**Bay Study: Datasheet Proofing**

**Initial Data Entry and QC by Lab Staff**

Lab staff completes pre-proofing tasks after each survey, which includes:

* Depth conversion: converts station depths from feet to meters
* Complete shrimp log
* Enter station and tow data into Field Entry database
* Check station and tow data through line-by-line by printing Access queries “Data entry check for Station” and “Data entry check for Tow”
* Edits data in the Field Entry database as appropriate

**First Proof by an Environmental Scientist**

In the Field Entry database, use the following queries and tables to proof field data:

*X = query written and tested*

*xQry Meter by date and time*

* In Design View, assign survey
* Check date/time/tow chronology
* Check that previous End Meter is the same as next tow’s Start Meter
* For discrepancies, check datasheets for comments or entry errors

*xQry Meters Small or Large*

* In Design View, assign survey
* Shows stations whose meter is <14000 or >30000
* Look at tide, distance, datasheet comments to determine if meter is reasonable
* Replace MWT meters with station average when:
  + Debris on flowmeter, or other noted problems
  + Meter is missing
* If OT meter is not reasonable or valid, delete (meter is not used to calculate OT CPUE)

*xQry Distance High or Low*

* In Design View, assign survey
* Shows stations whose tow distance is outside of the normal range
* Look at tide, direction, distance of other tow at the station, etc. to help determine where the error lies
* Replace OT distance with station average when:
  + Error cannot be resolved
  + Distance is missing
* If MWT distance is not reasonable or valid, delete (distance is not used to calculate MWT CPUE)

(uses calc dist query)

*xQry Distance Outliers*

* In Design View, assign survey
* Shows stations when the difference between calculated and field recorded distance is >0.04 nm
* Often, error is with the coordinates
  + Lat and Long numbers were swapped
  + Some digits are more commonly misrecorded, ex. 4s and 7s, 3s and 8s
* Database cannot calculate distance when degrees change
  + Ex. Station 345: 3759.715 🡪 3800.357
  + Use an online coordinate distance calculator to confirm the field distance, such as <http://boulter.com/gps/distance/#more>

(uses calc dist query)

*xQry* *Bearing Outliers*

* In Design View, assign survey
* Shows stations when the difference between calculated and field recorded bearing is > 45°
* Look at map, tow direction, tides, etc. to help determine which bearing seems more logical
* Often, calculated bearing is more logical

*xTide Bearing Crosstab*

* Run the query twice, each time with 2 lines of criteria:
  + Net = 2, Tide = 1 OR Net = 1, Tide = 2
  + Net = 1, Tide = 1 OR Net = 2, Tide = 2
* Compare station bearing to previous surveys
* Look at map, tow direction, tides, etc. when bearing seems too high/low
* Change to calculated bearing if it’s more logical than the field bearing
* Note: Tows on a slack tide are not included in the crosstab

*x**Depth Crosstab*

* Compare station depth to previous surveys
* If a depth range is reported on the datasheet, use the average
* If a depth seems out of the normal range, cross check against the Seabird log
  + If both are out of range, can use the average of datasheet and Seabird log depths
* If depth is missing, use depth from Seabird log if reasonable

*xSubstrate Crosstab*

* Compare station substrate to previous surveys
* If substrate is missing, look at nearby stations or previous surveys
* Leave blank if not obvious

*xSecchi Crosstab*

* Compare secchi to nearby stations
* If secchi is missing, use average of nearby stations of similar depth (shoals or channels)

After running all queries and making appropriate edits, review each datasheet for:

* Correct tow and catch codes
* Correct fish and crab alpha codes – highlight all crab codes
* Make sure subsample data and expansions are correct
* Add plus counts
* Make sure all fish that were returned for ID have been verified and logged
* Process all fish and crabs found in shrimp samples (use a red pencil to amend the datasheet)
* Legibility – writing should be neat and clear
* Scan for completeness and address any blank fields

**Second Proof**

A second ES also reviews the queries listed above to ensure all questionable data has been identified and addressed. They will scan each datasheet for completeness, paying particular attention to:

*Are species per net reasonable?*

* Is the species normally found at that station or time of year?
* Remove any anchovy or jellyfish data from otter trawl catches

*Are species lengths reasonable?*

* Are lengths too large or small for that species or time of year?
* Remove lengths that are less than their minimum length cutoff

*Do net, tide, direction and bearing agree?*

* Use station map and tide book to confirm

*Same species on front and back of datasheet?*

* Combine lengths if the same species is on front and back
* Designate a separate age class with S code “9” if necessary
* Ensure all subsampled species are on the back of datasheet

**Using the Field Sheet Proofing List**

Add your initials to the checklist as tasks are completed.

Log all edits made, and include both the original and edited data. Also include an explanation if necessary. Ex. “Bearing: 85 🡪 137, calculated bearing more reasonable” or “Secchi: blank 🡪 85, averaged from stations 320 and 322”

Log any edits to fish data in the boxes at the bottom of the page. This includes length or alpha code changes, combining species from front and back, or removal of undersized lengths.

Also log any entry errors so that there is a record of change to the Field Entry database.

Add your initials to the left column “Updated in Access” when edits are made in the Field Entry database.

Make sure all edits have been made in the appropriate places – database and/or datasheets.