Good procedures to manage OceanGliders data with Coriolis.

V1.3 – January 2020

Author: OceanGliders Technical Coordination

NOTE: US Gliders DAC users should follow the guidelines from <https://gliders.ioos.us/data/>

Contents

[General information 2](#_Toc30003618)

[About OceanGliders 2](#_Toc30003619)

[About this document 2](#_Toc30003620)

[OceanGliders data management structure 2](#_Toc30003621)

[Data mode 2](#_Toc30003622)

[Format 2](#_Toc30003623)

[Data policy 3](#_Toc30003624)

[The good procedures for PI/Operators 3](#_Toc30003625)

[Before glider deployment 3](#_Toc30003626)

[During glider deployment 3](#_Toc30003627)

[After glider deployment 3](#_Toc30003628)

[The good procedures for DAC 3](#_Toc30003629)

[Annexe: “Upload procedure” 4](#_Toc30003630)

[PIs/Operators 4](#_Toc30003631)

[Data Assembly Centres 4](#_Toc30003632)

[Annexe: “conversion from binary to ascii for Slocum data files” 5](#_Toc30003633)

[Annexe: Filling Metadata files for Coriolis glider data management tool. 6](#_Toc30003634)

# General information

## About OceanGliders

OceanGliders is the international glider program of the GOOS. It is dedicated to the promotion of the glider technology for sustained Ocean observing and scientific studies. The OceanGliders community promotes glider applications through coordination, training, liaison between providers and users, advocacy, and provision of expert advice. We intend to favour oceanographic experiments and operational monitoring of the oceans with gliders through scientific and international collaboration. We provide information about glider projects and glider data management, as well as resources related to gliders. All OceanGliders data are publicly available.

## About this document

This document specifies the good procedures for gliders data management if Coriolis is your Data Assembly Centre (DAC) or Global DAC.

This document relies on the « [*EGO Glider's user’s manual*](http://www.ego-network.org/dokuwiki/lib/exe/fetch.php?media=public:datamanagement:ego_gliders_netcdf_format_manual_v1.2.pdf) » that define EGO format.

OceanGliders identified 2 GDACs, Coriolis and US Glider DAC. For the glider groups working with the US glider DAC, please take a look at : <https://gliders.ioos.us/data/>

## OceanGliders data management structure

Data flow within OceanGliders is organized at three stages: operator/PI, DAC and GDAC.

* Operator are deploying, piloting and recovering the gliders. They are responsible for the operation at sea and deliver the data and metadata in real time.
* The Principal Investigator (PI), typically a scientist, is responsible for the mission.
* The DAC processes raw data and metadata into EGO files following requirements of the “EGO Glider user manual” that describe the EGO format. DAC should deliver the EGO files to the GDAC in near real time.
* The GDAC archives and distributes the best copy of the data set. When a higher quality data set (e.g. calibrated data) is available, it replaces the previous version of the data file.

In some cases, PI and DAC as well as DAC and GDAC can be the same person or institution.

## Data mode

There is three temporal mode for gliders data: Real-Time data, High Frequency data (for Slocums) and delayed mode data.

* The Real-Time mode (RT) concerns the data that are transmitted by the glider.
* The High-Resolution mode (HR), or recovery mode, is the full dataset collected after glider recovery. It is High Frequency and full parameter.
* The delayed mode (DM) are quality controlled data. The delayed mode data management procedures are not established yet.

## Format

The format processed by Coriolis is called EGO format.

It uses the NetCDF (network Common Data Form) system, a set of software libraries and machine-independent data formats. Our implementation of NetCDF is based on the community-supported Climate and Forecast (CF) specification, which supplies a standard vocabulary and some metadata conventions. EGO layers several more conventions above the CF standard. These are intended to make it easier to share in-situ data, to make it simpler for the GDACs to aggregate data from multiple sites, and to ensure that the data can be created and understood by the basic NetCDF utilities.

EGO includes standard terms for the short name of both coordinate and data variables (measurements).

EGO includes controlled vocabularies for the variable attributes and metadata.

EGO data are available through Coriolis FTP website : <http://www.ifremer.fr/co/ego/ego/v2/>

The EGO format is describe in the [EGO glider user's manual](http://www.ego-network.org/dokuwiki/lib/exe/fetch.php?media=public:datamanagement:ego_gliders_netcdf_format_manual_v1.2.pdf).

## Data policy

Real Time data in EGO format are publicly available through GDAC FTP, data selection interface and GTS.

