**QC & processing – IPHC profile data files**

**(2014 steps from raw files to distribution)**

* SeaSoft-SeaBird: Eric runs. Output \*.cnv files
* cnv-to-NetCDF: Shaun converted \*.cnv to NetCDF
* create 4-up plots: run plotplus routines from DKachel
  + plot4up\_epic\_TsalSigTpH.ppc & plot4up\_epic\_TsalSigTpH
  + convert pplus output to ps or gif, then to pdf. Combine in pdf file.
* View/QC of 4-up plots:
  + 1st: sal, temp, pH, Sig-T. 2nd: O2, Chl, temp.
  + Markings on pdfs on ipad. Hand-written notes per profile during QC
* Hand edit changes to profile NetCDF files (from /epic).
  + Ncdump \*.nc files to cdl format (\*.nc -> \*\_ed1.cdl).
  + Matlab to view points, while hand-edit of cdl files (linux-vi or Notepad++).
  + Missing vals in nc files = 1.0e+35. Missing vals in csv/txt file = 999999
  + Look at Lauri’s metadata notes before/when editing
  + Remove Sig-T from CDL file if temp or salinity were edited.
* Document QC [notes to tab-5 in xlsx master doc]:
  + Note patterns - pH cals applied or not. Fast descent, O2 too high, etc.
  + Note where var (sal, etc) is deleted.
* Convert cdl files to nc: Ncgen edited cdl files to nc (\_ed1.nc; ncgen –o \*.nc \*.cdl)
* Re-calculate Sigma-T: using EPIC:ctdvar (70). \*\_ed1.nc => \*\_ed1.var
  + ptr file for input to ctdvar (all done at once).
  + Rename \*ed1.var to \*ed2\_var.nc (see linux ‘rename’ cmd below)
  + Also, DK version my\_ctdvar.x (env setup precarious, precompiled for ?, not dependable)
  + Could be done in Matlab or Ferret?
* SigT cleanup. Replace “Infinityf” (missing data in ST\_70 generated from ctdvar tool)
  + ncdump \*\_ed2\_var.nc to \*\_ed3.cdl
  + convert values ‘no-data’ values (Infinityf, inf) to 1.0e+35 (see 1-line ‘rename’below)
  + convert cdl (ed2 phase) back to NetCDF (\*\_final.nc)
* rename \*ed3 to \*\_final.nc, move to final/ area, **verify (view) edits** (via ferret)
* convert filenames: nc to IPHC (perl & shell script)
  + Create Cast-Station file (2 col) from file list file:

sortFnames2Info.p ls\_fileList.txt > \*\_CastStn.txt

* + perl to make Script to rename files: renameFinalNC\_forIPHC\_2013.p Arg is cast-stn file. Uses nc files in current dir.
  + Run output script: renameFinalNCfiles\_IPHC.sh
  + IPHC filename format: IPHC2014\_shp\_stnn\_cNNN.csv (shp=3-ltr ship; stnn=4-digit station; NNN= 3-digit cast or ‘set’ number.
* Create csv file:
  + run EPIC:ctdlst to create txt files with IPHC filenames
  + Create shell script by hand to run EPIC:ctdlst (-> nc2txt\_ctdlst.sh). See below.
  + Reformat txt files: convert\_ctdlst2IPHC.p (see below)
  + Final check: vars ok, no-data values uniformly 999999 (not 1e+35, 1e+36,or NaN).

FINAL shell files:

renameFinalNCfiles\_IPHC.sh [from ../../renameFinalNC\_forIPHC\_2013.p \*CastStn.txt]

example: mv penomm036\_ed3.nc IPHC2013\_pen\_3052\_c036.nc

nc2txt\_ctdlst.sh [9-line input to EPIC:ctdlst for each \*.nc file]

runConvert\_lst2final.sh

../../../../convert\_ctdlst2IPHC.p IPHC2013\_pen\_3043\_c112\_nc.lst > IPHC2013\_pen\_3043\_c112.csv

**From Dave’s Processing:**

Plots created for QC

1) sal, temp, Sigma-T, pH

2) temp, Chl, O2, O2 saturation (%)

**Linux dir structure (suggested)**:

cnv/ - files from SeaBird

nc/ - orig nc files

edit/ - nc with added O2 from epic/, then all levels of editing. Filenames \*\_ed1, \*\_ed2, etc

plots/ - 4-up plots for QC

final/ - mv \*ed3.nc to here, then rename for IPHC, and create csv for IPHC.

**Useful Command-line Tools**:

rename \_ed1.var \_ed2\_var.nc \*\_ed1.var [change fnames from ed1.var to ed2\_var.nc]

Grep Infinityf \*\_ed3.cdl (if any of these, sub with next perl command)

perl –pi –e ‘s/Infinityf/1.0E+35/g’ \*\_ed3.cdl

perl –pi –e ‘s/0\.10000E\+36/999999/g’ \*\_ed3.cdl

**CODE**:

\*\*renameFinalNC\_forIPHC\_2013.p (\*\_ed3.nc to \*.nc)

[ at /home/pearl/Sullivan/projects/IPHC/2013/dataProcDKachel/iphc\_2013/ ]

Create shell script to rename nc files (arg \*\_CastStn.txt)

\*\_CastStn.txt input is 2-column list of cast station

Output script name: renameFinalNCfiles\_IPHC.sh

\*\*nc2txt\_ctdlst.sh

Create shell script to run EPIC:ctdlst on all NetCDF files for conversion to ascii

Output script name: nc2txt\_ctdlst.sh

[ls-1 all files, then hand-edit to add all input for ctdlst input is:

Filename.nc

Y

N

<< 3 carriage returns here >>

D

9 ]

\*\*convert\_ctdlst2IPHC.p

[ at /home/pearl/Sullivan/projects/IPHC/2013 ]

\*\*sortFnames2Info.p nnnnnn\_fileList.txt (list of orig. cnv files)

[ at /home/pearl/Sullivan/projects/IPHC/ ]

Output: nnnnnn\_castStn.txt (2-col list of cast station)

for use in renaming final files to IPHC required naming.

**Phyllis’ comments on data quality [more to be added here]**:

Sigma-T

pH –

* generally not wiggly. Smoother.
* Often all good or all bad.
* get Phyllis’ chart for expected pH at depth.

Inversion – general cause is instrument sent down too fast, too slow and/or coupled with rough water and heaving.

Temperature –

* data problems almost never T.
* Reliable.
* Thermistor slower adjustment to change in temp than conductivity cell.
* If no pump….

pH – values should be between 7 – 9 (low 8-ish good)

**for-the-moment STEPS:**

hand edits done

convert back to NetCDF (ed1)

calc Sigma-T (if needed)

dump to cdl to correct missing data value (ed2)

convert back to NetCDF final (ed3 = final)

rename final NetCDF (with cast,stn)

dump final-name NetCDF to txt (ctdlst)

correct Missing data values in \*txt

convert \*.txt to \*.csv

tar groups of final \*.nc & \*.csv files

gzip files

move to web & post

**FileName format, incoming** **files**:

PCS064606911

PCS = vessel, 0646 = station number; 069 = set/cast; 11 = year;

Vessel chars + station number (xxxx) + cast number (xxx) + year (xx)