**CTD Data Processing**

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

This document discusses processing of the rosette 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 SN7798 in an SBE 32 Carousel Water Sampler SN32-1173 (Rosette CTD)

Vessel: RV *William Kennedy* and its small boats

Cruise date: August 5-29, 2019

Spatial region: Hudson Bay

Notes:

* Many errors were found in the processed rosette bottle files produced previously in 2019. Users are discouraged from using those files.
* SPAR was disconnected after the first two casts.
* Station 7 Cast 2 did not have an acclimation period, and the CTD pump was likely not working when the first five bottles were fired.
* It was indicated in rosette logs that NMEA data stream was connected to PC after the first two casts, however, the raw data has no records of NMEA coordinates or NMEA time. Coordinates from the logbook were used in the final downcast files here.
* The 11\_pre\_final folder contains processed files that were not altered in any way besides following the regular processing steps; in other cases, these files would be considered final. However, several issues were found in the rosette bottle files after comparing them to the rosette paper logs (e.g., a file registered 5 bottles firing at 10m depth whereas the paper log indicates only 4 bottles were fired at that depth, or a file registered 14 bottles firing while the rosette only has 12, etc.). The 14\_final folder contains the rosette bottle file where the issues were resolved based on educated guesses and assumptions detailed below; however, field notes were lacking, facts remain unclear, and the assumptions could be wrong. Issues, solutions, and notes from comparison of rosette paper logs vs. bottle files are listed below:
  + Station 3 Cast 1: paper log shows 5 bottles were fired (85,85,60,60,60m), bottle file shows one more bottle was fired at 60m. Nonissue, no fix needed.
  + Station 5 Cast 1: paper log shows times in AM (2AM), bottle file times were the same but in PM (2PM). Nonissue, checked K.Yezhova’s logbook to confirm times, bottle files are correct, no fix needed.
  + Station 7 Cast 1: paper log recorded times in a wrong timezone. Nonissue, checked K.Yezhova’s logbook to confirm times, bottle files are correct, no fix needed.
  + Station 11 Cast 1: Rosette paper log has wrong times recorded (every 16:… time should be 18:…). Nonissue, bottle files are correct, no fix needed.
  + Station 11 Cast 2: paper log recorded times in a wrong time zone. Nonissue, checked K.Yezhova’s logbook to confirm times, bottle files are correct, no fix needed.
  + Station 12 Cast 1: Paper log indicates all 12 bottles were fired, but the bottle file only has record of 11 bottles fired (appears that one of the two bottles fired at 60m depth did not register in the software). No notes in logbooks about rosette coming back with an empty bottle. Assuming paper log is correct, adding +1 to bottle position number for bottles fired after bottle 3. Copying info from bottle 3 for the missing bottle 4.
  + Station 12 Cast 2: Paper log indicates 5 bottles were fired (10,10,0,0,0m), bottle file shows 6 bottles were fired with one extra at 10m (10,10,10,0,0,0). Assuming it was a last-minute decision by the rosette operator that did not get properly recorded in the paper log (paper log was missing info, did not have the times bottles closed recorded, so it was likely the target depths were written down before the cast began and not updated afterwards). Assuming bottle file is correct, no fix needed.
  + Station 14 Cast 2: Paper log indicates that bottles 10 and 11 were fired out of order, but this was not observed in the bottle file. It is unknown why the previous technician believed the bottles fired out of order. Nonissue, no fix needed.
  + Station 15 Cast 2: Paper log indicates all 12 bottles were fired, bottle file only has 4 (software seems to have stopped recording after the first bottle fired at 30m). Additionally, it appears that one of the four bottles fired at 40m depth did not register in the software. There were no notes about any empty bottles. Assuming paper log is correct, adding +1 to bottle position number for the bottle fired after bottle 3, copying info from bottle 3 for the missing bottle 4, copying info from bottle 5 (30m) for the missing bottle 6, and pulling data for bottles 7-12 from the upcast.
  + Station 20 Cast 2: paper log indicates 10 bottles were fired, bottle file indicates 11 bottles were fired (with one extra at 20m). Paper log is filled out fully, detailed, no indication of change of plans / firing of an additional bottle. K.Yezhova’s logbook agrees with the paper log. Assuming paper log is correct, deleting the third bottle fired at 20m and subtracting 1 from bottle position number for bottles 8-11.
  + Station 23 Cast 1: paper log has wrong date (Aug 24, should be Aug 25). Paper log indicates 4 bottles were fired (first 2 at bottom, following 2 at 60m), bottle file shows all 12 bottles were fired (first 2 at bottom, the rest at 60m). Probably a last-minute decision by the rosette operator that did not get recorded in the paper log. Nonissue, no fix needed.
  + Station 23 Cast 2: paper log has wrong date (Aug 24, should be Aug 25). Paper log indicates 12 bottles were fired, bottle file indicates 13 bottles were fired (with one extra at 10m). Rosette only has 12 bottles. Unable to tell which record is truly erroneous. K.Yezhova’s logbook indicates that bottle 11 was at 0m (agrees with paper log, not rosette file). Assuming paper log is correct, deleting the fifth bottle fired at 10m and subtracting 1 from bottle position number for bottles 12-13.
  + Station 26 Cast 1: paper log indicates 12 bottles were fired, bottle file registered 14 bottles (with extras at 30m and 10m). Unable to tell which records are truly erroneous. K.Yezhova’s logbook agrees with the paper log. Assuming paper log is correct, deleting the third bottles fired at 30m and 10m, and adjusting bottle numbers accordingly.

