IML Trouble Shooting Guide

## March 2019

### Emily Chisholm

|  |  |  |
| --- | --- | --- |
| **ISSUE** | **FILES INVOLVED** | **RESOLUTION** |
| Data files need to be read offline without querying BioChem | BCD2IML\_format.R  line 20-24 | Commented out BioChem query lines, replaced with read.csv() of pertinent data file pulled using SQL |
| Missing functions in MatLab | Textcell.m  Strmachi.m  Sw\_dpth.m  Remove\_doubtful.m  +++ | Required MatLab path to include MLI\_CTD\_Toolbox  SeawaterCsiro\_ver2\_0\_2  Rosette  All provided by Gordana Lazin or from IML (Seawater also available on shared drive, MLI\_CTD available on SVN) |
| NaN dates being produced in IML\_format.txt files | BCD2IML\_format.R | Edited line 105 to produce correct format, changed POSIXct format to ‘%d-%b-%Y %H%M’ from ‘%Y-%m-%d %H%M’ |
| Error loading ‘east\_mask’ file | B\_batch\_BIO.m  (line 53-57) in ‘B\_control\_Q.m’ | Commented out visual inspection (Stage 5), determined to not work for Scotain Shelf area  Change made permanent in B\_control\_Q\_GL.m’ |
| producing blank files ‘QC\_tests\_performed’ and ‘QC\_data…’ | B\_batch\_BIO.m | Ensure you are using B\_control\_Q\_GL.m  Resolved by commenting out visual inspection stage of QC |
| Modified to run without querying BioChem (offline) | Get\_flags\_IML2BCD.r | Commented out BioChem query lines, replaced with read.csv() of pertinent data file pulled using SQL  Changes made permanent in Get\_flags\_IML2BCD\_EC.r |
| file relates mission descriptors and cruise names but only for AZMP cruises | AZMP\_mission\_name\_descriptor.csv | Would be very difficult to replicate for entire dataset. Modified script to pull Mission descriptor for file names.  Scripts modified to run without AZMP\_mission\_name\_descriptor.csv |
| Failure to map particular data type | BCD\_IML\_MAP.csv | Ensure map is updated with all included data types. Modified by EC as BCD\_IML\_MAP\_upd.csv |
| Some loops are generating warnings about NA (invalid factors) being generated | BCD2IML\_format.r | Had to initiate new variables as factors, resolved in new script version.  Eg.  data$DATA\_TYPE\_METHOD <- factor(data$DATA\_TYPE\_METHOD, levels = c(DATA\_TYPE\_METHOD, 'O2\_Winkler\_Molar')) |
| DIS\_SAMPLE\_KEY\_VALUE is formatted differently for various cruises | BCD2IML\_format.r | Gordana’s script assumed it would be consistent (‘missiondescriptor\_event\_sampleID’) create sample key value from pieces at beginning of formatting from other data columns.  bcdo$DIS\_SAMPLE\_KEY\_VALUE=paste0  (bcdo$MISSION\_DESCRIPTOR,"\_",sprint  ("%03d",bcdo$EVENT\_COLLECTOR\_EVENT\_ID)  ,"\_",bcdo$DIS\_DETAIL\_COLLECTOR\_SAMP\_ID) |
| creating an empty dataframe  Some files were affected and would not run but some files were running through just fine | Read\_btl\_txtfile.m | Files without pressure data are unable to map onto correct depth bin, causing an error in IML script  Added a column zbouteille, following IML standards and matching it to BCD column, ‘START\_DEPTH’, ran files successfully through BCD2IML\_format.r and read\_btl\_txtfile.m |
| multiple BCD data types map to same IML code | Get\_flags\_IML2BCD\_EC.r | isolate BCD\_IML\_map to only the data types present in that particular mission  # load map file that matches IML fields to the BCD fields  map=read.csv("BCD\_IML\_map\_upd.csv", stringsAsFactors = FALSE)  #limit map to data types present in BCD to avoid confusion with non unique IML codes  #E chisholm Feb 2019  dtype <- unique(bcd$DATA\_TYPE\_METHOD)  tmap <- map[map$BCD\_FIELDS %in% dtype , ]  map <- tmap |
| Conversion of oxygen data in mmol/m\*\*3 to ml/l (unit which Oxygen QC is run in) | BCD2IML\_format.r | Added the DATA\_TYPE\_METHOD O2\_Winkler\_Molar to BCD files, changed O2\_Winkler data to O2\_Winkler\_molar if data maximum per cruise was greater than 14 (this could likely be refined using a visual inspection process)  O2\_Winkler\_Molar maps to IML code OXYM\_01 which is converted using factor of 44.66 to ml/l before being run through QC. \*\*\*see instructions  O2 molar variable and flag had to be added into B\_addQ2btl.m in order for OXYM flag to be created. New variable also created in data\_btl.m |
| S/T dimensions were not matching | B\_control\_Q.m | Removed duplicated PSAL codes from BCD\_IML\_map, so only one salinity value per IML file  Duplicated variables commonly cause problems with IML QC |
| Strmachi error | B\_control\_Q.m | Various matching/ dimensional errors solved by ensuring last data column is not CTD data. If so, move this data column so that all CTD columns are before labo columns |
| cruises did not appear to contain oxygen data | DATA | issue in the SQL code where if a cruise measured oxygen the entire mission was included but then data points were excluded based on geographic boundaries so some cruises were included who did not collect oxygen data within geographic boundaries  Cruises without oxygen data will also throw strange errors eg. Strmachi error in MatLab QC |
| where data frame dimensions were not matching | Get\_flags\_IML2BCD.r | Script could not find QC flag columns  Edit to line 81 to generalize formula for excluding lat, lon, date, time and not exclude any data or QC flag columns. |
| Test 3.6 nutrients  Index exceeds dimension | B\_control\_Q\_GL.m  B\_stage3\_Q.m | Solved by removing nutrient data columns (not key for processing oxygen data) – may need other solution if nutrients are focus and/or cannot remove |
|  |  |  |
|  |  |  |
|  |  |  |