Links:

* SBE data processing manual:

<https://www.seabird.com/asset-get.download.jsa?code=251446>

* Sea Bird data processing download (Seasoft V2):

[Sea-Bird Scientific Software - Sea-Bird Scientific (seabird.com)](https://software.seabird.com/)

* Sea Bird community toolkit for python:

*Seabird are currently updating the UI on their software, alongside this they are actively releasing and documenting a bunch of stuff regarding calling their data proc via python. We should look at whether this should be used instead of the current calling method which was written by an external scientific institute (hakai) prior. \*Noting here that the documentation I have seen here mentions lack of support for the instrument models we use.*

[Sea-BirdScientific/seabirdscientific: The Sea-Bird Scientific Community Toolkit (github.com)](https://github.com/Sea-BirdScientific/seabirdscientific)

* Files in:

WQS001.hex – Data file which comes off the instrument

4409\_20020701.xmlcon – instrument configuration file, outlines which sensor packages and calibration coefficients are tied to this CTD at the time of the cast, used to convert .hex data to a user readable .cnv file.

8 x \*.psa files – project setup files for each data processing step, these house information relevant to the processing parameters. These are used to “process” the .cnv file along a string of .cnv steps in this order: Convert – Filter – Align – CellTM – Loop edit – Wild edit – Derive – Bin Av

* Basic current workflow:

Instrument config files and PSA files are located by serial number and date in the config folder – these are updated once per year when CTD is serviced by manufacturer. Raw hex file is placed in the raw folder by user. When process button is clicked, an external database with deployment details is searched for the filename, to gather latitude metadata, if no database is connected or file not found the user is prompted for manual latitude entry. The hex file header info is scraped for a deployment date and a CTD Serial number. These two values are then used to search the config folder for the correct xmlcon file, and respective PSA files. Using these, the file then runs through all 8 processing steps, each time appending the name with a step value, and subsequently using that file for the next step. Once complete, the files used are gathered in a standardised folder structure, and move to a processed area.

* Notes

Database link is not included in github, lets put this to the side for now and manually enter -19 lat.

Current code limitations/todo:

