CTD Data Processing

Contributors to this document: Kate Yezhova, Janine Hunt, Pascal Guillot

This document discusses processing of the autonomous CTD dataset from the 2019 Southampton Island Marine Ecosystem Project (SIMEP). The procedure is based on what Janine and Kate were taught by Pascal, and the SBE Data Processing Manual (https://www.seabird.com/asset-get.download.jsa?id=55174002258). Page 20 of the manual outlines the steps for processing data.

Instrument: SBE 19plus V2 SeaCAT Profiler CTD SN7783 (Autonomous CTD)

Vessel: RV *William Kennedy* and its small boats

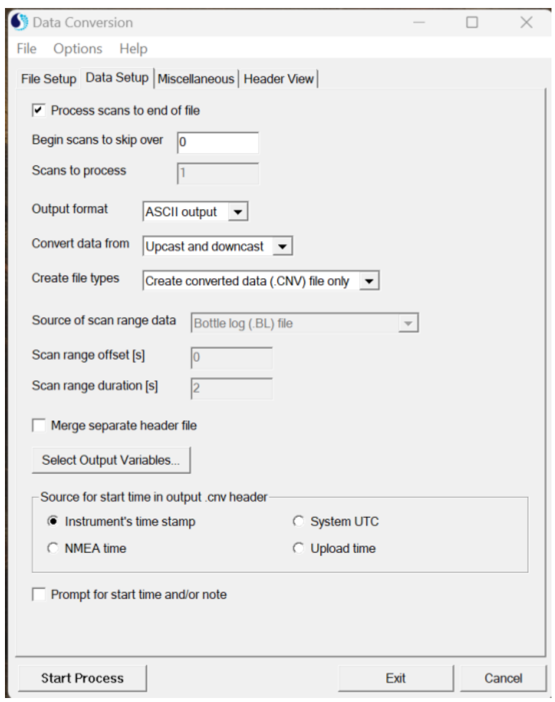
Cruise date: August 5-29, 2019

Spatial region: Hudson Bay

Notes: Proper CTD acclimation procedures were not followed during this cruise. The CTD pump only starts working once the minimum conductivity frequency is met and the pump delay elapses. For 157 out of 199 casts, due to improper acclimation procedures (not acclimating the CTD long enough, if at all), the pump did not have a chance to start working prior to the beginning of the downcast. Measurements collected prior to the pump turning on were excluded as most of the critical sensors (conductivity, temperature, oxygen) are pumped. When 25% or more of the downcast was lost for a cast, the upcast was used instead in the final files. Additionally, it does not appear WET Labs C-Star was functioning/set up properly during this cruise. Data was kept in the final file in case someone knows how to fix it.

The following steps were taken to process the data:

1. Create the following folder structure:
    a. 2019_wk_auto_ctd_sn7783
        i. logbooks
        ii. originals
        iii. r_scripts
        iv. seabird_psa_and_xmlcon
        v. data
            1. 00_raw
            2. 01_datacnv
            3. 02_section
            4. 03_filter
            5. 04_align
            6. 05_ctmass
            7. 06_loopedit
            8. 07_derive
            9. 08_binavg

        10. 09_split
        11. 10_final
2. Into the logbooks folder, place the ship logbook.
3. Into the originals folder, place all original data from the field (data files, logbooks, calibration files, etc.); zipped to prevent accidental modification.
4. Ensure CTD files all follow the same naming structure, and that the casts correspond to entries in the digital logbook.
5. The .xml file of the first cast was compared against the .xml file of the last cast, no differences in configuration or calibration values were found. The .xmlcon file was checked against calibration documents (except transmissometer calibration documents could not be found) to ensure all values were correct. The .xmlcon file was then checked against the first .xml file to ensure consistency; all checked out until volt channels, not sure where the .xml file takes the offset and slope values from for the volt channels.
6. Convert raw .hex files to .cnv files
    a. SBE Data processing → Run → Data Conversion (#1) → File Setup
        i. Open *01_DatCnv_SN7783_AUTOCTD.psa* file from the seabird_psa folder
        ii. Under Instrument configuration file, load the .xmlcon file in the "seabird_psa_and_xmlcon" folder
        iii. Under Input directory, select all .hex files from "00_raw" folder
        iv. Under Output directory, select "01_datacnv" folder
    b. … → Data Setup
        1. 

        2. Click Select Output Variables…, and choose the following:

a.

