AUSTRALIAN NATIONAL FACILITY FOR OCEAN GLIDERS (ANFOG)

Data Management

USER's MANUAL
Version 5.01
March 2017

Version	Date	Comments	Author
0.9	28/12/2007	Creation of document	Mun Woo
1.0	16/07/2008	Changed fill values from 99999 to NaN	Mun Woo
1.1	11/09/2008	Changed units of DOXY to mg/L Corrected POSITION_QC dimensions to NSURF	Mun Woo
1.2	8/10/2008	Added 6 QC tests Changed oxygen probe temperature range = 0 to 40 ^{o}C Changed pitch range = -90 to 90 deg	Mun Woo
1.3	31/10/2008	Added data transfer flow diagrams (figures 1 & 2). Included Appendix showing uncertainties for values taken from sensors. Amended Spike test & Gradient test to include considerations for data deeper than 500 db. Added sawtooth profile figure (Fig. 3).	Mun Woo
1.4	07/01/2009	Added sensor list for Seaglider (Ref Table 3.5)	Mun Woo
1.5	20/01/2009	Incorporated Seaglider data formats	Mun Woo
2.0	21/02/2009	Overhaul of document to conform to IMOS NETCDF USER'S MANUAL Version 1.0	Mun Woo
2.1	11/05/2009	Changes made to conform to IMOS NETCDF USER'S MANUAL Version 1.1, and IMOS FILE NAMING CONVENTION Version 1.2.	Mun Woo
2.2	21/05/2009	Changes made to conform to IMOS NETCDF USER'S MANUAL Version 1.2, and IMOS FILE NAMING CONVENTION Version 1.3.	Mun Woo
2.3	07/08/2009	Addition of DOXY time-lag correction, and corresponding parameter DOXY_raw to store original uncorrected values.	Mun Woo
3.0	12/04/2010	Inclusion of additional parameters HEAD, UCUR & VCUR derived from engineering parameters. Processed file names now include 'E' and 'V' to reflect inclusion of engineering parameter & seawater current velocities.	Mun Woo

3.01	01/08/2012	Minor modifications in the QC process Addition of Irradiance parameters (OCR: Ocean Colour Radiometer)	Claire Gourcuff & Christine Hanson
3.1	01/10/2012	New processing code for interpolation of raw engineering data. Addition of parameter PRES. Additional QC on OCR parameters data. DOXY unit correction. Table 1 updated. New values for QC range tests (CDOM, DOXY, DEPTH, TEMP, FLU2). Spike test on FLU2, CDOM and VBSC removed. Flag FLU2, CDOM, VBSC and OCR data < 0.5 m depth. Flag descending OCR profiles. Format change to the document: previous section 3. expanded and split into sections 3. and 4.	Claire Gourcuff & Christine Hanson
4.0	08/07/2014	Major changes - Generation of FV00 NetCDF files, which contain raw data on their own time dimension, following most of the EGO (Everyone's Gliding Observatories) recommendations. - Changes in parameters: FLU2 replaced by CPHL, DOXY removed, DOX1 and DOX2 added. Irradiance parameter names changed. - New parameters PHASE and PROFILE. - metadata regarding PLATFORM, DEPLOYMENT, and SENSOR stored as attributes of empty variables (containers). - Data corrections implemented for Slocums in FV01 files. - Seawater properties computed using Gibbs TEOS-10 library instead of EOS-80. - Minor changes in the global and variable attributes. - Section on uncertainties removed (work has to be done on the subject). - NetCDF 4 format	Claire Gourcuff
5.0	11/11/2016	Introduced a plethora of changes for compliancy with the IMOS 1.4 convention which comprise many changes to global and variable attributes. Additionally, a new variable, BBP has been introduced. The status of the fleet of ANFOG gliders in operation has also been updated.	Mun Woo
5.01	30/3/2017	Data taken during any period outside of the start and end time of a glider mission are now removed, rather	Mun Woo

		than flagged QC=0. Good data taken at bad or missing location fixes are now flagged QC=3 rather than 0.	

Table of Contents

1 Introduction

This document is the Australian National Facility for Ocean Gliders (ANFOG) data user's manual. ANFOG is a facility under Australia's Integrated Marine Observing System (IMOS). The document contains the description of the formats and files produced by ANFOG.

1.1 ANFOG program

The underwater ocean glider represents a technological revolution for oceanography. Autonomous ocean gliders can be built relatively cheaply, are controlled remotely and reusable allowing them to make repeated subsurface ocean observations at a fraction of the cost of conventional methods. ANFOG, with IMOS/NCRIS/EIF funding, currently deploys a fleet of 13 gliders around Australia. The data retrieved from the glider fleet will contribute to the study of the major boundary current systems surrounding Australia and their links to coastal ecosystem processes.

The ANFOG glider fleet consists of two types: Slocum gliders and Seagliders. Slocum gliders (named for Joshua Slocum the first solo global circumnavigator), manufactured by Teledyne Webb Research Corp, are optimised for shallow coastal waters (< 200 m) where high manoeuvrability is needed. ANFOG currently has 7 Slocum gliders for deployment on the continental shelf. Seagliders, originally designed and built at the University of Washington then by iRobot and now supplied by Kongsberg, are designed to operate most efficiently in the open ocean up to 1000 m water depth. ANFOG uses Seagliders to monitor the boundary currents surrounding Australia. The Seagliders are used to conduct repeated glider surveys across the boundary currents and continental shelves, which is valuable for gathering long-term environmental records of physical, chemical and biological data not widely measured to date. Whilst the Slocum gliders, due to their low cost and operational flexibility, will be of great use in intensive coastal monitoring, both types of gliders weigh only 52kg, enabling them to be launched from small boats. They have a suite of sensors able to record temperature, salinity, dissolved oxygen, turbidity, dissolved organic matter, fluorescence (proxy for chlorophyll) and light penetration against position and depth.

ANFOG data are provided to users via the IMOS portal (http://imos.aodn.org.au/webportal/), both in Near Real Time and in Delayed mode (Figure 1).

The ANFOG data formats are based on NetCDF because:

- It is a widely accepted data format by the user community,
- It is a self-describing format for which tools are widely available,
- It is a reliable and efficient format for data exchange.