* Button functions:
  + Path choice buttons do not function. Paths need to be defined manually in the config.py file
  + Stop and cancel buttons do not function correctly. – *cancel button hopefully fixed line 352 if derive\_latitude is not None:. This is working to skip the file but creating a layout issue issue where the define button on the main page below cancel is also clicked at the same time.*
* Assorted
  + ~~Change processed folder to “processing” - complete~~
  + ~~Change completed to “processed” – complete~~
  + ~~Code hangs if folders are already in processed / processing folders. This should likely throw an error instead (or perform a different function). – complete this now skips the file. Oddly the code hangs if starts with skipping a file, but not if file is skipped midway. Gets to print("CTD Serial Number:", ctd\_id) line 401 then hangs. – fixed, extra parsing added for cast\_date~~
  + Add more error prompts –
    - PSA file unreadable
* File streams
  + ~~Currently .psa files are found in config folder, used, then copied into cast folder in “~~*~~processing~~*~~”, then moved to completed folder at the end. I would like these to be copied into the cast folder in “~~*~~processing~~*~~”, used from here, then moved to “~~*~~processed~~*~~”. This is so that if we adjust the .psa file parameters during the proposed QAQC steps the change is only applied to this one cast, and then carries through with folder structure.~~
    - ~~Line 475~~
    - ~~# Create instance of SBE functions with config\_path files~~
    - ~~Place folder setup (def\_initsetup) prior to this step - complete~~
    - ~~Reroute SBE paths from config\_folder to “cwd, CONFIG["PROCESSING\_PATH"] + "./" + file\_name, xmlcon\_file” or similar - complete~~
  + Folder to only move from processing to processed if full processing is done and user has confirmed happy with the data. Install prompt prior to moving files.
    - Change so that .hex file is moved from raw folder, not copied (once above is done)
  + Future addition for processed folder to have a further layer for sampling trip number (processed/Trip\_8010/wqs001…)
    - This would require a call to the oceanDB to find what trip number cast was on.
* Processing step selection
  + Build in system where if hex file in raw folder already has a folder in “processing”,
    - Prompt user to select which dataproc steps to complete with tick boxes. Default to be all.
    - Dataproc steps selected will pick up from processing folder and replace downstream files from step selected rather than starting from raw.
    - Further notes below with user report
* Data report
  + After bin\_av step, code to pause and give user metrics on the bin\_av file.
    - Start with are there any negative values. – scrape # name and # span from file. User can just eyeball these, nice to have negative values highlighted if possible but not requirement.
    - Scan range for each bin – more complex to do. Requires table creation from derive step data file, stripping out badflag rows, then pulling max and min scan number for max and min depth value within depth ranges (0.5m-1.5m, 1.5m-2.5m ect).
    - At this point it would be good to see a plot of the data as well
    - More metrics potentially to follow
  + If user unsatisfied with report, user to adjust .psa files
    - This could be done manually or in GUI. Decision to be made.
    - File potentially to be removed from processing folder if user clicks no to satisfactory test
      * Issue with this is that .psa file needs to be adjusted, and psa files are a general use file until they are copied into processing folder and used. This is an argument to leave files in processing folder, and allow the option for reprocessing once .psa files have been changed.
    - Current proposed idea would be that files stay in processing/filename folder. User to manually edit psa files from this folder, and re-run steps as required. Code to recognise that folder & PSA files exist in processing folder, and to replace only files downstream of the indicated start point (ie if user selects re-run from CellTM, the cellTM appended file is the new start point, and loopedit, wildedit ect files (files only, not psa files) are overwritten.
* Version control log
  + Script to continually update a version control log,
    - File in
    - Config file used
    - Dataproc steps run
    - Parameters used? Scrape from cnv file with line search
    - See appendix 1 below
* There is a train of thought to have a system where the code gathers a bunch of information prior to starting any processing like Roxanas did. It would look at all the files in the “waiting to be processed” folder, and cross reference with oracle database looking for metadata, as well as in the completed folder to see if they had already been done. Once metadata was entered in the database, it would then process only these casts. – more to come
* Heavily document each step throughout code
  + JM is fine navigating, altering and writing small amounts of python code, but not power-user enough to create this whole thing properly. A point to note is that over time things like config file headers, instrument models ect change. We want this script created in a way that does not require us to bug you for little changes that seabird make or simple adjustments like naming convention changes ect.

Code data to scrape for log

* log xmlcon file
* log File\_name
* log Lat and Lon from DB
* log logreq from DB

Appendix 1, data to scrape for log from any cnv file. Cnv files all have the log that we require as they step through, but we want it all in a central location.

\* Sea-Bird SBE19plus Data File:

