

GRIIDC Glider Data Submission Guidance

In the last decade, autonomous underwater glider observations have increased and the technology has drastically improved providing a multidisciplinary approach to oceanographic studies. Gliders can allow coupling traditional oceanographic measurements such as temperature, salinity, and currents with acoustics, phytoplankton and zooplankton abundance, oxygen, nitrate, animal telemetry, and pH among others. Glider observations provide the vertical resolution that is often under-sampled by traditional oceanographic measurements while deployed for long periods of time and traveling up to thousands of kilometers per deployment.

In 2014, the Integrated Ocean Observing System (IOOS) proposed the creation of the IOOS Underwater Glider Network and National Glider Data Assembly Center (NGDAC). NGDAC receives glider data from operational glider in U.S. coastal waters and can also ingest data from glider missions that operate outside U.S. when data sharing is permitted. Currently, the NGDAC accepts physical oceanographic properties such as density, salinity, conductivity, temperature, and currents. Other biological and chemical measurements are not included in the archive, however future plans to receive these datasets are in progress.

GRIIDC can provide a venue for data management of glider data that is not currently submitted to NGDAC. GRIIDC will eventually submit the data to National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI) for long-term archiving. Therefore, the glider data received by GRIIDC should follow the same format and metadata requirements held by NGDAC to allow for data compatibility and later submission to NCEI. This guidance document outlines requirements for glider data submission to GRIIDC to facilitate data transfer to NCEI.

Glider Terminology Used by NGDAC:

- Profile: A single vertically oriented track of a glider, either upward or downward through the water column. A profile is one-half of a dive.
- Dive: A single vertical profile to depth followed by a vertical profile towards the surface. A dive does not necessarily begin with or terminate with a surfacing and/or GPS fix.
- Segment: The set of data collected between 2 GPS fixes obtained while the glider is on the surface of the water.
- Trajectory or Deployment: A series of one or more segments completed by a glider between the time of deployment and the time of recovery.

Glider Observation Data

For physical oceanographic properties such as density, salinity, conductivity, temperature, and currents, GRIIDC encourages the data to be submitted directly to NGDAC and provide a link to the dataset with the data submission to comply with GoMRI requirements.

Additional glider data that can be submitted to GRIIDC includes ocean optics data such as backscattering, photosynthetically available radiation (PAR), turbidity, fluorescence (e.g., chlorophyll-a, colored dissolved organic matter [CDOM]), biological data such as phytoplankton and/or zooplankton abundances, dissolved oxygen, dissolved nutrients, chemical data such as pH, and passive acoustic data. This list is not exhaustive and other datasets are welcome as far as they include the appropriate metadata and quality control information.

Data Submission Format

Data should be submitted in ASCII or NetCDF format. ASCII files written in space, comma, or tab delimited variables are easily read by both humans and computers. Common file extensions include *.txt and *.cnv. Spreadsheet data (e.g. Excel) should be saved as comma separated variables (*.csv) or other text formats for submission rather than Excel workbooks (*.xls,*.xlsx). NetCDF is a widely-used self-describing scientific data format that is compatible with a large number of common software analysis packages such as Matlab and is accepted by NCEI.

There should be sufficient information submitted with the data (preferably inside the individual files themselves) such that the data are completely understandable. This information includes variable and parameter names (e.g., t is temperature), the units of measure (e.g., degrees Celsius), the latitude and longitude coordinates and depth where collected, and date and time when collected. Date-time should include time zone or be expressed in UTC whenever possible and ideally written in ISO 8601 format (e.g., 2015-10-25T22:34:51Z).

Filename Format

The filename should include the glider name, start time of data acquisition, and whether the data is real-time (rt) or delayed mode. The date should be in ISO 8601 format and UTC.

Examples: glidername_yyyymmddTHHMMSSZ_rt.nc
Ru30_20180120T102030Z_rt.nc
Ru30_20180120T102030Z_delayed.nc

Global Attributes Information

While most of the information for the global attributes used by NGDAC are currently required during the metadata submission process it is important to make sure the following information is available: glider type (seaglider, spray glider, slocum glider), processing level (description of the processing and/or quality control), and references for the method. The metadata should include the title, abstract, keywords, contributor names and roles, the creator name, e-mail and URL, the date the file was created, issued and modified.