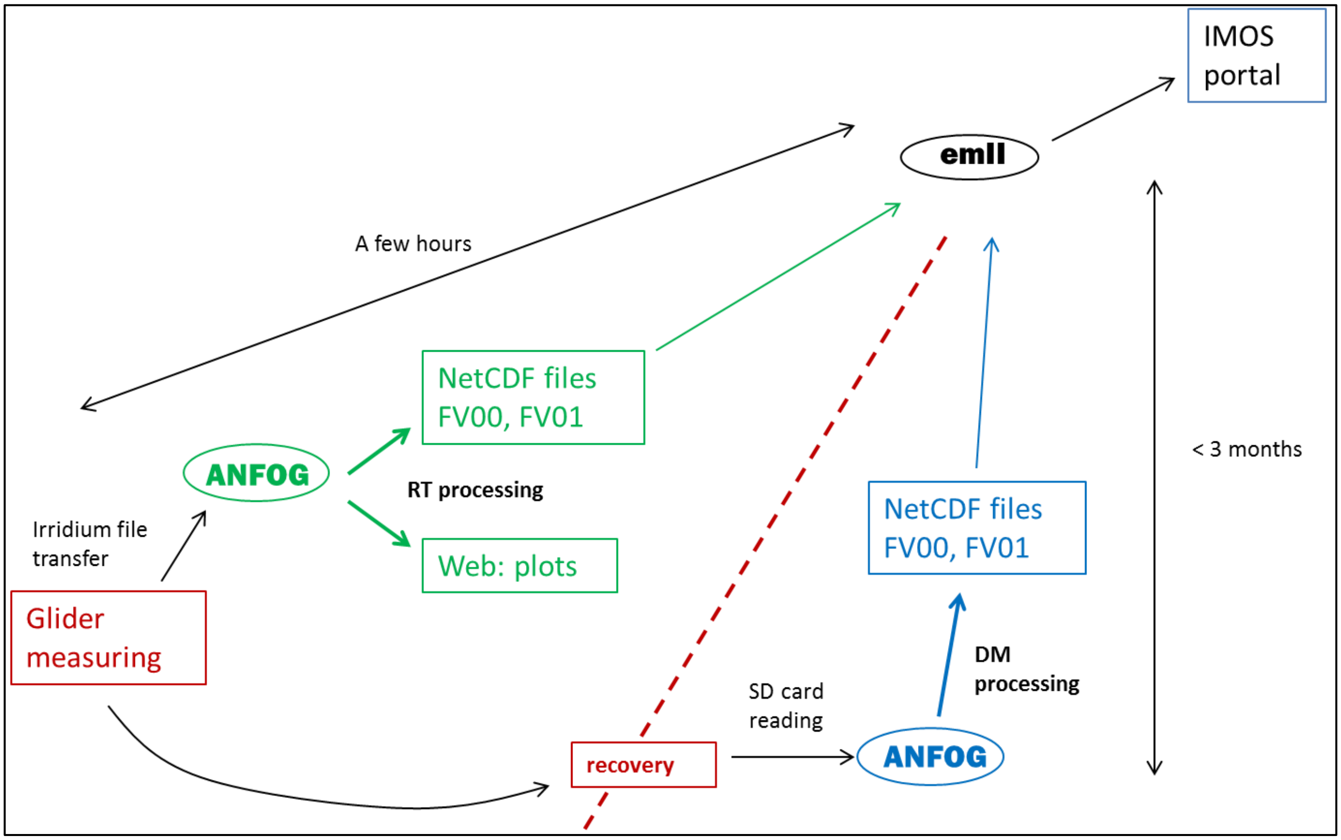

Figure 1. ANFOG data processing summary. emII stands for eMarine Information Infrastructure (an IMOS facility), RT for 'Real Time' and DM for 'Delayed Mode'.

1.2 Slocum glider data

Slocum gliders have an endurance of ~20 days and range of 1,500 km. Typically, the Slocum glider first records its position and the time stamp at the surface and then performs multiple dive-ascent cycles continuously in a saw-tooth pattern through the water column (Figure 2). After that, it surfaces again to transmit a sample of data, receive new instructions from the base station (if any), and finally it records its position and time stamp before diving into the next segment of saw-tooth sampling. ANFOG gliders are usually configured to sample up to 200 m depth.

Currently, all ANFOG Slocum gliders are instrumented with a Seabird-CTD, WETLabs FLBBCD 3 parameters optical sensor (measuring Chlorophyll-a, CDOM and Backscatter), an Aanderaa Oxygen Optode and a Satlantic 4-channel OCR (Ocean Colour Radiometer) downwelling irradiance sensor. Note that Slocum gliders of G1 type, used in the past did not have any irradiance sensor.

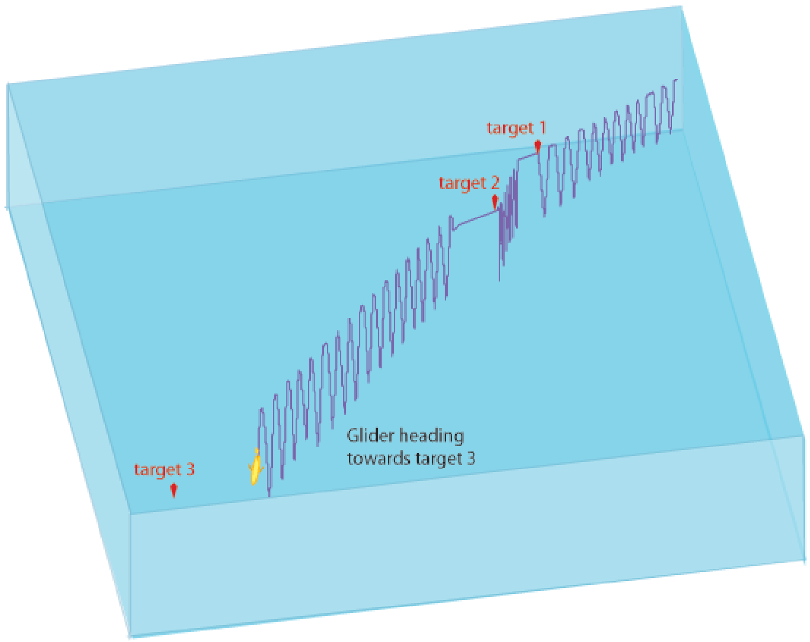

Figure 2. A Slocum glider performing 3 segments of saw-tooth lines. It surfaces at target points. This is different from a Seaglider, which usually surfaces after every descent-ascent cycle.