The following steps were taken to process the data:

1. Create the following folder structure:
   1. 2019\_wk\_ros\_ctd\_sn7798
      1. logbooks
      2. originals
      3. r\_scripts
      4. seabird\_psa\_and\_xmlcon
      5. 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\_bottlesum
         12. 11\_pre\_final
         13. 12\_stn3
         14. 13\_stn15\_cast2
         15. 14\_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. All .xmlcon files in the raw data were checked. Based on looking at .xmlcon files, there was a change in configuration after station 10, the SPAR sensor was removed. The .xmlcon files were identical for these earlier stations-casts: 3-1, 3-2, 5-1, 6-1, 7-1, 7-2, 8-1, 9-1, 9-2, 10-1. After the configuration change, the remaining .xmlcon files were identical for the later stations-casts.

The main .xmlcon file (non-cast specific) aligns with the earlier cast-specific .xmlcon files with SPAR added.

The main .xmlcon file was checked against calibration documents (except SPAR calibration documents could not be found) to ensure all values were correct. For SPAR, the units, conversion factor, and ratio multiplier were compared against the 2023 .xmlcon file, and the values were very similar. The main .xmlcon file was also compared to the one used in 2018; they were identical.

In rosette log scans (scans of rosette sheets filled out by hand by the rosette operator during rosette operations), it is noted on station-cast 5-1 that “casts 1 & 2 of stn 3 did not use NMEA to PC in XMLCON. Will be added for all further casts”. The addition of NMEA was not reflected in any cast-specific .xmlcon files, and in the next step, when the data conversion module was run, there were scan length errors for cast 5-1 onwards, confirming that the present .xmlcon files are not compatible with the raw data. Upon further investigation, it was found that SPAR was removed from the rosette configuration following station-cast 3-2, not 10-1 as suggested by cast-specific .xmlcon files.

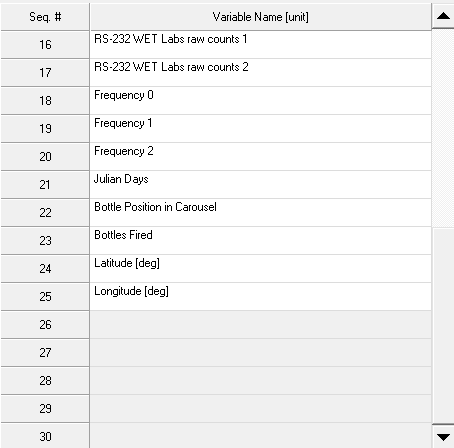
The main .xmlcon file was zipped and saved to the originals folder. A copy of the main .xmlcon file was modified to include NMEA and exclude SPAR. Station-casts 3-1 and 3-2 will be processed separately using a separate .xmlcon file with SPAR.

