitions of Quality Control sets.

If your facility is using a different set of quality control codes please supply AODN with the details. If appropriate, we will incorporate your codes into the IMOS convention and create a QC set that meets your needs.

Reference Table A : List of QC flags sets used in the IMOS project

Set Number	Description
1	IMOS standard flags (GTSPP data quality codes)
2	ARGO quality control procedure
3	BOM (SST and Air-Sea flux) quality control procedure
4	WOCE quality control procedure (Multidisciplinary Underway Network – CO_2 measurements)

5.2.1 -Set 1 - IMOS standard flags

Reference Table B : IMOS standard flags

Flag Value	Meaning	Description
0	No QC performed	The level at which all data enter the working archive. They have not yet been quality controlled
1	Good data	Top quality data in which no malfunctions have been identified and all real features have been verified during the quality control process

2	Probably good data	Good data in which some features (probably real) are present but these are unconfirmed. Code 2 date are also data in which minor malfunctions may be present but these errors are small and/or can be successfully corrected without seriously affecting the overall quality of the data.
3	Bad data that are potentially correctable	Suspect data in which unusual, and probably erroneous features are observed
4	Bad data	Obviously erroneous values are observed
5	Value changed [8]	Altered by a QC Centre, with original values (before the change) preserved in the history record of the profile
6	Not used	Flag 6 is reserved for future use
7	Not used	Flag 7 is reserved for future use
8	Not used	Flag 8 is reserved for future use
9	Missing value	Indicates that the element is missing

5.2.2 -Set 2 - ARGO quality control procedure

5.2.2.1 - ARGO measurement flag scale

Reference Table C : ARGO measurement flag scale

Flag Value	Meaning	Real-time comment	Delayed-mode comment
0	No QC was performed	No QC was performed	No QC was performed

[8] Where data values must be changed (e.g. smoothing of data sets) we strongly prefer that the original data be retained and an additional variable be added to accommodate the interpolated/corrected data values.

1	Good data	All ARGO real-time QC tests passed	The adjusted value is statistically consistent and a statistical error estimate is supplied
2	Probably good data	Probably good data	Probably good data
3	Bad data that are potentially correctable	Argo QC tests (15, 16 or 17, see Carval et al 2008) failed and all other real-time QC tests passed. These data are not to be used without scientific correction. A flag 3 may be assigned by an operator during additional visual QC for bad data that may be corrected in delayed mode	An adjustment has been applied, but the value may still be bad
4	Bad data	Data have failed one or more of the real-time QC tests, excluding Test 16 (see Carval et al 2008). A flag 4 may be assigned by an operator during additional visual QC for bad data that are not correctable	Bad data. Not adjustable
5	Value changed [9]	Value changed	Value changed
6	Not used	Not used	Not used
7	Not used	Not used	Not used
8	Interpolated value	Interpolated value	Interpolated value

[9] Where data values must be changed we strongly prefer that the original data be retained and an additional variable be added to accommodate the interpolated/corrected data values.

9	Missing value	Missing value	Missing value

5.2.2.2 - ARGO profile quality flags

Argo profile quality flags (A to F) are defined as the percentage of levels (N) with good data, where:

- QC flag values of 1, 2, 5 or 8 are counted as GOOD data
- QC flag values of 9 (missing) are NOT USED in the computation
- All other QC flag values are counted as BAD data

The computation should be taken from <PARAM>_ADJUSTED_QC if available and from <PARAM>_QC otherwise (Carval et al 2008).

This could also be applied to any type of data, not only profile data.

Reference Table D : ARGO profile quality flags

Flag	Meaning
" " (blank)	No QC performed
A	N = 100%; All profile levels contain good data
B	75% <= N < 100%
C	50% <= N < 75%
D	25% <= N < 50%
E	0% <= N < 25%
F	N = 0%; No profile levels have good data

Example:

A TEMP profile has 60 levels (3 levels contain missing values)

- 45 levels are flagged as 1
- 5 levels are flagged as 2
- 7 levels are flagged as 4

- 3 levels are flagged as 9 (missing)

Percentage of good levels = ((45+5)/57)*100 = 87.7%

TEMP_quality_control:quality_control_global = "B"

5.2.3 -Set 3 - BOM quality control procedure (SST and Air-Sea fluxes)

The following table summarises the different flags used by the Bureau of Meteorology to qualify the quality for different datasets, particularly SST and Air-Sea fluxes (Verein 2008).