1.3 Seaglider data

Seagliders have an endurance of 1-6 months and range of 6,000 km and can sample up to 1000 m depth. Typically, the Seaglider surfaces after every dive-ascent cycle, takes GPS fixes and sends data. On some occasions, pilots may set the Seaglider to surface after a pre-determined number of dives have been completed in order to navigate through strong surface currents. Unlike Slocum gliders for which only a subset of the dataset is transmitted while the glider is in the water (and the full dataset will only be available for processing upon vehicle retrieval), Seagliders always send all their data at each surfacing.

ANFOG Seagliders are equipped with a Seabird-CTD, WETLabs BBFL2VMT 3 parameters optical sensor (measuring Chlorophyll-a, CDOM and Backscatter) and a Seabird Oxygen sensor.

2 Glider data and meta-data format

Glider format files contain meta-data and data for one deployment of one glider. The observation data is stored as data timeseries in NetCDF format. Data are transmitted from the gliders, passed through processing and quality control procedures before being available to users through the IMOS portal. ANFOG data processing is summarized in Figure 1.

ANFOG produces data in Near Real Time, i.e. immediately after the glider begins reporting at surface, and data in Delayed Mode, after the glider has been recovered. Since the processing version 4.0, ANFOG generates NetCDF-4 format files (versus NetCDF-3 format previously), and releases 2 types of NetCDF files for each deployment: FV00 and FV01. These two files contain the same global attributes. The FV00 file contains all the raw data from the deployment (both in engineering and physical units), with each parameter following its own time dimension. This file is aimed at storing raw data in a self-defined way, and to follow as close as possible the EGO (Everyone's Gliding Observatories) format. Data in FV00 files are not Quality Controlled. The FV01 file is the one recommended to users, with all the parameters associated with Quality Controlled flags. It includes corrected data and all the parameters interpolated (slightly, as all the time dimensions are usually very similar) following the same unique time dimension. For clarity reasons, only the FV01 file format is described extensively in this document.

2.1 File naming convention

NetCDF files from ANFOG follow the following IMOS convention (also described in Mancini, Tattersall, Proctor, and Galibert (2012)), described in Table 1:

IMOS_ANFOG_<data-code>_<start-date> _ <platform-code> _FV <file-version> _<product-type>_END-<end-date>.nc

Examples:

IMOS_ANFOG_BCEOSTUV_20090210T044813Z_SG155_FV01_timeseries_END-20090210T050626Z.nc

IMOS_ANFOG_ER_20090210T044813Z_SG155_FV00_timeseries_END-20090210T050626Z.nc

Table 1. Elements of file-naming convention

Part of filename	Description
data-code	B: biology (plankton, fluorescence, nutrients, dissolved organic matter) C: conductivity O: oxygen concentration S: salinity of sea water T: temperature of sea water U:turbudity of sea water (including backscatter) V:velocity of sea water E: Engineering or technical parameters.

	R: Raw data.
start-time	Start date and time of the measurements in UTC. Date format is: yyyymmddTHHMMSSZ where T is the delimiter between date and time, and Z indicates that time is in UTC. Example: 20090211T114713Z is 11th February 2009, 11:47:13 AM
platform-code	The glider unit's name. See Table 2.
file-version	Value representing the version of the file. This value is preceded by two characters: 'FV'. 00 : 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 units or physical units. Time and location details can be in relative units and values can be pre-calibration measurements. Level 0 data is unsuitable for public access. 01 : Level 1 – quality controlled data. Quality controlled data have passed quality assurance procedures such as routine estimation of timing and sensor calibration or visual inspection and removal of obvious errors. The data are in physical units using standard SI metric units, with calibration and other pre-processing and corrections applied. All time and location values are in absolute coordinates to agreed standards and datum. Metadata exists for the data. This is the standard data level that is made available to the IMOS community.
product-type	(optional) This code gives information about the product included in the dataset. Example: timeseries
end-date	End date and time of the measurements in UTC. The code is preceded by four characters: 'END-'. Date format is: yyyymmddTHHMMSSZ where T is the delimiter between date and time, and Z indicates that time is in UTC. Example: 20090211T114713Z is 11th February 2009, 11:47:13 AM
creation-date	(optional) Creation date and time of the file. The data format is the same as the start and end date. The code is preceded by the two characters C-. In the case of ANFOG datasets, files that include a creation date have been re-processed with an updated version of the processing software. Example: C-20081112T231255Z

Listed in Table 2 are codes by which each glider is identified. Each glider unit's name is unique within the IMOS project, and used as its platform code. ANFOG's Seaglider codes begin with 'SG', while the Slocum codes begin with 'SL'.

Table 2. ANFOG platform codes.

Glider type	Platform code	Platform type	Active as of November 2016
Seagliders	*SG151*	*Seaglider UW*	*No (lost in 2010)*
	SG152	**Seaglider UW**	**Yes**
	SG153	**Seaglider UW**	**Yes**
	SG154	**Seaglider UW**	**Yes**
	SG155	**Seaglider UW**	**Yes**
	SG514	*Seaglider 1kA (iRobot)*	*No (lost in 2014)*
	SG516	*Seaglider 1kA (iRobot)*	*No (lost in 2016)*
	SG517	*Seaglider 1kA (iRobot)*	*No (lost in 2011)*
	SG519	*Seaglider 1kA (iRobot)*	*No (lost in 2011)*
	SG520	*Seaglider 1kA (iRobot)*	*No (lost in 2011)*
	SG521	*Seaglider 1kA (iRobot)*	*No (lost in 2013)*
	SG540	*Seaglider 1kA (iRobot)*	*No (lost in 2014)*
Slocum	*SL104*	*Slocum G1*	*No (lost in 2010)*
	SL106	*Slocum G1*	*No (lost in 2010)*
	SL109	*Slocum G1*	*No (lost in 2009)*
	SL130	*Slocum G1*	*No (lost in 2011)*
	SL209	**Slocum G2**	**Yes**
	SL210	**Slocum G2**	**Yes**
	SL239	**Slocum G2**	**Yes**
	SL248	**Slocum G2**	**Yes**
	SL281	**Slocum G2**	**Yes**
	SL286	**Slocum G2**	**Yes**
	SL287	**Slocum G2**	**Yes**
	SL502	**Slocum G2**	**Yes**

2.2 Global attributes

This section contains general information about the NetCDF files themselves (valid for FV01 files).