Seq. #	Variable Name [unit]
1	Scan Count
2	Descent Rate [m/s]
3	Pressure, Strain Gauge [db]
4	Depth [salt water, m]
5	Temperature [ITS-90, deg C]
6	Conductivity [mS/cm]
7	Salinity, Practical [PSU]
8	Density [sigma-theta, kg/m^3]
9	Specific Volume Anomaly [10^-8 * m^3/kg]
10	Oxygen raw, SBE 43 [V]
11	Oxygen, SBE 43 [umol/kg]
12	Oxygen, SBE 43 [ml/l]
13	Oxygen, SBE 43 [% saturation]
14	PAR/Irradiance, Biospherical/Licor [umol photons/m^2/sec]
15	Fluorescence, WET Labs CDOM [mg/m^3]

b.

Seq. #	Variable Name [unit]
16	Fluorescence, WET Labs ECO-AFL/FL [mg/m^3]
17	Beam Transmission, WET Labs C-Star [%]
18	Beam Attenuation, WET Labs C-Star [1/m]
19	Voltage 0
20	Voltage 1
21	Voltage 2
22	Voltage 3
23	RS-232 WET Labs raw counts 0
24	RS-232 WET Labs raw counts 1
25	RS-232 WET Labs raw counts 2
26	Frequency 0
27	Frequency 1
28	Frequency 2
29	Julian Days
30	

c. Click Start Process

7. Preparing an Excel sheet for taking notes
    a. Open the 01_datacnv folder
    b. Select all files (Ctrl+A)
    c. Hold shift, right click, select copy as path
    d. Go to Excel and paste (Ctrl+V)
    e. Select column A, go to Find & Select -> Replace -> Type out the beginning of the paths in "find what", and replace with blank. Type ".cnv" and replace with blank too.
    f. Add title row "cast id, start scan, end scan, notes"
    g. Save the file, titled "section.xlsx", into the logbooks folder
8. Plotting casts
    a. SBE Data Processing → Run → Sea Plot (#20) → File Setup
        i. For Program setup file, choose 02_SeaPlot_SN7783_AUTOCTD.psa

        ii. For Input directory, select all files in 01_datacnv folder
        iii. For Output directory, select any folder (the plots do not get automatically saved)
    b. → Plot Setup
        i. Title: datacnv
        ii. For variables, choose the following:
            1. y-axis: pressure
            2. x-axis 1: scan count
            3. hide other x-axes
    c. Click Start Process
    d. In the plot window select View → Show Cursor Position
    e. Record scan # of beginning of downcast (when the CTD comes back up to surface or begins a descend after acclimating at ~1-5m depth for a few minutes), and the end of the upcast (just before the CTD comes out of the water at the end) for each cast in the Excel sheet you created in the previous step.
        i. Almost all casts were acclimated at the surface and stayed at bottom for some time. For many, acclimation periods were very short. For some, there were no acclimation periods.
    f. It was anticipated that for many casts, due to improper acclimation procedures (not acclimating the CTD long enough, if at all), the pump would not have started working prior to the beginning of the downcast. The 02_pump_check.R script was partially run to get the scan counts at which the pump would have turned on for each cast (the pump only starts working once the minimum conductivity frequency is met and the pump delay elapses). The scan # of beginning of downcast was compared to the scan # at which the pump turned on to check whether the pump turned on prior to the start of the downcast.
        i. The pump did not have a chance to start working until the downcast was underway for most casts (157 out of 199). The "start scan" value in the section.xlsx spreadsheet was updated for the affected casts to exclude measurements collected prior to the pump turning on. The original "start scan" values can still be found in a new column, pump_check_notes, in the section.xlsx spreadsheet.
9. Cutting out soaking period (must go one file at a time)
    a. In SBE Data Processing: Run → Section (#16) → File Setup
        i. Program setup file: 03_Section_SN7783_AUTOCTD.psa
        ii. Input: one cast at a time from 01_datacnv. (Definitely not the most time efficient method but the simplest at this point, I am running out of time to test out different MATLAB/R scripts.)
        iii. Output: 02_section folder
    b. → Data Setup
        i. Section based on: scan count
        ii. Input minimum and maximum value for each cast and click Start Process, one cast at a time