Reference Table E : BOM Quality Control procedure flags (SST and Air Sea Fluxes)

Flag	Purpose
B	Value out of bounds
C	Time not sequential
D	Failed T, Tw and Td tests
E	Failed True wind recomputation test
F	Platform velocity unrealistic
G *	Value exceeds threshold
H **	Discontinuity
J	Erroneous value
L	Value located over land
M	Instrument malfunction
Q	Pre-flagged as suspect
S	Spike in data (visual)
T	Time duplicate
U *	Suspect data (statistical)
V *	Step in data (statistical)
X *	Spike in data (statistical)
Z	Value passed all test

Note: * - applied for SST, ** - applied for time

5.2.4 -Set 4 - WOCE quality control procedure (Multidisciplinary Underway Network – CO_2 measurements)

Reference Table F : WOCE quality control procedure flags (CO_2 measurements)

Flag	Meaning
2	good
3	questionable
4	bad

If data is flagged as questionable (flag number 3), the variable SUBFLAG is used to add more information.

Table FF: SUBFLAG variable for questionable measurements

Flag	Meaning
1	Outside of standard range
2	questionable/interpolated SST (Sea Surface Temperature)
3	questionable EQU temperature
4	Anomalous (EQU T-SST) (+- 1 C)
5	questionable sea-surface salinity
6	questionable pressure
7	low EQU gas flow
8	questionable air value
10	other, water flow

References: "Pierrot,D. et al. 2009, Recommendations for Autonomous Underway pCO2 Measuring Systems and Data Reduction Routines, Deep-Sea Research II, doi:10.1016/j.dsr2.2008.12.005"

5.3 - Cell methods

CF Cell methods are applied in generating the parameter value for a cell, particularly for gridded data sets.

In the Units column, u indicates the units of the physical quantity before the method is applied.

Reference Table G : List of different cell methods, derived from the CF convention

Cell method	Units	Description
point	u	The data values are representative of points in space or time (instantaneous). This is the default method for a quantity that is intensive with respect to the specified dimension.
sum	u	The data values are representative of a sum or accumulation over the cell. This is the default method for a quantity that is extensive with respect to the specified dimension
maximum	u	Maximum
median	u	Median
mid_range	u	Average of the maximum and minimum
minimum	u	Minimum
mean	u	Mean (average value)
mode	u	Mode (most common value)
standard_deviation	u	Standard deviation
variance	u^2	Variance

6 - REFERENCES

Bainbridge, S. 2008. *Data Standards Framework for the IMOS instrument Data.* AODCJF Internal Report.

Carval, T., Keeley, B., Takatsuki, Y, Yoshida, T., Lock, S., Schmid, C., Goldsmith, R., Wong, A., McCreadie, R., Thresher, A and Tran, A. 2008. *Argo Data Management User's Manual Version 2.1*. IFREMER.

Convention for the standardisation of NetCDF Files. 1995. Sponsored by the "Cooperative Ocean/Atmosphere Research Data Service, a NOAA/university cooperative for the sharing and distribution of global atmospheric and oceanographic research data sets. http://ferret.wrc.noaa.gov/noaa_coop/coop_cdf_profile.html. Accessed December 2008.

Eaton, B., Gregory, J., Drach, B., Taylor, K, Hnakin, S., Caron, J., Signell, R., Bentley, P. Rappa, G., Höck H., Pamment A. and Juckes M. 2009. *NetCDF Climate and Forecast (CF) Metadata Conventions v1.6* http://cf-pcmdi.llnl.gov/. Accessed December 2011.