1. Convert raw .hex files to .cnv files, and .bl files to .ros files
   1. SBE Data processing 🡪 Run 🡪 Data Conversion (#1) 🡪 File Setup
      1. Open *01\_DatCnv\_SN7798\_ROSCTD.psa* file from the seabird\_psa folder
      2. Under Instrument configuration file, load “Edited 2019 configuration file.XMLCON”
      3. Under Input directory, select all .hex files in “00\_raw” folder
      4. Under Output directory, select “01\_datacnv” folder
   2. … 🡪 Data Setup. The chosen scan range offset and duration mean the software will extract scans from the 5 seconds before each bottle is fired.

A screenshot of a computer screen

Description automatically generated

* + 1. Click Select Output Variables…, and choose the following:
       - 1. A screenshot of a computer

            Description automatically generated
         2. 
  1. Click Start Process

1. Preparing an Excel sheet for taking notes
   1. Open the 01\_datacnv folder
   2. Under Type, choose “CNV File” to display only the .cnv files
   3. Select all files (Ctrl+A)
   4. Right click, select copy as path
   5. Go to Excel and paste (Ctrl+V)
   6. Select column A, go to Find & Select -> Replace -> Type out the beginning of the paths (folder names) in “find what”, and replace with blank. Type “.cnv” and replace with blank too.
   7. Add title row “cast id, start scan, end scan, notes”
   8. Save the file, titled “section.xlsx”, into the logbooks folder
2. Plotting casts
   1. SBE Data Processing 🡪 Run 🡪 Sea Plot (#20) 🡪 File Setup
      1. For Program setup file, choose 02\_SeaPlot\_SN7798\_ROSCTD.psa
      2. For Input directory, select all files in 01\_datacnv folder
      3. For Output directory, select any folder (the plots do not get automatically saved)
   2. 🡪 Plot Setup
      1. Title: datacnv
      2. For variables, choose the following:
         1. y-axis: pressure
         2. x-axis 1: scan count
         3. hide other x-axes
   3. Click Start Process
   4. In the plot window select View 🡪 Show Cursor Position
   5. Record scan # of beginning of downcast (when the CTD begins a descend after acclimating at ~1-5m depth for some time), and the end of the upcast (just before the CTD comes out of water at the end) for each cast in the Excel sheet you created in the previous step.
      1. Will not be proceeding with processing the following casts:
         1. 7-2 (no acclimation period, rosette was lowered to ~2m, held there for a bit, bottles were fired, then rosette was brought back on deck). The .ros file for this cast will be processed further, but the .cnv file will not be.
         2. 18-2, cancelled cast.
      2. Station-cast 14-1 was “cancelled” due to the rosette line getting caught on ice, but the downcast is still usable, so the cast will be processed further.
3. Cutting out soaking period (must go one file at a time)
   1. In SBE Data Processing: Run 🡪 Section (#16) 🡪 File Setup
      1. Program setup file: 03\_Section\_SN7798\_ROSCTD.psa
      2. Input: one cast at a time from 01\_datacnv. (Definitely not the most time efficient method but the simplest at this point.)
      3. Output: 02\_section folder
   2. 🡪 Data Setup
      1. Section based on: scan count
      2. Input minimum and maximum value for each cast and click Start Process, one cast at a time
4. 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).
5. Typically, a pump check R script would be run at this point to ensure that the pump started working before the downcast began (the pump typically only starts working once the minimum conductivity frequency is met and the pump delay elapses). However, there is no information in .hex or .hdr files regarding pump delay. Possibly there was no pump delay, and the pump started working immediately upon the SBE19Plus being turned on. Only station-cast 7-2 did not have an acclimation period (and was not sectioned), all other casts had a good acclimation period.