\* FileName = D:\WQ\field\CTD\_files\2021\_04\_7603\_Aqr\_19Plus\WQQ323.hex

\* Software Version 1.59

**\* Temperature SN = 4525**

**\* Conductivity SN = 4525**

\* System UpLoad Time = Apr 13 2021 15:44:23

\*\* WQQ323

\* ds

\* SeacatPlus V 1.4D SERIAL NO. 4525 12 Apr 2021 21:43:25

\* vbatt = 12.9, vlith = 8.3, ioper = 61.8 ma, ipump = 132.8 ma,

\* iext01 = 14.7 ma, iext23 = 66.9 ma,

\* status = not logging

\* number of scans to average = 1

\* samples = 31695, free = 409810, casts = 15

\* mode = profile, minimum cond freq = 3137, pump delay = 40 sec

\* autorun = no, ignore magnetic switch = no

\* battery type = ALKALINE, battery cutoff = 7.5 volts

\* pressure sensor = strain gauge, range = 1450.0

\* SBE 38 = no, Gas Tension Device = no

\* Ext Volt 0 = yes, Ext Volt 1 = yes, Ext Volt 2 = yes, Ext Volt 3 = yes

\* echo commands = yes

\* output format = raw HEX

\* S>

\*

\* SeacatPlus V 1.4D SERIAL NO. 4525 12 Apr 2021 21:43:38

\* temperature: 23-oct-20

\* TA0 = 1.270134e-03

\* TA1 = 2.604666e-04

\* TA2 = 3.978012e-07

\* TA3 = 1.377431e-07

\* TOFFSET = 0.000000e+00

\* conductivity: 23-oct-20

\* G = -9.738302e-01

\* H = 1.403505e-01

\* I = -2.267605e-04

\* J = 3.822297e-05

\* CF0 = 2.637367e+03

\* CPCOR = -9.570000e-08

\* CTCOR = 3.250000e-06

\* CSLOPE = 1.000000e+00

\* pressure S/N = 3770, range = 1450 psia: 20-oct-20

\* PA0 = 6.297490e-01

\* PA1 = 4.435756e-03

\* PA2 = -1.994071e-11

\* PTCA0 = 5.217988e+05

\* PTCA1 = 6.595602e+00

\* PTCA2 = -3.583078e-02

\* PTCB0 = 2.487100e+01

\* PTCB1 = -9.999999e-04

\* PTCB2 = 0.000000e+00

\* PTEMPA0 = -5.420420e+01

\* PTEMPA1 = 5.714442e+01

\* PTEMPA2 = -6.930979e-01

\* POFFSET = 0.000000e+00

\* volt 0: offset = -4.684667e-02, slope = 1.248835e+00

\* volt 1: offset = -4.714667e-02, slope = 1.249249e+00

\* volt 2: offset = -4.669333e-02, slope = 1.249405e+00

\* volt 3: offset = -4.679667e-02, slope = 1.249395e+00

\* EXTFREQSF = 1.000005e+00

\* dh

\* cast 7 **07 Apr 2021 00:22:36** samples 19989 to 21220, avg = 1, stop = mag switch