International Organization for Standardization. 2009. *Numeric representation of Dates and Time.* http://www.iso.org/iso/support/faqs/faqs_widely_used_standards/widely_used_standards_other/date_and_time_format.htm. Accessed 27 Jan 2009

Mancini, S., Tattersall, K., and Proctor, R. 2008. *IMOS NetCDF File Naming Convention v1.3* http://imos.org.au/facility_manuals.html. Accessed 31 January 2009.

OceanSITES. 2007. *Taking the Pulse of the Global Ocean* http://www.oceansites.org. Accessed 27 January 2009.

OceanSites 2008. *OceanSites Parameters Table*. http://www.ifremer.fr/co/etc/oceansites/oceansites-user-manual-parameters.pdf . Accessed 27 January 2009.

Olsen, L.M, G. Major, K. Shein, J. Scialdone, R. Vogel, S. Leicester, H. Weir, S. Ritz, T. Stevens, M. Meaux,c. Salomon, R. Bilodeau, M. Holland, T. Northcutt, R.A. Restrepo, 2007. NASA/Global Change Master Directory (GCMD) Earth Science Keywords version 6.0.0.0.0. http://gcmd.nasa.gov/Resources/valids/archives/keyword_list.html . Accessed 18 Mar 2009.

Reed, G. 2008. *Marine Community Profile of ISO 19115 Version 1.4*. Australian Ocean Data Centre Joint Facility

Rew, R., Davis, G., Emmersson, S. and Davies, H. 2008. *UniData NetCDF User's Guide* http://www.unidata.ucar.edu/. Accessed 20 Dec 2008.

SeaDataNet 2009. *SeaDataNet measurand qualifier flags : list key L201.* http://seadatanet.maris2.nl/v_bodc_vocab/welcome.aspx. Accessed 27 Jan 2009.

Time and Date AS. 2009. *Time Zone in Hobart* http://www.timeanddate.com/worldclock/timezone.html?n=396. Accessed 27 Jan 2009.

Underwood, M. 2008. *IMOS Data streams and their Uncertainties v 3.1* http://imos.org.au/facility_manuals.html. Accessed 15 December 2008.

Unidata Program Centre of the University Corporation for Atmospheric Research. 2008. *Writing NetCDF Files: Best Practices* http://www.unidata.ucar.edu/software/netcdf/docs/BestPractices.html. Accessed 15 December 2008.

Unidata Program Centre of the University Corporation for Atmospheric Research. 2008. *UDUNITS Software Package*. http://www.unidata.ucar.edu/software/udunits/. Accessed 15 December 2008..

Verein, R. 2008. *Quality Control procedure for automated processing of IMOS ship and mooring observations (Facilities 2c, 2d and 3).* CAWCR. Australian Bureau of Meterology.

Wikipedia contributors. 2008. *Wikipedia, The Free Encyclopedia : ISO 8601* http://en.wikipedia.org/wiki/ISO_8601. Accessed Jan 20 2009.

World Ocean Circulation Experiment 2009. *WOCE Parameter Lists: Basic Parameters & Full List of Parameters with extended-GF3 codes.* http://woce.nodc.noaa.gov/woce_v3/wocedata_1/sss/documents/liste_param.htm . Accessed 27 January 2009.