10. Run the 01_section_check.R script to check that correct values were entered in the Section module and that none of the pressure/depth values ended up being negative (indicating measurements in the air).
11. Run the 02_pump_check.R script to ensure that the pump started working before the downcast began (the pump only starts working once the minimum conductivity frequency is met and the pump delay elapses).
    a. As mentioned above, this script was run prior to sectioning and the pump did not have a chance to start working until the downcast was underway for most casts (157 out of 199). The casts were sectioned to remove any periods where the pump was not working. Running this script again on the sectioned casts does not raise any flags.
12. The sectioned casts were visualized using Sea Plot to see where too much of the downcast was lost due to the pump not having a chance to start working until the downcast was underway (see above). In the section.xlsx spreadsheet, a new column was created titled "need_to_use_upcast", where values of n (no) or y (yes) were entered to indicate whether the upcast will have to be used instead of the downcast in the final file. When 25% or more of the downcast was lost for a cast, a value of y was entered in the "need_to_use_upcast" column.
13. Run the 03_sal_check.R to check the minimum conductivity measurements to ensure none of the samples were freshwater, as the processing steps are slightly different from seawater. SBE said in personal communication that the rough threshold for freshwater for data processing purposes is 0.6 S/m, i.e., 6 mS/cm.
    a. No casts were flagged.
14. Filtering
    a. SBE Data processing → Run → Filter (#2) → File Setup
        i. Program setup file: 04_Filter_SN7783_AUTOCTD.psa
        ii. Input directory: 02_section folder (all casts)
        iii. Output directory: 03_filter folder
    b. → Data Setup
        i. Low pass filter A, time constant (s): 1.0
        ii. Low pass filter B, time constant (s): 0.5
        iii. Specify Filters…
            1. Clear all
            2. Pressure, Strain Gauge (db): Low pass filter A
            3. Temperature (ITS-90, deg C): Low pass filter B
            4. Conductivity (mS/cm): Low pass filter B
    c. Click Start Process

15. Align CTD (advance parameters in time relative to pressure)
    a. SBE Data processing → Run → Align CTD (#3) → File Setup
        i. Program setup file: 05_Align_SN7783_AUTOCTD.psa
        ii. Input directory: 03_filter folder (all casts)
        iii. Output directory: 04_align folder
    b. → Data Setup → Enter Advance Values
        i. Clear all
        ii. Temperature (ITS-90, deg C): +0.5 seconds

            1. This is the recommended value for SBE19plusV2 in the data processing manual
    iii. Conductivity (mS/cm): +0.5 seconds

            2. Note that the manual gives contradicting statements. First statement is: "For an SBE 19plus or 19plus V2 with a standard 2000-rpm pump, do not advance conductivity." Second statement is: "If temperature is advanced relative to pressure and you do not want to change the relative timing of temperature and conductivity, you must add the same advance to conductivity."
            3. Pascal applies a +0.5 second advance to both temperature and conductivity, Pascal's method will be followed.

    iv. Oxygen raw, SBE43 (V): no advance
            4. The data processing manual suggests +3 to 7 seconds for an SBE19Plus
            5. Janine noted that this only works if the Oxygen raw, SBE43 (V) variable is being aligned. Janine tried several delays and the 0s delay seemed best for 2021 data.
            6. Pascal noted that at the beginning, he tried to estimate the right correction. You need to remove the gap between the downcast and upcast because of the long sensor response time. When you are going to apply a correction, you shift all oxygen values X seconds below their original place. Therefore, if you are moving at a speed of 1m/s, a +5 second shift would shift all values 5 m below where they were recorded. For oceanic waters with little variation, this could be okay. But for Arctic waters with chlorophyll maxima and oxygen peaks, this could create a big shift between these two events (SCM and O_2 peak). In Pascal's opinion, it is scientifically incorrect to create a gap between SCM and O_2 peak. Pascal either does not apply an oxygen correction, or he applies a 0.5s correction (same as for temperature and conductivity).
  c. Click Start Process

16. Cell Thermal Mass

  a. As per the data processing manual, "Perform conductivity cell thermal mass correction if salinity accuracy of better than 0.01 PSU is desired in regions with steep gradients. Note: do not use Cell Thermal Mass for freshwater data."
  b. SBE Data processing → Run → Cell Thermal Mass (#4) → File Setup
    i. Program setup file: 06_CTMass_SN7783_AUTOCTD.psa
    ii. Input directory: 04_align folder (all casts)
    iii. Output directory: 05_ctmass folder
  c. → Data setup → Correct primary conductivity values
    i. Thermal anomaly amplitude (alpha): 0.04
    ii. Thermal anomaly time constant (1/beta) = 8.0
  d. Click Start Process