\* S>

# nquan = 14

# nvalues = 13

# units = specified

# name 0 = prdM: Pressure, Strain Gauge [db]

# name 1 = tv290C: Temperature [ITS-90, deg C]

# name 2 = c0S/m: Conductivity [S/m]

# name 3 = sbeox0V: Oxygen raw, SBE 43 [V]

# name 4 = flECO-AFL: Fluorescence, WET Labs ECO-AFL/FL [mg/m^3]

# name 5 = CStarTr0: Beam Transmission, WET Labs C-Star [%]

# name 6 = par/sat/log: PAR/Logarithmic, Satlantic

# name 7 = scan: Scan Count

# name 8 = timeS: Time, Elapsed [seconds]

# name 9 = depSM: Depth [salt water, m], lat = -16.92

# name 10 = sal00: Salinity, Practical [PSU]

# name 11 = sbox0Mm/Kg: Oxygen, SBE 43 [umol/kg], WS = 2

# name 12 = nbin: number of scans per bin

# name 13 = flag: flag

# span 0 = 2.012, 14.083

# span 1 = 27.7623, 27.8229

# span 2 = 5.374402, 5.441685

# span 3 = 2.8356, 2.8815

# span 4 = 0.1510, 0.4135

# span 5 = 67.2525, 83.4491

# span 6 = 3.4698e+01, 1.0726e+03

# span 7 = 850, 1019

# span 8 = 212.177, 254.542

# span 9 = 2.000, 14.000

# span 10 = 33.3897, 33.8920

# span 11 = 188.316, 192.315

# span 12 = 1, 15

# span 13 = 0.0000e+00, 0.0000e+00

**# interval = meters: 1**

# start\_time = Apr 07 2021 00:22:36 [Instrument's time stamp, header]

# bad\_flag = -9.990e-29

**# <Sensors count="7" >**

# <sensor Channel="1" >

# <!-- Count, Temperature -->

**# <TemperatureSensor SensorID="58" >**

**# <SerialNumber>4525</SerialNumber>**

**# <CalibrationDate>23-Oct-20</CalibrationDate>**

# <A0>1.27013365e-003</A0>

# <A1>2.60466637e-004</A1>

# <A2>3.97801157e-007</A2>

# <A3>1.37743110e-007</A3>

# <Slope>1.00000000</Slope>

# <Offset>0.0000</Offset>

# </TemperatureSensor>

# </sensor>

# <sensor Channel="2" >

# <!-- Frequency 0, Conductivity -->

**# <ConductivitySensor SensorID="3" >**

**# <SerialNumber>4525</SerialNumber>**

**# <CalibrationDate>23-Oct-20</CalibrationDate>**

# <UseG\_J>1</UseG\_J>

# <!-- Cell const and series R are applicable only for wide range sensors. -->

# <SeriesR>0.0000</SeriesR>

# <CellConst>2000.0000</CellConst>

# <ConductivityType>0</ConductivityType>

# <Coefficients equation="0" >

# <A>0.00000000e+000</A>

# <B>0.00000000e+000</B>

# <C>0.00000000e+000</C>

# <D>0.00000000e+000</D>

# <M>0.0</M>

# <CPcor>-9.57000000e-008</CPcor>

# </Coefficients>

# <Coefficients equation="1" >

# <G>-9.73830214e-001</G>

# <H>1.40350510e-001</H>

# <I>-2.26760497e-004</I>

# <J>3.82229666e-005</J>

# <CPcor>-9.57000000e-008</CPcor>

# <CTcor>3.2500e-006</CTcor>

# <!-- WBOTC not applicable unless ConductivityType = 1. -->

# <WBOTC>0.00000000e+000</WBOTC>

# </Coefficients>

# <Slope>1.00000000</Slope>

# <Offset>0.00000</Offset>

# </ConductivitySensor>

# </sensor>

# <sensor Channel="3" >

# <!-- Count, Pressure, Strain Gauge -->

**# <PressureSensor SensorID="46" >**

**# <SerialNumber>4525</SerialNumber>**

**# <CalibrationDate>20-Oct-20</CalibrationDate>**

# <PA0>6.29749028e-001</PA0>

# <PA1>4.43575603e-003</PA1>

# <PA2>-1.99407101e-011</PA2>

# <PTEMPA0>-5.42042030e+001</PTEMPA0>

# <PTEMPA1>5.71444200e+001</PTEMPA1>

# <PTEMPA2>-6.93097775e-001</PTEMPA2>

# <PTCA0>5.21798758e+005</PTCA0>

# <PTCA1>6.59560232e+000</PTCA1>

# <PTCA2>-3.58307768e-002</PTCA2>

# <PTCB0>2.48710000e+001</PTCB0>

# <PTCB1>-1.00000000e-003</PTCB1>

# <PTCB2>0.00000000e+000</PTCB2>

# <Offset>0.000000</Offset>

# </PressureSensor>

# </sensor>

# <sensor Channel="4" >

# <!-- A/D voltage 0, Oxygen, SBE 43 -->

**# <OxygenSensor SensorID="38" >**

**# <SerialNumber>0288</SerialNumber>**

**# <CalibrationDate>29-Dec-20</CalibrationDate>**

# <Use2007Equation>1</Use2007Equation>

# <CalibrationCoefficients equation="0" >

# <!