   1. A test pump check (test\_pump\_check.R) was run with minimum conductivity frequency of 3257, pump delay of 60 seconds, and sample rate of 0.25 (4 samples/second), all casts passed (downcast began after the pump started working), except 7-2.
   2. Station-cast 7-2 did not have an acclimation period, and based on the test pump check, the pump did not start working until scan #587. This means that the pump was only working when bottle 6 was fired. For bottles 1-5, the pump was likely not yet working.
6. Run the 02\_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.
   1. No casts were flagged.
7. Filtering
   1. SBE Data processing 🡪 Run 🡪 Filter (#2) 🡪 File Setup
      1. Program setup file: 04\_Filter\_SN7798\_ROSCTD.psa
      2. Input directory: 02\_section folder (all casts)
      3. Output directory: 03\_filter folder
   2. 🡪 Data Setup
      1. Low pass filter A, time constant (s): 1.0
      2. Low pass filter B, time constant (s): 0.5
      3. 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
   3. Click Start Process
8. Align CTD (advance parameters in time relative to pressure)
9. SBE Data processing 🡪 Run 🡪 Align CTD (#3) 🡪 File Setup
   * 1. Program setup file: 05\_Align\_SN7798\_ROSCTD.psa
     2. Input directory: 03\_filter folder (all casts)
     3. Output directory: 04\_align folder
10. 🡪 Data Setup 🡪 Enter Advance Values
11. Clear all
12. Temperature (ITS-90, deg C): +0.5 seconds
    * + 1. This is the recommended value for SBE19plusV2 in the data processing manual
13. Conductivity (mS/cm): +0.5 seconds
14. 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.”
15. Pascal applies a +0.5 second advance to both temperature and conductivity, Pascal’s method will be followed.
16. Oxygen raw, SBE43 (V): no advance
17. The data processing manual suggests +3 to 7 seconds for an SBE19Plus
18. 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.
19. 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 O2 peak). In Pascal’s opinion, it is scientifically incorrect to create a gap between SCM and O2 peak. Pascal either does not apply an oxygen correction, or he applies a 0.5s correction (same as for temperature and conductivity).
20. Click Start Process
21. Cell Thermal Mass
22. 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.”
23. SBE Data processing 🡪 Run 🡪 Cell Thermal Mass (#4) 🡪 File Setup
24. Program setup file: 06\_CTMass\_SN7798\_ROSCTD.psa
25. Input directory: 04\_align folder (all casts)
26. Output directory: 05\_ctmass folder
27. 🡪 Data setup 🡪 Correct primary conductivity values
28. Thermal anomaly amplitude (alpha): 0.04
29. Thermal anomaly time constant (1/beta) = 8.0
30. Click Start Process
31. Loop Edit (flags scans with very low and backward velocity)
32. SBE Data processing 🡪 Run 🡪 Loop Edit (#5) 🡪 File Setup
33. Program setup file: 07\_LoopEdit\_SN7798\_ROSCTD.psa
34. Input directory: 05\_ctmass folder (all casts)
35. Output directory: 06\_loopedit folder
36. 🡪 Data Setup
37. Minimum velocity type: Fixed minimum velocity
38. 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).
39. Uncheck “Remove surface soak”
40. Check “Exclude scans marked bad”
41. Click Start Process
42. Derive (computes thermodynamic properties based on EOS-80 (practical salinity))
43. SBE Data processing 🡪 Run 🡪 Derive (#6) 🡪 File Setup
44. Program setup file: 08\_Derive\_SN7798\_ROSCTD.psa
45. Instrument configuration file: Edited 2019 configuration file.XMLCON
46. Input directory: 06\_loopedit folder (all casts)
47. Output directory: 07\_derive folder
48. 🡪 Data Setup 🡪 Select Derived Variables
49. A screenshot of a computer

    Description automatically generated
50. 🡪 Miscellaneous
51. Latitude when NMEA is not available: average starting latitude of all casts present in the merge\_key.xlsx (64.4718)
    1. When processing Station 3 casts separately, latitude of 63.5214 was used (average starting latitude of both Station 3 casts as seen in the merge\_key.xlsx).