# The good procedures for PI/Operators

### Before glider deployment

* Let Coriolis (codac@ifremer.fr) and JCOMMOPS Technical Coordinator (support@jcommops.org) know about the glider mission planning. Share the main details (WMO/glider name, start date).
* Prepare metadata files. See annexe “Filling Metadata files for Coriolis glider data management tool”.
* Upload metadata files on your Coriolis ftp account following the naming convention and proper procedure. See Annexe “Upload procedure”.

### During glider deployment

* Upload real time data file on your Coriolis ftp account following the procedure described in the Annexe “Upload procedure”

In case of an update of the metadata files during the glider mission, you can upload the updated files on your Coriolis ftp account following the same procedure.

### After glider deployment

* Update the metadata file with mission end information and upload the file on your Coriolis ftp account following the same procedure.
* Upload the full data set on your Coriolis ftp account following the same procedure.

# The good procedures for DAC

DAC should upload the data set in EGO format on the Coriolis FTP repository following the naming convention and proper procedure. See Annexe “Upload procedure”.

# Annexe: “Upload procedure”

Coriolis is using an ftp repository to collect real time gliders data. Every glider groups must have an account. To get an account, send a message to codac@ifremer.fr with support@jcommops.org in copy.

Glider groups are encourage to send a message to codac@ifremer.fr and support@jcommops.org when glider mission is about to start.

Data and metadata files must be submitted in the proper directory from users ftp account: **submit/<glider\_name>/<glider\_name>\_<glider\_mission\_start\_date>/**

* <glider\_name> refers to <platform\_code> (see annexe “Filling metadata files for Coriolis glider data management tool”)
* <glider\_mission\_start\_date> refers to <deployment\_start\_date> (see annexe “Filling metadata files for Coriolis glider data management tool”) and should follow YYYYMMDD format (no hours, minutes, seconds).

### PIs/Operators

**Raw data files managed by Coriolis:**

* For SeaGliders: the per dive \*.nc files.
* For SeaExplorer: The \*.gli, \*.gz and their associated \*.pl, \*.gz files
* For Slocums: The \*\_sbd.dat files with their associated \*\_sbd.m files. See annexe “conversion from binary to ascii for Slocum data files”

**Metadata files managed by Coriolis**

Without metadata files Coriolis can not process the data set. Metadata file are described in the annexe “Metadata files for Coriolis glider data management tool”.

### Data Assembly Centres

The EGO file names use the following naming convention for data files: <glider\_name>\_<deployment\_code>\_<data\_mode>.nc.

* <glider\_name> refers to <platform\_code> (see annexe “filling Metadata files for Coriolis glider data management tool” for more details).
* <deployment\_code> refers to name of the glider mission (see annexe “filling metadata files for Coriolis glider data management tool” for more details).
* <data\_mode> refer to this list:
	+ R: Real-time data
	+ P: Provisional data
	+ D: delayed mode
	+ M: mixed delayed mode and real-time

**Example**

milou/milou\_20150112/milou\_mooseperseust02\_08\_R.nc

This file contains observations and metadata from the Milou glider, from the deployment performed in January 2015.

# Annexe: “conversion from binary to ascii for Slocum data files”

Raw data send by the slocum glider in real time are binary files (\*.sbd, \*.tbd). However, for some reasons, raw data files managed by Coriolis for Slocum gliders are ascii files, the \*\_sbd.dat files with their associated \*\_sbd.m files.

Dockserver can convert automatically binary \*.sbd, \*.tbd files into \*.sbd.m and \*.sbd.dat using a Teledyne exec file (dbd2sc.exe, dba\_merge.exe, dba2orig\_matlab.exe).

However, glider groups may struggle to recover the ascii files directly from dockserver and may need to convert the binary raw file into ascii raw files.

We have developed a script in Matlab and Python to do the conversion from binary to ascii. It is available here : <https://www.ego-network.org/dokuwiki/lib/exe/fetch.php?media=public:datamanagement:code_conversion_raw_slocum.zip>. You can find three different version of the conversion script, depending on the use, and the Teledyne exec files. There is no guideline or manual to use those scripts.

This tool is a collaborative work and the OceanGliders community will appreciate to benefit from your improvements. Feel free to share any improvement at support@jcommops.org.

# Annexe: Filling Metadata files for Coriolis glider data management tool.

This section deserves a proper document you can find here: <https://www.ego-network.org/dokuwiki/lib/exe/fetch.php?media=public:datamanagement:v1.3:guidelines_to_fill_a_json_file_v1.1.3.docx>