Table 3. ANFOG NetCDF files global attributes

Name	Example	Definition
project	char('Integrated Marine Observing System (IMOS)');	The scientific project that produced the data

institution	char('ANFOG (Australian National Facility for Ocean Gliders)');	Name of the institute or facility where the original data was produced.
references	char(' http://www.imos.org.au'; "ANFOG Data Management Users Manual v4 July 2014");	Published or web-based references describing the data or methods used to produce it.
naming_authority	char('IMOS');	Naming authority will always be IMOS.
cdm_data_type	char('Trajectory')	Unidata Common Data Model data type used by THREDDS. More information on http://www.unidata.ucar.edu/projects/THREDDS/CDM/CDMTDS.htm
data_center	char(' Australian Ocean Data Network (AODN)')	Data center in charge of management and distribution of the data resource.
data_centre_email	char('info@aodn.org.au')	Data centre contact email address.
principal_investigator	char('Pattiaratchi, Charitha')	Name of principal investigator in charge of the glider unit.
principal_investigator_email	char('chari.pattiaratchi@uwa.edu.au')	Principal investigator's email address.
institution_references	char('http://imos.org.au/emii.html; http://imos.org.au/anfog.html')	Reference to the data provider and producer.
citation	char('Citation to be used in publications should follow the format: "IMOS.[year-of-data-download],[Title],[Data access URL],accessed [date-of access]".')	Citation used for usage of this data.
acknowledgement	char('Any users of IMOS data are required to clearly acknowledge the source of the material in the format: "Data was sourced from the Integrated Marine Observing	Any users (including re-packers) of IMOS data are required to acknowledge the source of the data in this format.

	System (IMOS) - IMOS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS) and the Super Science Initiative (SSI).”’)	
disclaimer	char(' Data, products and services from IMOS are provided "as is" without any warranty as to fitness for a particular purpose.')	Statement describing data distribution policy.
license	char('http://creativecommons.org/licenses/by/4.0/')	Reference for the license for the data
standard_name_vocabulary	Char('NetCDF Climate and Forecast (CF) Metadata Convention Standard Name Table Version 29')	Reference for CF standard names
source	char('Glider observation');	Method of production of the original data.
geospatial_lat_units	char('Degrees_north')	Units used for geospatial_lat_min/max attributes.
geospatial_lon_units	char('Degrees_east')	Units used for geospatial_lon_min/max attributes.
geospatial_vertical_units	char('meter')	Units used for geospatial_vertical_min/max attributes.
geospatial_vertical_positive	char('down')	Direction of vertical coordinates
Conventions	char(‘CF-1.6,IMOS-1.4’);	Format convention used by the dataset
title	char([‘Slocum G1 glider data from the mission SpencerGulf20080716’]);	Short description of the dataset indicating the type of glider from which data has come from, and the corresponding name of the mission.
date_created	char(‘2014-06-11T09:16:22z’);	Date and time at which the data was created.

		Format: yyyy-mm-ddTHH:MM:SSz' Example: 2008-12-10T09:35:36z : December 10th 2008 9:35:36AM
abstract	char(' This NetCDF file contains data from the SpencerGulf20080716 glider deployment. The platform deployed was UNIT109 (Slocum G1) with the following instruments on board: SBE CTD CTD41CP (Conductivity, Temperature, Depth), Wetlabs BBFL2S 3 parameters optical sensor (measuring Chlorophyll-a, CDOM and 660nm Backscatter)');	A paragraph describing the dataset: type of data contained, how it was created, who collected it, what instruments were used, etc.
comment	char(' This file contains data from the following specific instruments: SBE CTD CTD41CP, SN 106, cal 9/09/2007; Wetlabs bio-optical sensor BBFL2S, SN 498, cal 5/02/2008; ');	Additional information about the data, methods used to produce it, instruments used (model, serial number and calibration date), and correction applied to the data.
keywords	char('Water Temperature, Conductivity, Salinity, Oxygen, Turbidity, Chlorophyll, Organic Matter, Fluorescence, Scattering, Downwelling spectral irradiance');	A comma separated list of key words and phrases.
geospatial_lat_min	double(-35.5568);	Southernmost latitude (positive north) from which dataset was obtained; a value between -90 and 90 degrees.
geospatial_lat_max	double(-35.4248);	Northernmost latitude (positive north) from which dataset was obtained; a value between -90 and 90 degrees.

geospatial_lon_min	double(136.3502);	Westernmost longitude (positive east) from which dataset was obtained; a value between -180 and 180 degrees.
geospatial_lon_max	double(136.9128);	Easternmost longitude (positive east) from which dataset was obtained; a value between -180 and 180 degrees.
geospatial_vertical_min	double(-1.5731);	Minimum depth of measurements, in metres.
geospatial_vertical_max	double(100.559);	Maximum depth of measurements, in metres.
time_coverage_start	char('2008-07-16T00:47:11Z')	Start date and time in UTC, for which data in the dataset was collected. Format: yyyy-mm-ddTHH:MM:SSZ' Example: 2008-12-10T09:35:36Z : December 10th 2008 9:35:36AM
time_coverage_end	char('2008-07-24T02:37:22Z')	End date and time in UTC, for which data in the dataset was collected. Format: yyyy-mm-ddTHH:MM:SSZ' Example: 2008-12-10T09:35:36Z : December 10th 2008 9:35:36AM
author	char('Gourcuff, Claire')	Name of person responsible for the creation of the dataset.
processing	char('Written by MATLAB processing code version 5.00 - processing_ID=2216')	Version of ANFOG processing used to generate the file, and processing ID number (useful internally for ANFOG in case of reprocessing).
deployment_code	char('SpencerGulf20080716')	IMOS mission name
netcdf_version	char('4.0')	NetCDF file version
platform_code	char('SL109');	ANFOG glider unit number (as listed in Table 2). These are platform codes unique within the IMOS project.
local_time_zone	double(10)	Local time zone (UTC+)
author_email	char('anfog-ecm@uwa.edu.au')	NetCDF file author contact email address.

