IML QC Instructions V1

# March 2019

## Emily Chisholm

**Step 1 – Data Collection**

Collect data in BCD format

Headers should read:

|  |
| --- |
| [1] "DIS\_DATA\_NUM" "MISSION" "MISSION\_DESCRIPTOR"  [4] "EVENT\_COLLECTOR\_EVENT\_ID" "EVENT\_COLLECTOR\_STN\_NAME" "DIS\_HEADER\_START\_DEPTH"  [7] "DIS\_HEADER\_END\_DEPTH" "DIS\_HEADER\_SLAT" "DIS\_HEADER\_SLON"  [10] "DIS\_HEADER\_SDATE" "DIS\_HEADER\_STIME" "DIS\_DATA\_TYPE\_SEQ"  [13] "DATA\_TYPE\_METHOD" "DIS\_DETAIL\_DATA\_VAL" "DIS\_DETAIL\_DATA\_QC\_CODE"  [16] "DIS\_DETAIL\_DETECTION\_LIMIT" "DIS\_DETAIL\_DETAIL\_COLLECTOR" "DIS\_DETAIL\_COLLECTOR\_SAMP\_ID"  [19] "CREATED\_BY" "CREATED\_DATE" "DATA\_CENTER\_CODE"  [22] "INSTITUTE" "PROCESS\_FLAG" "BATCH\_SEQ"  [25] "DIS\_SAMPLE\_KEY\_VALUE" |
|  |
| |  | | --- | |  | |

Ensure data includes, temp, salinity, nutrient, depths, and all relevant metadata

**Step 2 – Pre Processing in R**

1. Ensure you have a working directory established, with a folder named “IML\_QC” for storing IML format files being produced.
2. Make sure package ‘reshape’ is called into your library

* require(reshape)

1. Make sure you have files:

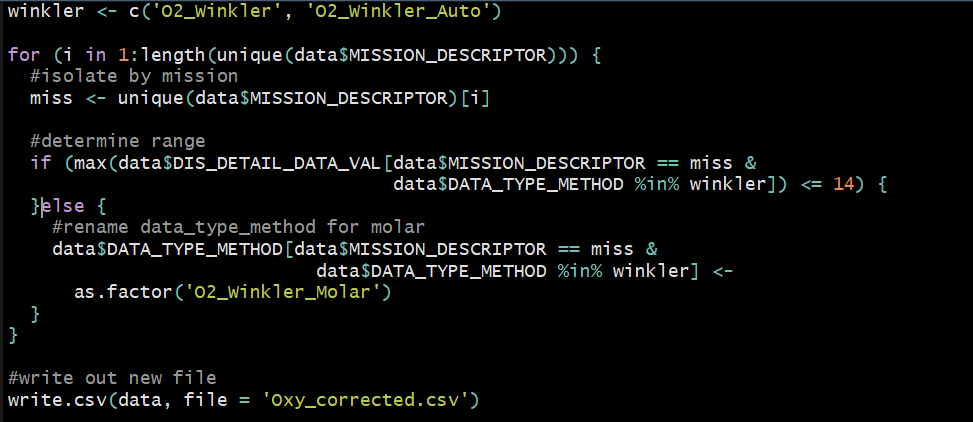
* convertOxy.R
* BCD2IML\_format.r
* Get\_flags\_IML2BCD\_EC.r

*Note:* If QC involves oxygen data in mmol/m\*\*3, convert values to ml/l before running QC using convertOxy.R function.