APPENDIX 1: EXAMPLE NETCDF FILE

This is an example IMOS formatted netCDF file for ANMN timeseries temperature and pressure logger data.

```
netcdf IMOS_ANMN-NSW_TZ_20110620T125500Z_PH100_FV01_PH100-1106-Aqualogger-520PT-104_END-20110831T133000Z_C-20200703T041240Z {
dimensions:
        TIME = 22345 ;
variables:
        double TIME(TIME) ;
                TIME:axis = "T" ;
                TIME:calendar = "gregorian" ;
                TIME:long_name = "time" ;
                TIME:standard_name = "time" ;
                TIME:units = "days since 1950-01-01 00:00:00 UTC" ;
                TIME:valid_max = 90000. ;
                TIME:valid_min = 0. ;
        int TIMESERIES ;
                TIMESERIES:cf_role = "timeseries_id" ;
                TIMESERIES:long_name = "unique_identifier_for_each_timeseries_feature_instance_in_this_file" ;
        double LATITUDE ;
                LATITUDE:axis = "Y" ;
                LATITUDE:long_name = "latitude" ;
                LATITUDE:reference_datum = "WGS84 geographic coordinate system" ;
                LATITUDE:standard_name = "latitude" ;
                LATITUDE:units = "degrees_north" ;
                LATITUDE:valid_max = 90. ;
                LATITUDE:valid_min = -90. ;
        double LONGITUDE ;
                LONGITUDE:axis = "X" ;
                LONGITUDE:long_name = "longitude" ;
                LONGITUDE:reference_datum = "WGS84 geographic coordinate system" ;
                LONGITUDE:standard_name = "longitude" ;
                LONGITUDE:units = "degrees_east" ;
                LONGITUDE:valid_max = 180. ;
                LONGITUDE:valid_min = -180. ;
        float NOMINAL_DEPTH ;
                NOMINAL_DEPTH:axis = "Z" ;
                NOMINAL_DEPTH:long_name = "nominal depth" ;
                NOMINAL_DEPTH:positive = "down" ;
                NOMINAL_DEPTH:reference_datum = "sea surface" ;
                NOMINAL_DEPTH:standard_name = "depth" ;
                NOMINAL_DEPTH:units = "m" ;
                NOMINAL_DEPTH:valid_max = 12000.f ;
                NOMINAL_DEPTH:valid_min = -5.f ;
        float TEMP(TIME) ;
```

TEMP:ancillary_variables = "TEMP_quality_control" ;
TEMP:coordinates = "TIME LATITUDE LONGITUDE NOMINAL_DEPTH" ;
TEMP:_FillValue = 999999.f ;
TEMP:long_name = "sea_water_temperature" ;
TEMP:standard_name = "sea_water_temperature" ;
TEMP:units = "degrees_Celsius" ;
TEMP:valid_max = 40.f ;
TEMP:valid_min = -2.5f ;
byte TEMP_quality_control(TIME) ;
TEMP_quality_control:_FillValue = 99b ;
TEMP_quality_control:flag_meanings = "No_QC_performed Good_data Probably_good_data Bad_data_that_are_potentially_correctable Bad_data Value_changed Not_used Not_used Not_used Missing_value" ;
TEMP_quality_control:flag_values = 0b, 1b, 2b, 3b, 4b, 5b, 6b, 7b, 8b, 9b ;
TEMP_quality_control:long_name = "quality flag for sea_water_temperature" ;
TEMP_quality_control:quality_control_conventions = "IMOS standard flags" ;
TEMP_quality_control:quality_control_global = "A" ;
TEMP_quality_control:quality_control_global_conventions = "Argo reference table 2a (see http://www.cmar.csiro.au/argo/dmqc/user_doc/QC_flags.html), applied on data in position only (between global attributes time_deployment_start and time_deployment_end)" ;
TEMP_quality_control:standard_name = "sea_water_temperature status_flag" ;
float PRES(TIME) ;
PRES:ancillary_variables = "PRES_quality_control" ;
PRES:coordinates = "TIME LATITUDE LONGITUDE NOMINAL_DEPTH" ;
PRES:_FillValue = 999999.f ;
PRES:long_name = "sea_water_pressure" ;
PRES:standard_name = "sea_water_pressure" ;
PRES:units = "dbar" ;
PRES:valid_max = 12000.f ;
PRES:valid_min = -5.f ;
byte PRES_quality_control(TIME) ;
PRES_quality_control:_FillValue = 99b ;
PRES_quality_control:flag_meanings = "No_QC_performed Good_data Probably_good_data Bad_data_that_are_potentially_correctable Bad_data Value_changed Not_used Not_used Not_used Missing_value" ;
PRES_quality_control:flag_values = 0b, 1b, 2b, 3b, 4b, 5b, 6b, 7b, 8b, 9b ;
PRES_quality_control:long_name = "quality flag for sea_water_pressure" ;
PRES_quality_control:quality_control_conventions = "IMOS standard flags" ;
PRES_quality_control:quality_control_global = "A" ;
PRES_quality_control:quality_control_global_conventions = "Argo reference table 2a (see http://www.cmar.csiro.au/argo/dmqc/user_doc/QC_flags.html), applied on data in position only (between global attributes time_deployment_start and time_deployment_end)" ;
PRES_quality_control:standard_name = "sea_water_pressure status_flag" ;
float DEPTH(TIME) ;
DEPTH:ancillary_variables = "DEPTH_quality_control" ;
DEPTH:comment = "depthPP: Depth computed using the Gibbs-SeaWater toolbox (TEOS-10) v3.06 from latitude and absolute pressure measurements to which a nominal value for atmospheric pressure (10.1325 dbar) has been substracted." ;
DEPTH:coordinates = "TIME LATITUDE LONGITUDE NOMINAL_DEPTH" ;
DEPTH:_FillValue = 999999.f ;
DEPTH:long_name = "actual depth" ;
DEPTH:positive = "down" ;