Data_mode	Char('Delayed Mode')	Mode of data processing

2.3 Dimensions

ANFOG Glider data are collected as time-series, where three-dimensional coordinates of latitude, longitude and depth coordinates are recorded, along with various measured parameters. A few computed parameters included in the files allow users to easily identify up and down profiles within the time-series. The length of the time-series data varies according to the duration spent underwater and frequency of sampling, with deeper dives typically producing longer data time-series. FV01 files only include one dimension: TIME (Table 4). FV00 files can contain up to four time dimensions among the ones described in Table 5, with each parameter associated with one of the available dimensions (see EGO (2013)).

Table 4. ANFOG FV01 NetCDF files dimension

Dimension	Definition
TIME	Number of time steps over which data was sampled: corresponds to the Science Bay time stamp for Seagliders, and to the CTD time stamp if recorded or to the Science Bay time stamp if not for Slocums.

Table 5. ANFOG FV00 NetCDF files dimensions

Possible Dimension	Definition
EVENT_SCIENCE	Number of time steps over which some of the data was sampled. Always present, both for Seagliders and Slocums.
EVENT_NAVIGATION	Number of time steps over which some of the data was sampled. Always present, both for Seagliders and Slocums (main Bay)
EVENT_GPS	Number of time steps over which Seaglider GPS data was sampled. Always present for Seagliders, never for Slocums.
EVENT_CTD	Number of time steps over which the Slocum CTD data was sampled. Never present for Seagliders, present for Slocums if the glider was configured to record the CTD time stamp.
EVENT_OXY	Number of time steps over which the Slocum Optode data was sampled. Never present for Seagliders, present for Slocums if the glider was configured to record the Optode time stamp.
EVENT_WETLABS	Number of time steps over which the Slocum ECO-triplet data was sampled. Never present for Seagliders, present for Slocums if the glider was configured to record the ECO-triplet time stamp.

2.4 Variables

This section contains information about the various types of data recorded during data sampling. Table 6 lists ANFOG variables with their attributes. Please refer to Mancini et al. (2012) for further details.

2.4.1 Variables and attributes in FV01 NetCDF datafiles

Table 6. Variable attributes in FV01 NetCDF files

Variable	Attributes	Definition
TIME	double TIME (TIME); TIME.standard_name = 'time'; TIME.long_name = 'time from the CTD'; TIME.units = 'days since 1950-01-01 00:00:00 UTC'; TIME.axis = 'T'; TIME.valid_min = 0; TIME.valid_max = 90000.0; TIME.calendar = 'gregorian'; TIME.ancillary_variables = 'TIME_quality_control'; TIME.observation_type = 'measured'; TIME_quality_control_set = 1; TIME.quality_control_indicator = 1;	Time at which <PARAM> measurements were made. Values are recorded as days since 12 am of 1st January 1950. The quality_control_set is as listed in Table 8 (usually '1', i.e. IMOS standard using IODE flags) The quality_control_indicator values are as listed in 9.
LATITUDE	double LATITUDE(TIME); LATITUDE.standard_name = 'latitude'; LATITUDE.long_name = 'latitude'; LATITUDE.units = 'degrees_north'; LATITUDE.axis = 'Y'; LATITUDE.valid_min = -90; LATITUDE.valid_max = 90; LATITUDE._FillValue = 99999.0; LATITUDE.comment = 'obtained from GPS fixes'; LATITUDE.reference_datum = 'geographical coordinates, WGS84 projection'; LATITUDE.ancillary_variables = 'LATITUDE_quality_control'; LATITUDE.observation_type = 'measured'; LATITUDE.quality_control_set = 1; LATITUDE.quality_control_indicator = 1;	Estimates of latitudinal position based on the position at the surface. The quality_control_set is as listed in Table 8 (usually '1', i.e. IMOS standard using IODE flags) The quality_control_indicator values are as listed in 9.
LONGITUDE	float LONGITUDE(TIME); LONGITUDE.standard_name = 'longitude'; LONGITUDE.long_name = 'longitude'; LONGITUDE.units = 'degrees_east';	Estimates of longitudinal position based on the position at the surface.

	LONGITUDE.axis = 'X'; LONGITUDE.valid_min = -180; LONGITUDE.valid_max = 180; LONGITUDE._FillValue = 99999.0; LONGITUDE.comment = 'obtained from GPS fixes'; LONGITUDE.reference_datum = 'geographical coordinates, WGS84 projection'; LONGITUDE.ancillary_variables = 'LATITUDE_quality_control'; LONGITUDE.observation_type = 'measured'; LONGITUDE.quality_control_set = 1; LONGITUDE.quality_control_indicator = 1;	The quality_control_set is as listed in Table 8 (usually '1', i.e. IMOS standard using IODE flags) The quality_control_indicator values are as listed in 9.
PRES	float PRES(TIME); PRES.standard_name = 'sea_water_pressure'; PRES.long_name = 'sea_water_pressure'; PRES.units = 'dbar'; PRES._FillValue = 99999.0; PRES.valid_min = -5; PRES.valid_max = 1100; PRES.coordinates = 'TIME LATITUDE LONGITUDE DEPTH'; PRES.comment = 'pressure measured by the CTD'; PRES.ancillary_variables = 'PRES_quality_control'; PRES.observation_type = 'measured'; PRES.quality_control_set = 1; PRES.quality_control_indicator =1;	Pressure values are recorded in decibars downward from the sea surface (when available, the pressure recorded is the one measured by the CTD). The quality_control_set is as listed in Table 8 (usually '1', i.e. IMOS standard using IODE flags) The quality_control_indicator values are as listed in 9.
DEPTH	float DEPTH(TIME); DEPTH.standard_name = 'depth; DEPTH.long_name = 'depth'; DEPTH.units = 'm'; DEPTH._FillValue = 99999.0; DEPTH.valid_min = -5; DEPTH.valid_max = 1000; DEPTH.positive = 'down'; DEPTH.axis = 'Z'; DEPTH.reference_datum = 'sea surface'; DEPTH.ancillary_variables = 'DEPTH_quality_control'; DEPTH.observation_type = 'computed'; DEPTH.quality_control_set = 1;	Depths computed from latitude and relative pressure measurements. The quality_control_set is as listed in Table 8 (usually '1', i.e. IMOS standard using IODE flags) The quality_control_indicat

