Glider Toolbox Quick Start

The glider toolbox (GTB) is a set of MATLAB/Octave scripts and functions developed at SOCIB to manage the data collected by a glider fleet. They cover the main stages of the data management and data processing both in real time and delayed time mode.

# Requirements

* Linux or Windows OS
* Matlab 2012, Matlab 2014 or Octave.
* C++ compiler
* [Imagemagick](https://sourceforge.net/projects/imagemagick)
* External libraries
  + mexnc and snctools (Commit #4053):   
    NetCDF library preferences ( [Sourfoge/mexcdf](https://sourceforge.net/projects/mexcdf/files/mexcdf/mexcdf.r4053.zip/download)). Octave users may install octcdf instead of snctools.
  + m\_map 1.4 library: Mapping toolbox ([m\_map1.4.tar.gz](http://www.eos.ubc.ca/~rich/m_map1.4.tar.gz))
  + seawater 3.3 library: CSIRO Sea Water Library ([seawater\_ver3\_3.1.zip](http://www.marine.csiro.au/cgi-bin/marlin-dl/seawater/seawater_ver3_3.1.zip))
  + m2html: inline html documentation generator
* Webb Research package for Slocum database   
  (github: [kerfoot/slocum](https://github.com/kerfoot/slocum/tree/master/linux-bin) for linux or [coolcloud.mote/slocum](http://coolcloud.mote.org/glider_data/slocum_tools/windows/) for windows)
  + dba2asc: converts Slocum binary data to ascii format
  + dba\_merge: merge ascii navigation and science files
  + dba\_sensor\_filter and dba2\_time\_filter: filter data based on sensor or timestamps
  + dba2\_orig\_matlab: converts ascii Slocum files to Matlab/Octave data files
* Matlab jar libraries:
  + Data base drivers 9.4 ([postgresql-9.4.1212.jre6.jar](https://jdbc.postgresql.org/download/postgresql-9.4.1212.jar)). MATLAB interface with the PostgreSQL JDBC (Java Database Connectivity) driver. An Octave alternative would be needed. It may be installed from the distribution repositories as a package libpg-java (renamed to libpostgresql-jdbc- java in Debian repositories on March 2012).
  + NetCDF 4.2 library ([netcdfAll-4.2.jar](ftp://ftp.unidata.ucar.edu/pub/netcdf-java/v4.2/netcdfAll-4.2.jar) from Unidata)

# Installation

* Clone repository at <https://github.com/socib/glider_toolbox>.   
  This creates a directory glider\_toolbox with the following structure:
  + bin: binary scripts supporting the GTB
  + config: contains configuration files.
  + ext\_lib: contains external libraries and binaries.
  + glider\_data: contains files for local processing when using default configuration
  + m: contains Matlab/Octave code
  + notes: contains schematics and manuals
* Download and uncompress external libraries, binary files and matlab jars. Users can run the gtb\_install\_extlibs under the bin folder. This script downloads and uncompresses these files automatically to the right locations.
  + External libraries to ext\_lib/lib resulting in the following folders under ext\_lib/lib: m2html, m\_map, mexcdf/mexnc, mexcdf/snctools, seawater.
  + Webb Research binaries to ext\_lib/bin
* Download jar libraries and configure Matlab. Users that run the installation script can find the jar files under ext\_lib/matlab
  + Copy jar files to $MATLAB\_ROOT/java/jarext
  + Add path to $MATLAB\_ROOT/toolbox/local/classpath.txt
    - $matlabroot/java/jarext/netcdfAll-4.2.jar
    - $matlabroot/java/jarext/postgresql-9.4.1212.jre6.jar
* Setup Mex libraries from Matlab. Run matlab glider\_toolbox and call the following commands.  
   >> cd m  
   >> mex -setup (select 0 to answer prompt question)  
   >> configGliderToolboxPath  
   >> setupMexPosixtime  
   >> setupMexSFTP  
   >> exit
* Create GTB library documentation: Run **make doc** from the glider\_toolbox. Additionally, an online version of the of the documentation is available [online](http://www.socib.es/users/glider/glider_toolbox/doc).
* Edit configuration files for real time (config/configMainRT.txt) and delayed mode (config/configMainDT.txt).
  + Local and public paths
  + Database information
  + Dockserver