A screenshot of a computer

Description automatically generated

1. Click Start Process
2. Optional step that was not done: Use Derive TEOS-10 (absolute salinity) module to derive variables based on TEOS-10.
3. Bin Average
4. SBE Data processing 🡪 Run 🡪 Bin Average (#8) 🡪 File Setup
5. Program setup file: 09\_BinAvg\_SN7798\_ROSCTD.psa
6. Input directory: 07\_derive folder (all casts)
7. Output directory: 08\_binavg folder
8. 🡪 Data Setup
9. Bin type: Pressure
10. Bin size = 0.5
11. A screenshot of a computer screen

    Description automatically generated
12. Click Start Process
13. Split (splitting the downcast from upcast)
14. SBE Data processing 🡪 Run 🡪 Split (#17) 🡪 File Setup
15. Program setup file: 10\_Split\_SN7798\_ROSCTD.psa
16. Input directory: 08\_binavg folder (all casts)
17. Output directory: 09\_split folder
18. 🡪 Data Setup
19. 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)
20. Check “Exclude scans marked bad”
21. Click Start Process
22. Follow the same steps to process Station 3 casts in the “12\_stn3” folder.
    1. 3-1 scan range for sectioning: 1795-4505
    2. 3-2 scan range for sectioning: 1548-5484
23. Merging with logbook (using R)
24. Run the 03\_pre\_final\_file.R script to merge CTD data with the logbook and output Excel and ODV files.
25. 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
26. It was found that none of the casts actually had NMEA data recorded. NMEA latitude and longitude was 0 for all casts; something must have been wrong with how things were configured on the ship. Logbook coordinates will be used.
27. Create bottle files
    1. SBE Data Processing 🡪 Run 🡪 Bottle Summary (#9) 🡪 File Setup
       1. Program setup file: 11\_BottleSum\_SN7798\_ROSCTD.psa
       2. Instrument configuration file: Edited 2019 configuration file.XMLCON
       3. Input directory: select all .ros files in 01\_datacnv folder
       4. Output directory: 10\_bottlesum folder
    2. 🡪 Data Setup
       1. Check Output min/max values for averaged variables
          1. A screenshot of a computer

             Description automatically generated
       2. Under Select Averaged Variables, Select All, then uncheck Latitude and Longitude given that NMEA data was not actually recorded (see above).
       3. Under Select Derived Variables, select the following:
          1. A screenshot of a computer

             Description automatically generated
       4. Start Process
       5. Repeat for Station 3 using its configuration file
28. Create the pre-final bottle data file
    1. Run the 04\_read\_bottle\_files.R script to output a user-friendly rosette bottle data file into the 11\_pre\_final folder.
29. To get the missing bottle data for depths 20m, 10m, and 0m for Station 15 Cast 2 (software did not record these bottles firing for unknown reasons), the following steps were done:
    1. The .bl, .hdr, and .hex files were copied from 00\_raw to 13\_stn15\_cast2 folder. By looking at the .bl file, it was observed that bottles that fired without issues (1 and 2) were 12 scans in length. The .cnv file from the 01\_datacnv folder was plotted to visualize this cast, and it was observed where on the flat parts of the upcast (rosette stopped at sampling depths) the bottles for which we do have bottle data were fired. Similar points on the flat parts of the upcast at 20m, 10m, and 0m depths were chosen. Timestamp of the end scan was used, taken from the .cnv file from the 01\_datacnv folder. The following lines replaced the previous contents of the .bl file in the 13\_stn15\_cast2 folder:

*7, 7, Aug 18 2019 13:55:47, 5880, 5892*

*9, 9, Aug 18 2019 13:58:01, 6415, 6427*

*11, 11, Aug 18 2019 14:00:06, 6915, 6927*

* 1. Steps 6 and 22 were repeated for Station 15 Cast 2 using the modified .bl file, with all output files being saved to the 13\_stn15\_cast2 folder. At step 6.b, under Create file types, “Create bottle (.ROS) file only” was chosen.

1. Errors were found in the bottle files. The errors and solutions to the errors are described in the Notes section at the beginning of this document. To correct the errors, and to attach Station 15 Cast 2 bottle file data obtained from the previous step, run the 05\_error\_correction.R script to output final, corrected files into the 14\_final folder.