	DEPTH.quality_control_indicator = 1; DEPTH.comment = 'Depth computed using the Gibbs-SeaWater toolbox (TEOS-10) v3.02, from latitude and relative pressure measurements';	or values are as listed in 9.
<PARAM>	float <PARAM>(TIME); <PARAM>.standard_name[1] = '<X>'; <PARAM>.long_name = '<X>'; <PARAM>.units = '<X>'; <PARAM>._FillValue = 99999.0; <PARAM>.valid_min = <X>; <PARAM>.valid_max = <X>; <PARAM>.coordinates = <X>; <PARAM>.ancillary_variables = '<X>'; <PARAM>.observation_type = 'measured'; <PARAM>.quality_control_set = 1; <PARAM>.quality_control_indicator = 1; <PARAM>.comment = ' ';	<PARAM> contains the values of a parameter listed in reference table 7. <X> : this field is specified in the reference table 7. The observation_type is either 'measured' if the parameter directly comes from the sensor, or 'computed' if not. The quality_control_set is as listed in Table 8 (usually '1', i.e. IMOS standard using IODE flags) The quality_control_indicator values are as listed in 9.
<PARAM_quality_control>	double <PARAM_quality_control> (TIME); <PARAM_quality_control> .standard_name[1] = ' <PARAM>.standard_name status_flag'; <PARAM_quality_control> .long_name = 'quality flag for <PARAM>.standard_name'; <PARAM_quality_control>.quality_control_set = 1; <PARAM_quality_control>.quality_control_conventions = 'IMOS standard flags'; <PARAM_quality_control>._FillValue = 99; <PARAM_quality_control>.valid_min = 0; <PARAM_quality_control>.valid_max = 9;	Quality flag applied on the <PARAM> values. Information on flag meanings is found in 9.

[1] *<PARAM>.standard_ name* is not used if the parameter does not have a CF standard name.

	<PARAM_quality_control>.flag_values = [0,1,2,3,4,5,6,7,8,9]; <PARAM_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 interpolated_values missing_values';	

2.4.2 Parameters

The parameter codes used for ANFOG data management (FV01 files) other than LONGITUDE, LATITUDE, PRES, DEPTH and TIME are defined in Table 7. Some of these parameters are directly measured by the sensors, whereas some others are calculated from raw parameters.

Table 7. Parameters measured as timeseries by sensors on ANFOG gliders

code	standard_name / long_name (for non-CF)	observation _type	_FillValue	valid _min	valid _max	coordinates	Ancillary _variables	units
CNDC	sea_water_electrical_conductivity	measured	99999	0	60	TIME LATITUDE LONGITUDE DEPTH	CNDC_quality_control	S m^{-1}
TEMP	sea_water_temperature	measured	99999	-2.5	40	TIME LATITUDE LONGITUDE DEPTH	TEMP_quality_control	Celsius
PRES	sea_water_pressure	measured	99999	-5	1100	TIME LATITUDE LONGITUDE DEPTH	PRES_quality_control	dbar
BBP	particle_backscattering_coefficient	computed	99999	0	1	TIME LATITUDE LONGITUDE DEPTH	BBP_quality_control	m^{-1}
PSAL	sea_water_salinity	computed	99999	2	41	TIME LATITUDE LONGITUDE DEPTH	PSAL_quality_control	$1e^{-3}$
DOX1	mole_concentration_of_dissolved_molecular_oxygen_in_sea_water	measured (SG) computed (SL)	99999	0	650	TIME LATITUDE LONGITUDE DEPTH	DOX1_quality_control	umol l^{-1}
DOX2	moles_of_oxygen_per_unit_mass_in_sea_water	computed	99999	0	650	TIME LATITUDE LONGITUDE DEPTH	DOX2_quality_control	umol l^{-1}

CPHL	mass_concentration_of_chlorophyll_in_sea_water	computed	99999	0	100	TIME LATITUDE LONGITUDE DEPTH	CPHL_quality_control	mg m^{-3}
╪CDOM	concentration_of_coloured_dissolved_organic_matter	computed	99999	0	400	TIME LATITUDE LONGITUDE DEPTH	CDOM_quality_control	$1e^{-9}$
╪VBSC	volumetric_scattering_function	computed	99999	0	0.1	TIME LATITUDE LONGITUDE DEPTH	VBSC_quality_control	m^{-1} sr^{-1}
╪IRRADxxx*	downwelling_spectral_irradiance_in_sea_water_beam**n**+	computed	99999	0	1000	TIME LATITUDE LONGITUDE DEPTH	IRRADxxx*_quality_control	uW cm^{-2} nm^{-1}
╪HEAD	vehicle_heading	measured	99999	0	360	TIME LATITUDE LONGITUDE	HEAD_quality_control	Degrees
UCUR	eastward_sea_water_velocity	computed	99999	-10	10	TIME LATITUDE LONGITUDE	UCUR_quality_control	m s^{-1}
VCUR	northward_sea_water_velocity	computed	99999	-10	10	TIME LATITUDE LONGITUDE	VCUR_quality_control	m s^{-1}
╪UCUR_GPS	eastward_surface_sea_water_velocity	computed	99999	-10	10	TIME LATITUDE LONGITUDE	UCUR_GPS_quality_control	m s^{-1}
╪VCUR_GPS	northward_surface_sea_water_velocity	computed	99999	-10	10	TIME LATITUDE LONGITUDE	VCUR_GPS_quality_control	m s^{-1}
╪PHASE	glider_trajectory_phase_code	computed	99	0	6	TIME LATITUDE LONGITUDE	PHASE_quality_control PROFILE	1
╪PROFILE	glider_trajectory_profile_number	computed	999999	0	900000	TIME LATITUDE LONGITUDE	PROFILE_quality_control PHASE	1

╪ indicates a non-CF parameter. These parameters do not have associated standard names. Instead, *long_name* is listed.