DEPTH:reference_datum = "sea surface" ;
DEPTH:standard_name = "depth" ;
DEPTH:units = "m" ;
DEPTH:valid_max = 12000.f ;
DEPTH:valid_min = -5.f ;
byte DEPTH_quality_control(TIME) ;
DEPTH_quality_control:_FillValue = 99b ;
DEPTH_quality_control:flag_meanings = "No_QC_performed Good_data Probably_good_data Bad_data_that_are_potentially_correctable Bad_data Value_changed Not_used Not_used Not_used Missing_value" ;
DEPTH_quality_control:flag_values = 0b, 1b, 2b, 3b, 4b, 5b, 6b, 7b, 8b, 9b ;
DEPTH_quality_control:long_name = "quality flag for depth" ;
DEPTH_quality_control:quality_control_conventions = "IMOS standard flags" ;
DEPTH_quality_control:quality_control_global = "A" ;
DEPTH_quality_control:quality_control_global_conventions = "Argo reference table 2a (see http://www.cmar.csiro.au/argo/dmqc/user_doc/QC_flags.html), applied on data in position only (between global attributes time_deployment_start and time_deployment_end)" ;
DEPTH_quality_control:standard_name = "depth status_flag" ;

// global attributes:
:abstract = "NSW-IMOS Port Hacking 100m Mooring" ;
:acknowledgement = "Any users of IMOS data are required to clearly acknowledge the source of the material derived from IMOS in the format: \"Data was sourced from Australia\'s Integrated Marine Observing System (IMOS) - IMOS is enabled by the National Collaborative Research Infrastructure Strategy (NCRIS).\"" ;
:author = "Austin, Tim" ;
:author_email = "t.austin@unsw.edu.au" ;
:citation = "The citation in a list of references is: \"IMOS [year-of-data-download], [Title], [data-access-url], accessed [date-of-access]\"" ;
:comment = "Geospatial vertical min/max information has been filled using the DEPTH min and max." ;
:Conventions = "CF-1.6,IMOS-1.4" ;
:data_centre = "Australian Ocean Data Network (AODN)" ;
:data_centre_email = "info@aodn.org.au" ;
:date_created = "2020-07-03T04:12:40Z" ;
:deployment_code = "PH100-1106" ;
:disclaimer = "Data, products and services from IMOS are provided \"as is\" without any warranty as to fitness for a particular purpose." ;
:featureType = "timeSeries" ;
:file_version = "Level 1 - Quality Controlled Data" ;
:file_version_quality_control = "Quality controlled data have been through quality assurance procedures such as automated routines and sensor calibration or visual inspection and flag of obvious errors. The data are in physical units using standard SI metric units with calibration and other pre-processing routines applied, all time and location values are in absolute coordinates to comply with standards and datum. Data includes flags for each measurments to indicate the estimated quality of the measurement. Metadata exists for the data or for the higher level dataset that the data belongs to. This is the standard IMOS data level and is what should be made available to AODN and to the IMOS community." ;
:geospatial_lat_max = -34.1203333333 ;
:geospatial_lat_min = -34.1203333333 ;
:geospatial_lon_max = 151.22415 ;
:geospatial_lon_min = 151.22415 ;
:geospatial_vertical_max = 106.5675f ;
:geospatial_vertical_min = -0.4290929f ;
:geospatial_vertical_positive = "down" ;