The NGDAC also requires an ID and the World Meteorological Organization (WMO) ID. The WMO ID is a unique identifier for the dataset very similar to the filename which includes the glider name and the deployment time (e.g., glider-YYYYmmddTHHMM, ru30-20140101T0000). The WMO ID is a unique value

assigned to a specific glider deployed/location. The WMO ID can be requested from the NGDAC administrator (see link in references for more information).

Variables and Metadata

Glider data variables can be divided into three main categories: time-series, profiles, and container. Time-series variables are dependent of time, profile variables are dimensionless, and container variables are used to capture metadata regarding the platform and instrumentation. Individual profiles can be submitted to the NGDAC, who will aggregate all individual NetCDF profiles containing the same trajectory value.

Each variable should have the following information: long name (not an acronym), units, filling value (e.g., 999), and instrument. Other variables such as accuracy, precision, resolution, name of ancillary variables (e.g., quality control), comments, valid minimum and maximum values are encouraged to be included in the NetCDF files or metadata whenever they are available. Information about the variables can be included in the NetCDF, readme file (ASCII text file), or metadata. Please do not submit readme files in Microsoft Word format.

Time, latitude, longitude, and depth are required for all time-series variables such as temperature, salinity, and chlorophyll-a. Profile variables are not required with the submissions to GRIIDC but if they are available please included the profile_id and the mid-point latitude, longitude, and time.

Table 1: Summary of the variable types, description, and a NetCDF example of the attributes.

Type	Description	Variables	Examples (NetCDF)
Trajectory variables	Character array that identifies the deployment	<ul style="list-style-type: none"> Trajectory variable contains text using the format: Glider-YYYYmmddTHHMM. 	<pre>char trajectory(traj_strlen) ; trajectory:cf_role = "trajectory_id" ; trajectory:comment = "A single deployment..." ; trajectory:long_name = "Trajectory/Deployment Name" ;</pre>
Profile variables	Dimensionless variables used by NGDAC to provide access to individual profiles from the aggregated dataset	<ul style="list-style-type: none"> Unique identifier for the profile (1 to number of profiles in a trajectory). Latitude, longitude, and time refer to the mid-point of the profile. Depth-average current (u- and v-components) of the net current measured while the glider is underwater. Quality control if applicable. 	<pre>int profile_id ; profile_id:FillValue = -999 ; profile_id:comment = "Sequential profile number ..." ; profile_id:long_name = "Profile ID" ; profile_id:valid_max = 2147483647 ; profile_id:valid_min = 1 ; double u ; u:_FillValue = -999. ; u:comment = "The depth-averaged current ..." ; u:long_name = "Depth-Averaged Eastward ..." ; u:observation_type = "calculated" ; u:platform = "platform" ; u:standard_name = "eastward_sea_water_velocity" ; u:units = "m s-1" ; u:valid_max = 10. ; u:valid_min = -10. ;</pre>
Time-series	Dimensioned along the	<ul style="list-style-type: none"> Time stamp corresponding to the acquisition of the 	<pre>double time(time) ; time:ancillary_variables = "time_qc" ; time:calendar = "gregorian" ;</pre>