-- Coefficients for Owens-Millard equation. -->

# <Boc>0.0000</Boc>

# <Soc>0.0000e+000</Soc>

# <offset>0.0000</offset>

# <Pcor>0.00e+000</Pcor>

# <Tcor>0.0000</Tcor>

# <Tau>0.0</Tau>

# </CalibrationCoefficients>

# <CalibrationCoefficients equation="1" >

# <!-- Coefficients for Sea-Bird equation - SBE calibration in 2007 and later. -->

# <Soc>4.3672e-001</Soc>

# <offset>-0.5877</offset>

# <A>-4.2196e-003</A>

# <B> 1.7967e-004</B>

# <C>-2.7672e-006</C>

# <D0> 2.5826e+000</D0>

# <D1> 1.92634e-004</D1>

# <D2>-4.64803e-002</D2>

# <E> 3.6000e-002</E>

# <Tau20> 1.8500</Tau20>

# <H1>-3.3000e-002</H1>

# <H2> 5.0000e+003</H2>

# <H3> 1.4500e+003</H3>

# </CalibrationCoefficients>

# </OxygenSensor>

# </sensor>

# <sensor Channel="5" >

# <!-- A/D voltage 1, PAR/Logarithmic, Satlantic -->

**# <PARLog\_SatlanticSensor SensorID="76" >**

**# <SerialNumber>223</SerialNumber>**

**# <CalibrationDate>14-Nov-2020</CalibrationDate>**

# <a0>1.4503e+000</a0>

# <a1>8.8369e-001</a1>

# <Im>1.3589e+000</Im>

# <Multiplier>1.0000e+000</Multiplier>

# </PARLog\_SatlanticSensor>

# </sensor>

# <sensor Channel="6" >

# <!-- A/D voltage 2, Fluorometer, WET Labs ECO-AFL/FL -->

**# <FluoroWetlabECO\_AFL\_FL\_Sensor SensorID="20" >**

**# <SerialNumber>FLRT-108</SerialNumber>**

**# <CalibrationDate>28-Dec-2020</CalibrationDate>**

# <ScaleFactor>5.00000000e+000</ScaleFactor>

# <!-- Dark output -->

# <Vblank>0.0660</Vblank>

# </FluoroWetlabECO\_AFL\_FL\_Sensor>

# </sensor>

# <sensor Channel="7" >

# <!-- A/D voltage 3, Transmissometer, WET Labs C-Star -->

**# <WET\_LabsCStar SensorID="71" >**

**# <SerialNumber>CST-711PR</SerialNumber>**

**# <CalibrationDate>30-Dec-2020</CalibrationDate>**

# <M>21.5970</M>

# <B>-0.0430</B>

# <PathLength>0.250</PathLength>

# </WET\_LabsCStar>

# </sensor>

# </Sensors>

# datcnv\_date = Apr 03 2023 16:36:05, 7.26.7.129 [datcnv\_vars = 9]

**# datcnv\_in = \\mac\home\\_DATA\DATA\_AIMS\CTD\_Data\CTD\_DataProcessing\PROCESSING\202104\_7603\_Aqr\_SBE19plus\_4525\raw\wqq323.hex \\mac\home\\_DATA\DATA\_AIMS\CTD\_Data\CTD\_DataProcessing\PROCESSING\202104\_7603\_Aqr\_SBE19plus\_4525\conf\SBE19plus\_4525\_20201002.xmlcon**

# datcnv\_skipover = 0

# datcnv\_ox\_hysteresis\_correction = yes

# filter\_date = Apr 03 2023 16:39:54, 7.26.7.129

**# filter\_in = \\mac\home\\_DATA\DATA\_AIMS\CTD\_Data\CTD\_DataProcessing\PROCESSING\202104\_7603\_Aqr\_SBE19plus\_4525\edit\wqq323.cnv**

**# filter\_low\_pass\_tc\_A = 0.500**

**# filter\_low\_pass\_tc\_B = 1.000**

**# filter\_low\_pass\_A\_vars = tv290C c0S/m**

**# filter\_low\_pass\_B\_vars = prdM**

# alignctd\_date = Apr 03 2023 17:19:24, 7.26.7.129

**# alignctd\_in = \\mac\home\\_DATA\DATA\_AIMS\CTD\_Data\CTD\_DataProcessing\PROCESSING\202104\_7603\_Aqr\_SBE19plus\_4525\edit\wqq323\_filter.cnv**

# alignctd\_adv = tv290C 0.500, sbeox0V 2.500

# celltm\_date = Apr 03 2023 17:21:54, 7.26.7.129

**# celltm\_in = \\mac\home\\_DATA\DATA\_AIMS\CTD\_Data\CTD\_DataProcessing\PROCESSING\202104\_7603\_Aqr\_SBE19plus\_4525\edit\wqq323\_filter\_align.cnv**

**# celltm\_alpha = 0.0400**, 0.0000

**# celltm\_tau = 8.0000**, 0.0000

# celltm\_temp\_sensor\_use\_for\_cond = primary,

# loopedit\_date = Apr 03 2023 17:24:42, 7.26.7.129

**# loopedit\_in = \\mac\home\\_DATA\DATA\_AIMS\CTD\_Data\CTD\_DataProcessing\PROCESSING\202104\_7603\_Aqr\_SBE19plus\_4525\edit\wqq323\_filter\_align\_celltm.cnv**

# **loopedit\_minVelocity = 0.250**

**# loopedit\_surfaceSoak: minDepth = 2.0, maxDepth = 5, useDeckPress = 1**

**# loopedit\_excl\_bad\_scans = yes**

# wildedit\_date = Apr 03 2023 17:27:17, 7.26.7.129

# wildedit\_in = **\\mac\home\\_DATA\DATA\_AIMS\CTD\_Data\CTD\_DataProcessing\PROCESSING\202104\_7603\_Aqr\_SBE19plus\_4525\edit\wqq323\_filter\_align\_celltm\_loop.cnv**