17. Loop Edit (flags scans with very low and backward velocity)
  a. SBE Data processing → Run → Loop Edit (#5) → File Setup

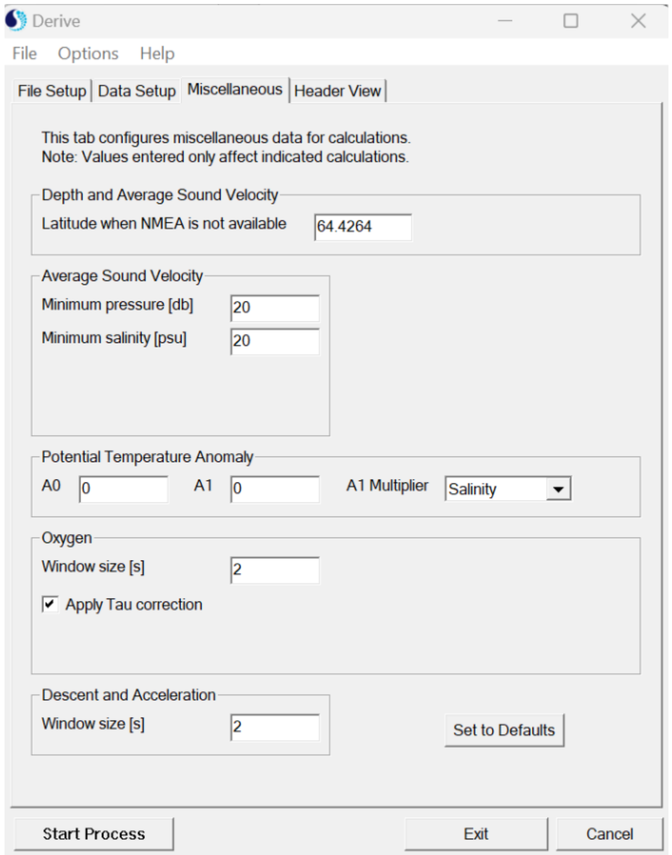
        i. Program setup file: 07_LoopEdit_SN7783_AUTOCTD.psa
        ii. Input directory: 05_ctmass folder (all casts)
        iii. Output directory: 06_loopedit folder
    b. → Data Setup
        i. Minimum velocity type: Fixed minimum velocity
        ii. Minimum CTD velocity (m/s): 0.05. Note that Pascal recommended using velocity < 0.1 m/s (as opposed to the SBE recommended 0.25 m/s).
        iii. Uncheck "Remove surface soak"
        iv. Check "Exclude scans marked bad"
    c. Click Start Process

18. Derive (computes thermodynamic properties based on EOS-80 (practical salinity))
    a. SBE Data processing → Run → Derive (#6) → File Setup
        i. Program setup file: 08_Derive_SN7783_AUTOCTD.psa
        ii. Instrument configuration file: 19-7783 CTD w transmissometer.xmlcon
        iii. Input directory: 06_loopedit folder (all casts)
        iv. Output directory: 07_derive folder
    b. → Data Setup → Select Derived Variables
        i.

Select Derived Variables

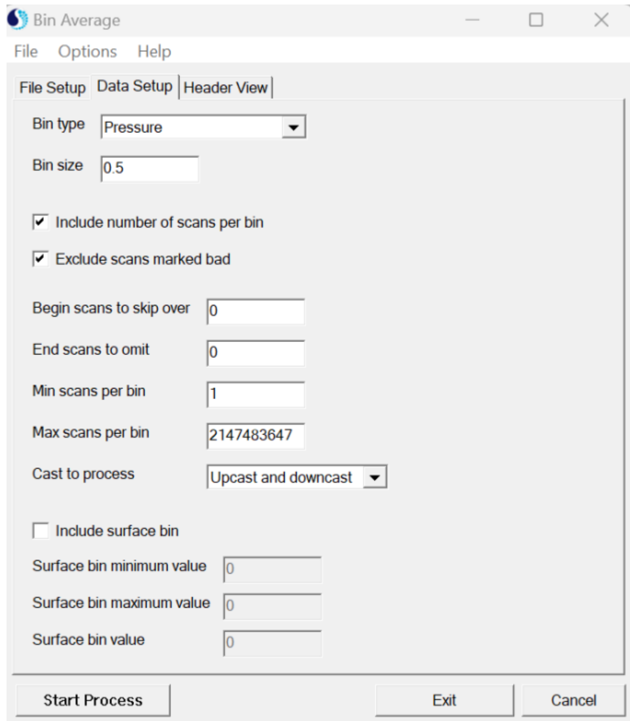
Seq. #	Variable Name [unit]
1	Density [density, kg/m^3]
2	Density [sigma-theta, kg/m^3]
3	Depth [salt water, m]
4	Oxygen, SBE 43 [ml/l]
5	Oxygen, SBE 43 [umol/kg]
6	Oxygen, SBE 43 [% saturation]
7	Potential Temperature [ITS-90, deg C]
8	Salinity, Practical [PSU]
9	Specific Volume Anomaly [10^-8 * m^3/kg]