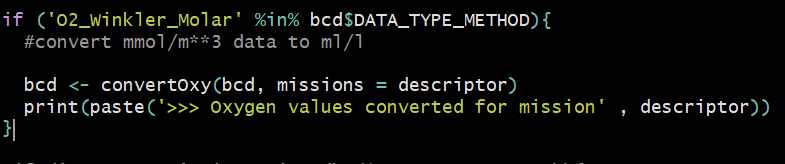
**Sub Step 2.1 – Converting Oxygen Data**

**Determine which cruises in data set contain data in mmol/m\*\*3 units**

* + 1. **Rename with DATA\_TYPE\_METHOD set to O2\_Winkler\_Molar to distinguish between measurement units**



* + 1. Run convertOxy.R function inside loop of BCD2IML\_format.R



* + 1. Oxygen can then be run through QC in ml/l and flags will map back to original mmol/m\*\*3 data (no rounding or conversion errors in data set)

1. Ensure BCD\_IML\_map\_upd.csv includes all data types that are present in your data set mapping to appropriate IML codes and units.
2. Ensure sample key value is created properly following standard structure

Eg.

bcdo$DIS\_SAMPLE\_KEY\_VALUE=paste0(bcdo$MISSION\_DESCRIPTOR,"\_",sprintf("%03d",bcdo$EVENT\_COLLECTOR\_EVENT\_ID),"\_",bcdo$DIS\_DETAIL\_COLLECTOR\_SAMP\_ID)

1. Loop through BCD2IML\_format.r for each cruise included in your data set.

**Step 3 – Processing in MATLAB**

1. Ensure your path includes

* MLI CTD package
* Seawater Package
* Rosette folder with IML QC scripts (modified at BIO)
* You are working out of you IML\_QC folder with all your IML formatted data

1. Run file B\_batch\_BIO.m

Note: ensure you are using B\_control\_Q\_GL.m (modified by Gordana Lazin for BIO) NOT B\_control\_Q.m

1. Move all QC files produced to a distinct folder in working directory (eg. **QC\_V1)**

**Step 4 – Post Processing in R**

1. Run through get\_flags\_IML2BCD\_EC.r

This script has been modified since Gordana’s 2016 version with bug fixes and generalizations

*Note:* ensure that line 67

# file name with path to the QC data

fnp=file.path(getwd(),"**QC\_V1**",fn)

“**BOLD**” is the name of the folder where you have put your IML QC files, specific to your working directory

**Step 4 - Summary**

1. BCD files will be created in distinct cruise folders in the file “BCD” in your working directory, the csv summary of all flags will be created in your main working directory (not a subfolder)
2. Check IMLTroubleshooting guide if errors are encountered

**Step 5 – Plotting**

Run through Plotting scripts

This requires

* + Plot\_flagged\_profiles\_EC.R
  + Plot\_profiles.R
  + Flagcol2.R
  + Lubridate package
  + BCD files
  + QC data files
  + IML BCD map file

Edit Plot\_flagged\_profile\_EC.R to run through missions of interest, can be performed individually or in batch. Ensure file paths are updated and correct, source the file to produce profile plots of each variable with flags highlighted.

* **Set Working directory**
* **Choose mission or batch list of missions (line 14)**
* **Set file paths to find BCD files (line 16)**
* **Set file paths to find qc files (line 27)**
* **Set file path to find IML to BCD map (line 41)**

This step will produce profile plots for each flagged profile, with a separate file for each event and a distinct plot per variable within each file. The files will be stored in the BCD folder for each distinct cruise.

**Step 6 – Flag review**

Flag review takes careful consideration and should be done by an expert who has extensive knowledge on the region and the variables being considered.

Flagged profiles should be analyzed and compared to BCD files. Then flags in BCD files can be updated as needed.

Table 1. IML flag meanings

|  |  |
| --- | --- |
| Flag | Meaning |
| 0 | no quality control |
| 1 | value seems correct |
| 2 | value appears inconsistent with other values |
| 3 | value seems doubtful |
| 4 | value seems erroneous |
| 5 | value was modified |
| 6 | reserved for future use |
| 7 | possible problem with data point—further investigation required (IML temporary flag) |
| 8 | reserved for future use |
| 9 | value missing |

The most common flag requiring review is the 7 flag, which should not be left in the final BCD files. Other flags should be verified but not changed without significant consideration and reason. When investigating data points flagged 7 it is important to look for large differences between CTD and bottle measurements, frequently a poor CTD calibration/cast can make the bottle measurements appear faulty. These 7 flags should be changed in the final BCD files to reflect the data as erroneous or valid.