* the codes are wave length dependant, different on each sensor. In total we have 7 generic wavelengths, i.e. ANFOG data contain the following parameters: IRRAD443, IRRAD490, IRRAD555, IRRAD670, IRRAD456, IRRAD 470 and IRRAD520. For instance for platform SL210, the four irradiance codes are: IRRAD443, IRRAD490, IRRAD555 and IRRAD670. The real wavelength, that is sometimes slightly different from the generic one (and differs from one sensor SN to another), is specified in the 'comment' attribute of the corresponding parameter.

+ **n** indicates the beam number, e.g. 1, 2, 3 or 4.

Description of calculated parameters:

- **DEPTH**: the depth is computed from measured pressure and latitude (linearly interpolated between GPS fixes), using Gibbs TEOS-10 equations.
- **PSAL**: the salinity is computed from measured conductivity, pressure and temperature, using Gibbs TEOS-10 equations.
- **DOX1**: for Slocum gliders, when the glider was configured to output the optode phases measurements, DOX1 is recomputed using these phase measurements, calibration coefficients, and the temperature measured by the CTD. Please refer to Gourcuff (2014b) for more details.
- **DOX2**: the oxygen concentration per unit mass is computed as DOX1 divided by the potential density ρ, where ρ is computed from temperature and salinity, using Gibbs TEOS-10 equations.
- **CPHL, CDOM, VBSC & IRRADxxx**: these parameters are computed using the raw measurements, in counts ($PARAM_{counts}$, stored in the FV00 files), together with calibration coefficients, following the equation:

  $$PARAM_{units} = SF \times (PARAM_{counts} - DC)$$

  where SF is the scale factor coefficient and DC the dark count coefficient, available after a new sensor calibration.
- **UCUR, VCUR:** UCUR and VCUR are depth-mean velocities of the seawater over all the water that the glider travels through between surfacing. The values are approximate estimates derived from engineering parameters.
- **UCUR_GPS, VCUR_GPS:** these velocities are estimates of the surface current, computed between two GPS fixes while the glider is drifting at surface. The values are approximate estimates derived from engineering parameters.
- **PHASE:** the phase of the trajectory at a given time is defined following EGO (see table 9 in EGO (2013)). Values used: 0 for surface drifting, 1 for descending profile, 4 for ascending profile, 3 for inflexion. Phase values are defined using a script based on depth rate.
- **PROFILE:** the first profile number is 1, and is increased at each phase change. It is set to 0 when the glider is at surface (PHASE=0) or during inflexion (PHASE=3) or if the profile is too short (less than 10 pressure measurements).
- **BBP:** Computed using a Xp factor = 1.1 and backscatter from seawater, using methods from Zhang (2009) code.

2.4.3 Quality control set and indicator

The attribute *quality_control_set* enables users to define the type of quality control flagging used for the dataset. Table 8 defines the quality control set numbers as declared in Mancini et al. (2012). When all flags in <PARAM>_QC are constant, the value is indicated using the attribute QC_indicator.

Table 8. List of QC procedure flags used in the IMOS project

Set Number	Description
1	IMOS standard set using the IODE flags
2	ARGO quality control procedure
3	BOM (SST and Air-Sea flux) quality control procedure

Table 9 shows the meanings of the flags used in ANFOG quality control. Note that for simplicity and consistency within the whole database, and to avoid subjectivity as much as possible in the QC flag values, we mainly use 0, 1, 4, 8 and 9 flag values. Data manually flagged have been attributed mostly flag 4.

Note also that all the LONGITUDE_quality_control and LATITUDE_quality_control when the glider is underwater are flagged with value 8 (longitude and latitude linearly interpolated between surfacings). However, if some of the data are slightly interpolated to match the same time series in FV01 files, the values are not flagged to 8.

Table 9. QC set number 1: IMOS standard set using the IODE 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 data 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	Altered by a QC centre, with original values (before the change) preserved in the history record of the profile.
6	Not used	Reserved for future use.
7	Not used	Reserved for future use.
8	Interpolated value	Indicates that data values are interpolated.
9	Missing Value	Indicates that the element is missing.

2.4.4 Metadata containers

In each NetCDF datafile, information regarding the Platform used, the Deployment itself, and the Sensors used to get the data present in the file is stored in empty and dimensionless variables, named 'containers'. These empty variables PLATFORM, DEPLOYMENT and SENSOR are associated with attributes that fully describe the dataset, as shown in Table 8.

Table 8. Containers attributes. For details please refer to EGO user's manual (paragraph 2.6). Note: if an attribute value is empty (information unknown), the attribute does not exist in the NetCDF file.

Variable	attributes	example
PLATFORM	trans_system_id	'Irridium'
	positioning_system	'GPS'
	platform_type	'Slocum G2'
	long_name	'Platform information'
	platform_maker	'Teledyne Webb Research'
	firmware_version_navigation	7.16
	firmware_version_science	7.16
	glider_serial_no	'209'
	battery_type	'Alkaline'
	glider_owner	'ANFOG'

	operating_institution	'ANFOG'
DEPLOYMENT	deployment_start_date	'2016-03-30-T24:23:29Z'
	deployment_start_latitude	-29.467 *(decimal degrees)*
	deployment_start_longitude	121.8523 *(decimal degrees)*
	long_name	'Deployment information'
	deployment_start_technician	'Thomson, Paul'
	deployment_end_date	'2016-04-03-T02:45:17Z'
	deployment_end_latitude	-17.1947 *(decimal degrees)*
	deployment_end_longitude	121.726 *(decimal degrees)*
	deployment_end_technician	'Thomson, Paul'
	deployment_end_status	'recovered'
	deployment_pilot	'Kiat Hong, Kah'
	Deployment_anomaly	'Fin rudder removed by shark. Mission terminated early and glider recovered'
SENSOR1	sensor_type	'CTD'
	sensor_maker	'Seabird'
	sensor_model	'GPCTD'
	sensor_serial_no	'96'
	long_name	'Sensor1 information'
	sensor_calibration_date	‘2011-11-30’ *(yyyy-mm-dd)*
	sensor_parameters	'TEMP, CNDC, PRES, PSAL'
SENSOR2	sensor_type	'ECO Puck'
	sensor_maker	'Wetlabs'
	sensor_model	'FLBBCDSLK '
	sensor_serial_no	‘2375’
	long_name	'Sensor2 information'
	sensor_calibration_date	'2011-10-06' *(yyyy-mm-dd)*
	sensor_parameters	'CPHL, CDOM, VBSC, BBP'
SENSORn	…	