# Processing

There are two ways to process glider data for each data mode (real or delayed time). First, users may process glider data by running the main\_glider\_data\_processing\_***X***t scripts where ***X*** defines “r” or “d” for real or delayed time respectively. These scripts use the configMain***X***R.txt configuration files to overwrite the default configurations set by the Matlab functions. The default configuration of the configuration files is enough to start processing except for the dockserver and/or database information. If necessary, users must complete the information of their database and dockservers before start processing. There are no edition required to the Matlab software as in previous versions of the GTB. In addition, users can call the gliderDataProcessing or the deploymentDataProcessing functions to run processing of glider data. The input variables of these functions include the information of the data sets, configuration and processing mode.

## *Real time script*

1. The [main\_glider\_data\_processing\_rt](http://www.socib.es/users/glider/glider_toolbox/doc/m/main_glider_data_processing_rt) script executes the realtime processing. The script uses most of the default configuration but custom configuration may be set by editing configMainRT.txt. The usual custom configuration are in the sample file from the repository. It includes the paths for the local and public locations, the database information access and the dockserver connection information. Notice that the processing\_mode is defined in the configuration file but this value should remain to rt indicating real time processing mode.

## *Delayed time script*

1. The [main\_glider\_data\_processing\_dt](http://www.socib.es/users/glider/glider_toolbox/doc/m/main_glider_data_processing_dt) executes the delayed time processing. The script uses most of the default configuration but custom configuration may be set by editing configMainDT.txt. The usual custom configuration are in the sample file from the repository. It includes the paths for the local and public locations, the database information access and the format conversion option. Notice that the processing\_mode is defined in the configuration file but this value should remain to dt indicating delayed time processing mode. The main\_glider\_data\_processing\_dt script takes the deployment indexes from the configMainDT.txt configuration file contrary to previous versions of the GTB where the user was required to edit the configDTDeploymentInfoQueryDB.m file. The indexes are defined as db\_access.deployment\_ids with deployment numbers separated by the “|” symbol.

## *Processing functions*

The GTB recently includes a set of processing functions that allows its users to input custom information for specific runs. The first function [deploymentDataProcessing](http://www.socib.es/users/glider/glider_toolbox/doc/m/processing_tools/deploymentDataProcessing) takes the information of the deployment, the paths to the data, the configuration definition and the processing mode which are used to process a specific and single deployment. Deployment information may be input directly into the function as explained in the Deployment Metadata section. A configuration file or structure can be input to define the processing configuration. Another difference with respect to the processing scripts is that users can choose what products are made by setting the appropriate paths in the input data\_paths. For more information refer to the documentation of the code.

The second processing function is [gliderDataProcessing](http://www.socib.es/users/glider/glider_toolbox/doc/m/processing_tools/gliderDataProcessing). It wraps [deploymentDataProcessing](http://www.socib.es/users/glider/glider_toolbox/doc/m/processing_tools/deploymentDataProcessing) and it allows managing datasets from multiple deployments. This function is specially useful if a metadata database is set with the information of the deployments. However, users can also defined a list of deployments by inserting an deployment\_list array with the list of deployments or using the deployment file. See the Deployment Metadata section for more information about how to define deployment information.

# Configuration files

Previous versions of the GTB required editing Matlab configuration functions to customize the configuration to the needs of the users. Currently, the approach of the GTB differs in the sense that the configuration files are set to a default configuration that allows the user to process data in a basic mode without the need to edit any code. For a more complex setup, users add or edit text files that are read by the processing scripts and functions to overwrite the default values. With this approach, different types of data processing are performed by running it using different configuration files. Configuration files may live in any place in the disk but it is recommended to use the config folder in the installation path for consistency. The configMainRT.txt and configMainDT.txt are used by the real and delayed time mode processing scripts. Custom files may be created in the config folder with no arbitrary name format. No parameter is required to be defined in the configuration file except for the database and dockserver information which are not set to any default value. Any parameter that is present in the configuration file will be used for the processing over default or input variables of the functions. The available parameters are defined in the configTemplate.txt file that also describe the parameters using commented lines. Users can use the “#” to write comments in their own configuration files.

