**CBS-MEA: Rosette CTD Data**

**Prepared by A. Niemi**

**April 2023**

**Objective:** At each CBS-MEA station CTD profile data is collected is collected by the rosette-integrated CTD, or backup up instrument. The CTD profile data is QAQCed by an established protocol at IOS, with Bill Williams being the responsible Principle Investigator (PI) for the data. Information derived from the CTD data is central to the analysis of ecosystem data collected during CBS-MEA. In 2022/23 the Ecosystem Stressors team developed a wish list of summary and derived CTD variables required to 1) support annual reporting of ocean conditions and, 2) use as part of multivariate analyses supporting multiple research questions. A R script, to read and manipulate the CTD data, was developed and a R data frame was established to create a summary output file supporting these needs. The output list of summary and derived variables is intended to be evergreen, supporting new/revised needs of the research team.

This SOP is meant to centralize and streamline the FWI/Ecosystem Stressors processing of the CBS-MEA CTD data to more efficiently and comprehensively support the needs of the team, including data management, data requests, and archiving.

FWI R CTD data frame

* The R data frame was established in 2023 by Ryan Galley
* Ecosystem Stressors team will keep a working version of the R data frame for annual processing of the CBS-MEA CTD data
* Ryan will provide an accessible repository (e.g., GitHub) where updated versions of the data frame can be accessed/uploaded.

FWI CBS-MEA CTD data receiving:

* Each year the keeper/administrator of the R CTD data frame (possibly Lauren moving forward) will request the cleaned and verified CTD from that field season. An email should be sent to Paul Macoun at IOS, cc’ing Bill Williams. The data would not likely be available until January of the following year, given the time for chlorophyll and bottle salt analyses which are used in the QAQC process, and the follow-on data processing.
* Request only complete casts (surface to bottom; exclude casts from rosette deployments that sample only a portion of the water column, i.e. if the rosette is deployed twice at a deep station)
* The files will be received with downcast only
* If there were issues at sea with the rosette CTD, we need to request data for the specific stations where the backup CTD (likely SBE19plus moving forward) was used. The cruise report should identify stations where the backup CTD was used.

CTD data input at FWI: done by R data frame administrator (Lauren)

* Add the new year section, and individual stations as done for previous years.
* Verify with core team users if any new or different derived variables are needed as outputs.
* If the data frame needs to be updated with new code to address new/different requests, these changes should be communicated to Ryan.
* Request zooplankton tow metadata from zooplankton PI as these are needed for the zooplankton output requests (total distance towed and multi-net strata).
* Update the summary output Word file and document the rationale, and any calculations associated with new or changed outputs. The Word file contains a list of all the desired output variables generated from each CTD cast.
* Generate the csv output file, with all the summary variables.
* Update the info tabs in the excel output file to include the CTD Data Processing Notes for that year (from the .ctd file).
* Send the output file to Paul and Bill to see if they have any comments or concerns.
* Share the output file with core team members who have requested the data (i.e., IOS, Fish, Benthos, Zoop, Food web). Core team members must be in contact with Bill if using the data in a report or publication to verify how Bill and his team should be acknowledged as part of the work (e.g., co-author, acknowledgements).
* Provide 1) the output file, 2) individual CTD profile data in csv format, and 3) the .ctd files provided by IOS to Mark for archiving.

CTD data requests at FWI: Mark

* External/partner (not core team) CTD data requests should be directed to the Data Manager (Mark), not the R CTD data frame administrator.
* Collaborators requesting part, or all, of the CTD summary output file should fill out a data request form identifying the variables they want. (Is there a current data request form that can be modified for this?)
* Share the request with Paul and Bill for their approval and verify how they want to be acknowledged in the request (e.g., co-authors of a paper, or in acknowledgement section). They are the PIs for the data, not us.
* Once approved, share the Excel CTD summary output file with the requested variables.
* Update data request tracking file (data manager file).
* If the request is for profile data there are 2 options depending on what the user wants,

1) Provide the csv files with the cast data. For years 2012-2022 provide the files received for the creation of the R data frame (received 2022), not older version of the files.

2) Direct the requestors to the Water Properties website where they can access the .ctd cleaned data provided by IOS.