# wildedit\_pass1\_nstd = 2.0

# wildedit\_pass2\_nstd = 20.0

# wildedit\_pass2\_mindelta = 0.000e+000

# wildedit\_npoint = 100

# wildedit\_vars = prdM tv290C c0S/m sbeox0V flECO-AFL CStarTr0 par/sat/log

# wildedit\_excl\_bad\_scans = yes

# Derive\_date = Apr 03 2023 17:29:52, 7.26.7.129 [derive\_vars = 3]

**# Derive\_in = \\mac\home\\_DATA\DATA\_AIMS\CTD\_Data\CTD\_DataProcessing\PROCESSING\202104\_7603\_Aqr\_SBE19plus\_4525\edit\wqq323\_filter\_align\_celltm\_loop\_wild.cnv \\mac\home\\_DATA\DATA\_AIMS\CTD\_Data\CTD\_DataProcessing\PROCESSING\202104\_7603\_Aqr\_SBE19plus\_4525\conf\SBE19plus\_4525\_20201002.xmlcon**

# derive\_time\_window\_docdt = seconds: 2

# derive\_ox\_tau\_correction = no

# binavg\_date = Apr 03 2023 17:32:41, 7.26.7.129

**# binavg\_in = \\mac\home\\_DATA\DATA\_AIMS\CTD\_Data\CTD\_DataProcessing\PROCESSING\202104\_7603\_Aqr\_SBE19plus\_4525\edit\wqq323\_filter\_align\_celltm\_loop\_wild\_derive.cnv**

# **binavg\_bintype = meters**

**# binavg\_binsize = 1**

**# binavg\_excl\_bad\_scans = yes**

# binavg\_skipover = 0

# binavg\_omit = 0

# binavg\_min\_scans\_bin = 1

# binavg\_max\_scans\_bin = 2147483647

# binavg\_surface\_bin = no, min = 0.000, max = 0.000, value = 0.000

# file\_type = ascii

\*END\*

2.012 27.8229 5.374402 2.8399 0.1510 81.0280 1.0726e+03 850 212.177 2.000 33.3897 188.777 5 0.0000e+00

3.018 27.7968 5.397007 2.8356 0.2154 67.2525 6.7780e+02 863 215.605 3.000 33.5661 188.316 14 0.0000e+00

4.024 27.7802 5.408875 2.8448 0.2855 67.8901 4.6998e+02 877 218.934 4.000 33.6607 189.051 12 0.0000e+00

5.030 27.7743 5.409251 2.8513 0.3137 69.8421 3.2783e+02 889 221.960 5.000 33.6672 189.626 12 0.0000e+00

6.036 27.7711 5.408903 2.8556 0.3271 71.0086 2.4227e+02 902 225.156 6.000 33.6668 190.022 14 0.0000e+00

7.042 27.7686 5.408802 2.8616 0.3686 71.2318 1.7595e+02 916 228.637 7.000 33.6676 190.552 9 0.0000e+00

8.048 27.7677 5.408793 2.8676 0.4000 74.4500 1.3783e+02 931 232.601 8.000 33.6679 191.085 11 0.0000e+00

9.054 27.7668 5.408856 2.8734 0.4135 75.7667 1.0542e+02 947 236.415 9.000 33.6687 191.588 15 0.0000e+00

10.060 27.7644 5.409267 2.8742 0.3795 78.8478 8.3838e+01 962 240.135 10.000 33.6729 191.683 13 0.0000e+00

11.066 27.7633 5.410310 2.8815 0.3650 80.1401 6.6568e+01 976 243.716 11.000 33.6808 192.315 13 0.0000e+00

12.072 27.7623 5.416283 2.8786 0.3397 83.4491 5.3430e+01 990 247.240 12.000 33.7230 192.045 10 0.0000e+00

13.078 27.7683 5.428583 2.8638 0.3211 80.5461 4.4446e+01 1005 251.039 13.000 33.8045 190.709 14 0.0000e+00

14.083 27.7740 5.441685 2.8493 0.3654 70.0896 3.4698e+01 1019 254.542 14.000 33.8920 189.394 1 0.0000e+00