    c. → Miscellaneous
        i. Latitude when NMEA is not available: average starting latitude of all casts (64.4264)



d. Click Start Process
e. Optional step that was not done: Use Derive TEOS-10 (absolute salinity) module to derive variables based on TEOS-10.

19. Bin Average
    a. SBE Data processing → Run → Bin Average (#8) → File Setup
        i. Program setup file: 09_BinAvg_SN7783_AUTOCTD.psa
        ii. Input directory: 07_derive folder (all casts)
        iii. Output directory: 08_binavg folder
    b. → Data Setup
        i. Bin type: Pressure
        ii. Bin size = 0.5

iii. 

c. Click Start Process

20. Split (splitting the downcast from upcast)
    a. SBE Data processing → Run → Split (#17) → File Setup
        i. Program setup file: 10_Split_SN7783_AUTOCTD.psa
        ii. Input directory: 08_binavg folder (all casts)
        iii. Output directory: 09_split folder
    b. → Data Setup
        i. Output files: upcast and downcast (it will rename each file for downcast with a "d" and upcast with a "u" in front of the file name)
        ii. Check "Exclude scans marked bad"
    c. Click Start Process
21. Merging with logbook (using R)
    a. Run the 04_final_file.R script to merge CTD data with the logbook and output Excel and ODV files.
        i. A file named "merge_key.xlsx" was created and saved in the logbooks folder. This file lists which cast filenames correspond to which entries in the logbook, determined by the CTD start times.
            1. A few errors in the digitization of the logbook were found in the time columns for the CTD entries, so the scanned copy of the original logbook was cross-referenced when matching up casts to entries in the logbook. The digitized logbook was not updated as the errors were insignificant for the final output (times noted in the logbook are only used to match to casts, not used for anything else). There were also some errors in the time columns for CTD entries in the physical logbook itself – for some of them, the correct times were obvious given the context (previous and following entries, or other information for that cast (beginning, bottom, end times)), but some the situation was less clear:

a. Times appear to be written down wrong for CTD157 in the original and digitized logbooks. Time BE: 1:13, BO: 1:18, EN: 00:23 in original logbook, 1:23 in digitized logbook. Time zone unknown but implied CDT. Likely an error was made on the ship converting from UTC (some lab monitors) or Eastern (ship time) to CTD. Cast SBE19plus_01907783_2019_08_11_0010_0157.cnv times are BE: 7:14 UTC, BO: 7:19 UTC, EN: 7:23 UTC. The minutes and duration of the downcast/upcast match, so this cast and entry in the logbook will be matched.

b. Times appear to be written down wrong for CTD199 in the original and digitized logbooks. Time BE: 11:30, BO: 11:30, EN: 11:31 (time zone unknown but implied CDT). It is likely an error was made on the ship converting from UTC (some lab monitors) or Eastern (ship time) to CTD. Cast SBE19plus_01907783_2019_08_11_0010_0201.cnv remains the last unmatched cast, and its times are BE: 15:32 UTC, BO: 15:32 UTC, EN: 15:33 UTC. The minutes and duration of the downcast/upcast match, so this cast and entry in the logbook will be matched.

2. Two CTDs were labeled as CTD163 in the logbook, the second occurrence will be changed to CTD163B in the R script for matching.

ii. CTD87 (SBE19plus_01907783_2019_08_11_0010_0087.cnv) was cancelled, it is possible the ship started moving while the CTD was still in the water. The cast will be excluded from the final file.

iii. CTD102 (SBE19plus_01907783_2019_08_11_0010_0102.cnv) was cancelled, scanned copy of the original logbook has a note “get away from ice”, it is likely the ship started moving while the CTD was still in the water. The cast will be excluded from the final file.

iv. It does not appear WET Labs C-Star was functioning/set up properly during this cruise. Data was kept in the file in case someone knows how to fix it.