:history = "2020-07-03T04:12:57Z - depthPP: Depth computed using the Gibbs-SeaWater toolbox (TEOS-10) v3.06 from latitude and absolute pressure measurements to which a nominal value for atmospheric pressure (10.1325 dbar) has been substracted." ;

:institution = "ANMN-NSW" ;

:institution_references = "http://www.imos.org.au/aodn.html" ;

:instrument = "Aquatec Aqualogger 520PT" ;

:instrument_nominal_depth = 104. ;

:instrument_nominal_height = 6. ;

:instrument_sample_interval = 300. ;

:instrument_serial_number = "023-564" ;

:keywords = "Aqualogger 520PT, TIME, TIMESERIES, LATITUDE, LONGITUDE, NOMINAL_DEPTH, TEMP, PRES, DEPTH" ;

:keywords_vocabulary = "IMOS parameter names. See https://github.com/aodn/imos-toolbox/blob/master/IMOS/imosParameters.txt" ;

:license = "http://creativecommons.org/licenses/by/4.0/" ;

:naming_authority = "IMOS" ;

:platform_code = "PH100" ;

:principal_investigator = "Roughan, Moninya" ;

:principal_investigator_email = "mroughan@unsw.edu.au" ;

:project = "Integrated Marine Observing System (IMOS)" ;

:quality_control_log = "imosImpossibleDateQC(dateMin=01/01/2007, dateMax=21/04/2017) did not fail on any TIME sample.\nimosImpossibleLocationSetQC(distanceKmPlusMinusThreshold=2.5) did not fail on any LATITUDE sample.\nimosImpossibleLocationSetQC(distanceKmPlusMinusThreshold=2.5) did not fail on any LONGITUDE sample.\nimosInOutWaterQC(in=20/06/11 12:55:00, out=31/08/11 13:30:00) flagged 1601 TEMP samples with flag Bad_data.\nimosInOutWaterQC(in=20/06/11 12:55:00, out=31/08/11 13:30:00) flagged 1601 PRES samples with flag Bad_data.\nimosInOutWaterQC(in=20/06/11 12:55:00, out=31/08/11 13:30:00) flagged 1601 DEPTH samples with flag Bad_data.\nimosGlobalRangeQC(min=-2.5, max=40) did not fail on any TEMP sample.\nimosGlobalRangeQC(min=-5, max=12000) did not fail on any PRES sample.\nimosGlobalRangeQC(min=-5, max=12000) did not fail on any DEPTH sample.\nimosImpossibleDepthQC(zNominalMargin=15, maxAngle=70 => min=99.7923, max=136.0703) did not fail on any PRES sample.\nimosImpossibleDepthQC(zNominalMargin=15, maxAngle=70 => min=89, max=125) did not fail on any DEPTH sample." ;

:references = "http://www.imos.org.au" ;

:site_code = "PH100" ;

:site_nominal_depth = 110. ;

:source = "Thermistor String" ;

:standard_name_vocabulary = "NetCDF Climate and Forecast (CF) Metadata Convention Standard Name Table 45" ;

:time_coverage_end = "2011-09-02T13:00:00Z" ;

:time_coverage_start = "2011-06-16T23:00:00Z" ;

:time_deployment_end = "2011-08-31T13:30:00Z" ;

:time_deployment_end_origin = "TimeLastInPos" ;

:time_deployment_start = "2011-06-20T12:55:00Z" ;

:time_deployment_start_origin = "TimeFirstInPos" ;

