Reporting Decimal Line.Sta Numbers

Parent Category: References (/references.html)
Category: Methods (/about-calcofi/methods.html)
Last Updated: 20 July 2018

Historically, when designing the original CalCOFI Line.Stations pattern, the transect lines were plotted perpendicular to the coastline with .0 decimal accuracy. Since reporting lines and stations with decimal accuracy was unnessary at the time, line and station numbers were rounded to whole numbers to save data column space. Computer cards had a limited number of character columns such as 128 columns in 1983 when the IEH ascii format was adopted. Data storage was also limited.

With the addition of 9 SCCOOS stations **in 2004**, **it became necessary to report the Line & Sta numbers with .0 decimal accuracy to resolve stations**. For example, to resolve sta 93.3 26.7 from SCCOOS sta 93.4 26.4, we have to report the decimal line.sta numbers in all the data. Since SCCOOS station data have been integrated into the CalCOFI time series, decimal line & station numbering was applied to the entire time series, even before 2004. Lines & stations of older data on ERDDAP or other biological datasets may not be reported to .0.
Line Number Reporting (Pre-2004 = Current): 93 = 93.3, 90 = 90.0, 87 = 86.7, 83 = 83.3, 80 = 80.0, 77 = 76.7, 73 = 73.3, 70 = 70.0, 67 = 66.7, 63 = 63.3, 60 = 60.0
And between line stations like Santa Barbara Basin: 82.47 = 81.8 46.9. So station designation has changed from 90.120 to 90.0 120.0, 93.27 to 93.3 26.7. It should be noted that the Line.sta calculator & algorithm (http://calcofi.org/field-work/station-positions/calcofi-line-sta-algorithm.html) will derive decimal, not integer, line & station numbers from latitude and longitude.

Differences in CalCOFI Line.Sta numbering in the CalCOFI time-series have been discussed especially when relating historical net tow data to hydrographic data.
When comparing time-series datasets, there are times when there are no matches between biological (zooplankton, ichthyoplankton) and physical (temperature, salinity, oxygen) data. Some data may not have matches, but other data may not match because the criteria used miss data, such as CalCOFI Line and Station numbers reported as integer vs decimal..
Here are some CalCOFI practices that may cause data to not match:

- early (1950's) CalCOFI net tows were often done between stations ie nets were tow as the boat transited between stations. So using specific latitude-longitude or CalCOFI Line.Sta as criteria may not match data.
- there are often net tows but no hydro (bottle or CTD) cast & vice versa. On some early cruises, only one 10m bottle was collected on station. On cruises such as in 1979, net tows were done but no bottle casts. NOAA fisheries performs other surveys in the CalCOFI area where net tows are performed, samples collected, & data added to their dataset without matching physical data. CalCOFI cruise reference: http://calcofi.org/field-work/station-info/20-survey-coverage.html
- If the ocean conditions are bad, station operations may only include a bottle or CTD cast and no net tows or vice-versa.
- During the 70's CalCOFI cruises were every three years so there are gaps in the time-series.

Finding criteria to match data should be easier from 1983 to present since cruise & data practices were improved.
To improve matching biological to physical data, especially when including data earlier than 1983, it is recommended to use criteria ranges instead of exact numbers when possible. Date-time windows of 2-4 hours or lat-lon ranges that are within ~2nm of each other. Two nautical miles is still CalCOFI's "on-station" criteria.

One option that may improve data matches in older data is to use *Rpt_Line* & *Rpt_Sta* in the hydrographic dataset for CalCOFI Line & Station since these may be rounded.
Another option is to exclude SCCOOS station data when querying rounded (to integer) CalCOFI Line & Station data.