# Deployment Metadata

1. Metadata of the deployment may be input in three different ways. By setting to 1 the active parameter of the db\_access configuration, users can retrieve the information from their database server. They must setup the server and user information in the db\_access configuration. Additionally, deployment information can be input by using deployment configuration files. The approach for the deployment files is similar to the one of the configuration files. However, all the information must be set since there are no default values as for the configuration. Users can use the “#” to write comments in their own configuration files. Lastly, the deployment information can be input to the processing function as matlab structures. For all of the options, the fields that must be defined by the deployment are in the table 1.

|  |  |
| --- | --- |
| **Structure field** | **Database field** |
| deployment\_id | deployment\_id |
| deployment\_name | deployment\_name |
| deployment\_start | deployment\_initial\_date |
| deployment\_end | deployment\_end\_date |
| glider\_name | platform\_name |
| glider\_serial | instrument\_serial |
| glider\_model | instrument\_model |

**Table1 - Deployment information**

Additional fields may be described to overwrite the values of the global attributes of the NetCDF metadata. These fields must be named identically as the attribute to be filled.

## *Metadata Database*

The GTB provides tools to retrieve the deployment metadata from a database. By default, a table (named deployment) in a database (named instrumentation) is query to retrieve the keywords that match the fields of the deployment structure as described by table 1. Potentially, users can use any database, table and field names. However, this is not allowed yet using the configuration files. Therefore, they are required for this purpose to edit the configDTDeploymentInfoQueryDB and/or configRTDeploymentInfoQueryDB accordingly.

## Metadata Deployment Structure

The deployment metadata may be input to the [gliderDataProcessing](http://www.socib.es/users/glider/glider_toolbox/doc/m/processing_tools/gliderDataProcessing) function as an array of structures using the deployment\_list option or as a structure to the [deploymentDataProcessing](http://www.socib.es/users/glider/glider_toolbox/doc/m/processing_tools/deploymentDataProcessing) function. Processing scripts do not provide this option. The required fields of the structure are deployment\_id, deployment\_name, deployment\_start, deployment\_end, glider\_name, glider\_serial and glider\_model. Additionally, other fields may be used to overwrite values of the global attributes of the NetCDF metadata.

## Metadata Deployment Text Files

Deployment information can also be read from a configuration file. The two processing scripts ([main\_glider\_data\_processing\_dt](http://www.socib.es/users/glider/glider_toolbox/doc/m/main_glider_data_processing_dt) and [main\_glider\_data\_processing\_rt](http://www.socib.es/users/glider/glider_toolbox/doc/m/main_glider_data_processing_rt)) read deployments from the config/deploymentDT.txt and config/deploymetnRT.txt files (respectively) if the db\_access.active parameter is set to 0. Any custom file defining the deployments may be read by the [gliderDataProcessing](http://www.socib.es/users/glider/glider_toolbox/doc/m/processing_tools/gliderDataProcessing) function when db\_access.active is 0. Users should input the file name when calling the function. The deployment file should contain the information of the deployment (same as the structure) and any optional field to overwrite global attributes of the NetCDF metadata. The deployment file can contain any amount of deployments by using the index approach. Below we present an example of the definition of two deployments:

deployment\_list(1).deployment\_id = 9

deployment\_list(1).deployment\_name = NOV2017\_TERESA\_GFMR0064

deployment\_list(1).deployment\_initial\_date = 7.3700e+05

deployment\_list(1).deployment\_end\_date = NaN

deployment\_list(1).platform\_name = teresa

deployment\_list(1).instrument\_serial = 518

deployment\_list(1).instrument\_model = Slocum G2 Deep

deployment\_list(2).deployment\_id = 10

deployment\_list(2).deployment\_name = NOV2017\_SDEEP04\_GFMR0065

deployment\_list(2).deployment\_initial\_date = 7.3700e+05

deployment\_list(2).deployment\_end\_date = NaN

deployment\_list(2).platform\_name = sdeep04

deployment\_list(2).instrument\_serial = 520

deployment\_list(2).instrument\_model = Slocum G2 Deep

deployment\_list(2).abstract = Deployment for scientific tests

deployment\_list(2).author = John Smith