2.4.5 Slocum data corrections

In version 4.0 and above of the ANFOG data processing, the person in charge of the processing can choose to apply corrections on Slocum CTD and Optode data. Indeed, the Seaglider processing chain has always included scripts developed by the University of Washington where thermal lag and small scale time lags are taken into account in salinity computation. However, in Slocum data processing, before version 4.0, these considerations were missing. Please refer to Gourcuff (2014a), Gourcuff (2014b), and Gourcuff (2014c) for more information.

3 Quality Control

3.1 Automatic tests

The following are quality control tests applied to ANFOG glider data, both in Real Time and in Delayed Mode.

1. Impossible date test

This test checks if time values are within the timeframe associated with the deployment (defined in the attributes of the variable DEPLOYMENT). Only data found within the period of the mission are quality controlled. Data outside of this timeframe are removed.

2. Impossible location test

The test requires that the observation latitude and longitude be sensible.
Latitude in range -90 to 90
Longitude in range -180 to 180

Good science data measured at locations that are bad or missing receive a flag of QC=3.

3. Range test

This test applies a gross filter on observed values for the measured scientific parameters. It needs to accommodate all of the expected extremes encountered in the oceans around Australia. Valid maxima and minima as listed in Table 7 are tested, except for the engineering parameters HEAD, UCUR and VCUR for which no QC is performed.

4. Spike test

This test is based on Argo Data Management (Argo, 2013) for temperature and salinity. A difference between sequential measurements, where one measurement is quite different than adjacent ones, is a spike in both size and gradient. The test does not consider the differences in depth, but assumes a sampling that adequately reproduces the temperature and salinity changes with depth. The algorithm is used on both the temperature and salinity profiles.

Test value = $| V2 - (V3 + V1)/2 | - | (V3 - V1) / 2 |$

where V2 is the measurement being tested as a spike, and V1 and V3 are the values above and below.

Temperature:
The V2 value is flagged when

- the test value exceeds 6.0 degree C. for pressures less than 500 db or
- the test value exceeds 2.0 degree C. for pressures greater than or equal to 500 db

Salinity:
The V2 value is flagged when

- the test value exceeds 0.9 PSU for pressures less than 500 db or
- the test value exceeds 0.3 PSU for pressures greater than or equal to 500 db

6. Gradient test

This test is based on Argo Data Management (Argo, 2013). It fails when the difference between vertically adjacent measurements is too steep. The test does not consider the differences in depth, but assumes a sampling that adequately reproduces the temperature and salinity changes with depth. The algorithm is used on both the temperature and salinity profiles.

Test value = | V2 - (V3 + V1)/2 |

where V2 is the measurement being tested as a spike, and V1 and V3 are the values above and below.

Temperature:
The V2 value is flagged when

- the test value exceeds 9.0 degree C for pressures less than 500 db or
- the test value exceeds 3.0 degree C. for pressures greater than or equal to 500 db

Salinity:
The V2 value is flagged when

- the test value exceeds 1.5 PSU for pressures less than 500 db or
- the test value exceeds 0.5 PSU for pressures greater than or equal to 500 db

7. Surface data

Bio-optical parameters (CDOM, CPHL, VBSC and BBP) and IRRADxxx data are consistently flagged as bad (quality_control = 4) above 0.5db pressure (0.5 m depth) where these parameters are often be noisy due to surface water interactions.

8. Descending IRRAD profiles

As the radiometer is angled at -20° to the horizontal plane, it is only level when the glider is ascending through the water column. Consequently, all descending IRRADxxx data are flagged 4 (bad data).

3.2 Manual QC

Manual QC is by subjective visual inspection of data values by an operator. To avoid delays, this is not mandatory before real-time distribution, and is only performed in Delayed Mode.

The parameters that are visually checked and manually flagged are: PSAL, CNDC, TEMP, DOX1 (and thus DOX2), CPHL, CDOM, VBSC and IRRADxxx. The consistency with PSAL is also checked knowing that PSAL is calculated from CNDC and TEMP, as well as the consistency with DOX1 and DOX2 in case oxygen has been recomputed using TEMP, the temperature from the CTD, and in any case for DOX2, that depends on density, i.e. PRES, TEMP and PSAL.

Examples of conditions where additional QC is applied:

- glider out of the water
- glider sitting in the seabed
- noise experienced due to bio-fouling or other reasons
- coarse outliers

Slocum Oxygen Quality Control:
For old Slocum mission (pre-2012), new oxygen corrections can't be applied due to BPHASE measurements not ouput by the glider.

3.3 QC Report

Each dataset distributed in Delayed Mode is associated with a data processing and Quality Control report, together with the NetCDF file. This report contains information about the corrections applied to the data, and their manual quality control. This report is not exhaustive and is distributed for information purposes only.

References:

Argo. 2013. Argo User's Manual. (Version 3.03).

EGO. 2013. EGO User's Manual, Version 1.01.

Gourcuff, C. 2014a. ANFOG Slocum CTD data correction.

Gourcuff, C. 2014b. ANFOG Slocum Oxygen data: new computation.

Gourcuff, C. 2014c. QC Officer's Manual. version 2.0.

Mancini, S., Tattersall, K., Proctor, R., & Galibert, G. 2012. IMOS NetCDF User's Manual, Version 1.3.

Zhang, X., Hu L, & He, M. 2009. Scatteirng by pure seawater: Effect of salinity, Optics Express, Vol. 17, No. 7, 5698-5710.