Steps to run discard for CAMS

SQL

1. Merge CAMS apportionment and Trip Attributes
2. Generate Observer tables of species discards by calendar year
3. Prorate observer discards
4. Combine Catch and Observer data
5. Generate support tables (once)

R

1. Define time period (e.g. calendar year)
2. Define set of species; estimation period should commonly be based on time period
   1. Import data
3. Define stratification variables (STRATA)
4. Loop over species list
   1. Pull data from support tables based on species
   2. Set up data frames and run discaRd on focal year
   3. Use transition rates by running discaRd on previous year
   4. Apply transition rates
   5. Match discard rates to individual trips by STRATA
   6. Observed trips receive the observed discard value for that trip
   7. Run GLM model for unobserved strata
   8. Substitute modeled rate where necessary
   9. determine final discard/trip $d\_{final}\*KALL\_{subtrip}$
5. Groundfish considerations
   1. Discard estimates are for groundfish trips only as determined by VMS declaration
   2. Electronic Monitoring estimates take precedent over other sources

Table description

The use of a combined catch and observation table allows for a single source table to be used in discard estimation. Previous methods took a two table approach, where catch information and observer records were stratified independently, and then matched to calculate discard (e.g., $D = K\*d/k$). This approach had the possibility of mismatches between observed strata and trip strata. In reality, this cannot occur. We therefore do trip by trip matching, using gear, mesh, statistical area, and LINK1, to match observer recorded species discards with commercial trip activity.

The primary driver for this approach was to use the trip recorded metrics as the stratification source. This reduces the possibility of mismatches and removes much of the hard-coding that has been used to date. Upfront matching also allows observed discards to easily be used as the official discard for a particular trip. Furthermore, we recognize that data errors either from the catch, or observer data, will result in a non-match. This likely reduces the total pool of observed trips that are being used, but we feel the benefits of using outweigh a reduced sample size. Quality control of these data are outside the purview of the CAMS project. Last, we only use observed trips (LINK1) where valid LINK3 (hauls) occurred to alleviate issues of multiple LINK1 records for a single subtrip.

IMPORTANT!: The table itself is a hybrid. For trips that were not observed, there will be a single row with all trip metrics and a total KALL per subtrip. When a trip was observed, there are multiple rows where the trip metric information is repeated, and each row shows species, discarded amount, and other observer recorded information for each row. Total KALL CANNOT be calculated without filtering rows by LINK1 to indicate an observed trip or not. These steps are outlined in subsequent R modules used to run disacRd.