variables	time axis	<p>sensor data for the profile.</p> <ul style="list-style-type: none"> • Latitude, longitude, depth, and pressure along the time variable. • All measurements collected along track (time) such as temperature, salinity, chlorophyll-a, etc. • Quality control variables if applicable. 	<pre>time:long_name = "Time" ; time:observation_type = "measured" ; time:standard_name = "time" ; time:units = "seconds since 1970-01-01T00:00:00Z" ; Double temperature(time); temperature:FillValue = -999. ; temperature:accuracy = " " ; temperature:ancillary_variables = "temperature_qc" ; temperature:coordinates = "time latitude longitude depth" ; temperature:instrument = "instrument_ctd" ; temperature:ioos_category = "Temperature" ; temperature:long_name = "Sea Water Temperature" ; temperature:observation_type = "measured" ; temperature:platform = "platform" ; temperature:standard_name = "sea_water_temperature" ; temperature:units = "Celsius" ; temperature:valid_max = 40. ; temperature:valid_min = -5. ;</pre>
Container variables	Used to store metadata about glider and instruments	Platform = metadata about the glider Instrument variables (e.g., CTD) to store information about the instruments.	<pre>int platform ; platform:_FillValue = -999 ; platform:comment = "Slocum ..." ; platform:id = "ruXX" ; platform:instrument = "instrument_ctd" ; platform:long_name = "Glider name" ; platform:type = "platform" ; platform:wmo_id = " " ; # WMO ID specific to this glider int instrument_ctd ; instrument_ctd:_FillValue = -999 ; instrument_ctd:calibration_date = "Date ISO 8601 " ; instrument_ctd:calibration_report = " " ; instrument_ctd:comment = "pumped CTD" ; instrument_ctd:factory_calibrated = " Date ISO 8601 " ; instrument_ctd:long_name = "Seabird Glider CTD" ; instrument_ctd:make_model = "Seabird GPCTD" ; instrument_ctd:platform = "platform" ; instrument_ctd:serial_number = "xxxx" ; instrument_ctd:type = "platform" ;</pre>

Quality Control Flags

While GRIIDC does not require quality control flags, users are encouraged to provide them when available. This variable is an array that contains values conveying information on the data quality status of the value in the variable. The size of the quality control flag matrix should be the same as the variable of interest. Table 2 provides the simple set of flags and associated descriptions adopted by U.S. IOOS. For additional information regarding flags, see the Manual for the Use of Real-Time Oceanographic Data Quality Control Flags (U.S. IOOS 2014) posted on the U.S. IOOS QARTOD website. The U.S. IOOS QARTOD

website also provides quality control guidelines for all the different types of real-time in-situ oceanographic observations.

Table 2: Quality control flags for real-time data (UNESCO 2013).

QC Flag	Description
Pass=1	Data have passed critical real-time quality control tests and are deemed adequate for use as preliminary data.
Not evaluated=2	Data have not been QC-tested, or the information on quality is not available.
Suspect or high interest = 3	Data are considered to be either suspect or of high interest to data providers and users. They are flagged suspect to draw further attention to them by operators.
Fail = 4	Data are considered to have failed one or more critical real-time QC checks. If they are disseminated at all, it should be readily apparent that they are not of acceptable quality.
Missing data=9	Data are missing; used as a placeholder.

Example of temperature quality flags in NetCDF file:

```
byte temperature_qc(time) ;
    temperature_qc:_FillValue = -127b ;
    temperature_qc: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_value missing_value" ;
    temperature_qc:flag_values = 0b, 1b, 2b, 3b, 4b, 5b, 6b, 7b, 8b, 9b ;
    temperature_qc:long_name = "temperature Quality Flag" ;
    temperature_qc:standard_name = "sea_water_temperature status_flag" ;
    temperature_qc:valid_max = 9b ;
    temperature_qc:valid_min = 0b
```

References

Manual for the Use of Real-Time Oceanographic Data Quality Control Flags (U.S.IOOS 2014;
https://cdn.ioos.noaa.gov/media/2017/12/QARTOD-Data-Flags-Manual_Final_version1.1.pdf)

NetCDF file specifications: <https://github.com/ioos/ioosngdac/wiki>

NGDAC website: <https://gliders.ioos.us/>

Paris. Intergovernmental Oceanographic Commission of UNESCO, 2013. Ocean Data Standards, Vol.3: Recommendation for a Quality Flag Scheme for the Exchange of Oceanographic and Marine

Meteorological Data. (IOC Manuals and Guides, 54, Vol. 3.) 12 pp.
http://www.nodc.noaa.gov/oceanacidification/support/MG54_3.pdf

U.S. IOOS QARTOD website: <https://ioos.noaa.gov/project/qartod/>

WMO ID information: <http://www.wmo.int/pages/prog/amp/mmop/wmo-number-rules.html>