:toolbox_input_file = "/home/holiveira/volumes/gg_hd/GG_stuff/IMOS_toolbox/data_files_examples/NSW/DATA/MOORINGS/RAW/PH100/019_PH100_20Jun2011/AQUAtec/csv/PH100_520PT_20Jun2011_564.csv" ;

:toolbox_version = "2.6.7 - GLNXA64" ;

}}

APPENDIX 2: IMOS TOOLBOX OPTIONAL GLOBAL ATTRIBUTES

The IMOS-Toolbox has been developed by the Australian National Mooring Network, supported by IMOS and the eMarine Information Infrastructure. It is written in MATLAB and Java with a graphical user interface and aims at converting oceanographic data files into pre-processed and quality controlled netCDF files. This IMOS toolbox is freely available in a standalone executable with its source code. More documentation and downloads can be found on the imos-toolbox GitHub page. When correctly configured, this toolbox automatically documents these optional global attributes:

Non-CF attribute	Name	Type	Example	Definition
	What			
	toolbox_input_file	S	toolbox_input_file = "E:\Documents and Settings\ggalibert\My Documents\IMOS_toolbox\data_files_examples\NSW\DATA\MOORINGS\SYD100\RAW\021_SYD100_15Dec2010\WQM\ASCII\WQM0141_031.DAT"	Full local path and filename of the original data file processed by the toolbox.
	toolbox_version	S	toolbox_version = "2.1b"	Version of the toolbox used to process data.

deployment_code	S	deployment_code = “NRSMAI-0807”	Deployment unique code within IMOS project, usually composed of a site_code and a date information in format YYMM or YYMMDD. Its content can be adapted to what suits best any particular needs of a facility. A deployment refers to a specific time period over which a specific platform (e.g. mooring) is continuously collecting measurements at a specific location.
instrument_sample_interval	N	instrument_sample_interval = 1	Sampling interval in seconds performed by the instrument on this data set.
instrument_beam_angle	N	instrument_beam_angle = 25	Angle between a transducer beam’s main axis and the vertical axis of the ADCP.
instrument_burst_interval	N	instrument_burst_interval = 900	Interval in seconds between two bursts of data collection performed by the instrument on this data set. Note that not all instruments are sampling data in bursts.
instrument_burst_duration	N	instrument_burst_duration = 120	Duration in seconds of a burst. Note that not all instruments are sampling data in bursts.
cruise	S	cruise = “125456”	Reference to the cruise during which the dataset was collected.
station	S	station = “1”	Reference to the station where the dataset (cast) was collected.

Non-CF attribute

Where

instrument_nominal_depth	N	instrument_nominal_depth = 90	Instrument nominal depth below sea surface in metres.
instrument_nominal_height	N	instrument nominal_height = 0	Instrument nominal depth above sea floor in metres.
site_nominal_depth	N	site_nominal_depth = 90	Site nominal depth below sea surface in metres.
site_depth_at_deployment	N	site_depth_at_deployment = 84	Measured depth on site at deployment in metres.
Non-CF attribute	**When**		
time_deployment_start	S	time_deployment_start = “2008-11-23T10:16:00Z”	Start date in UTC when the instrument is deployed in its nominal position. See section 3.2.4 - on time format.
time_deployment_start_origin	S	time_deployment_start_origin = "TimeFirstInPos"	Origin of time_deployment_start. Can be : TimeFirstGoodData TimeFirstInPos TimeFirstWet TimeSwitchOn
time_deployment_end	S	time_deployment_end = “2009-01-06T09:43:00Z”	Final date in UTC when the instrument is retrieved from its nominal position. See section 3.2.4 - on time format.

time_deployment_end_origin	S	time_deployment_end_origin = "TimeLastInPos"	Origin of time_deployment_end. Can be : TimeLastGoodData TimeLastInPos TimeOnDeck TimeSwitchOff

APPENDIX 3: IMOS FILE NAMING CONVENTION

IMOS NETCDF FILE NAMING CONVENTION