**Historical comparison**

**Northern Bering Sea**

**CPUE:** There are 439 unique taxa included in HAEHNR.CPUE\_NBS (314,324 records). GAP\_PRODUCTS.CPUE has 435 unique observed taxa (311,460 records).

Of the shared 311,460 records, there are 26 mismatched records in 2022 that result from record modifications during the fish vouchering period in 2023.

**BIOMASS**: There are 439 unique taxa included in HAEHNR.BIOMASS\_NBS\_SAFE (6,585 records). GAP\_PRODUCTS.BIOMASS has 435 unique observed taxa (6,525 records).

Of the shared 6,525 records, there are 63 mismatched records. Of those 63 mismatched records, 47 are due to missing count data. In the Bering Sea, the sample variance of the numerical CPUE excludes hauls with missing count data, however the estimated variance of the mean numerical CPUE (sample variance of numerical CPUE / number of hauls) does not exclude hauls with missing count data in the denominator of that calculation, which presents a slight negative variance bias for these records. The remaining 16 records contain a CPUE that has been identified as being mismatched from the CPUE comparison.

**SIZECOMP**: There are 22 unique taxa included in HAEHNR.SIZECOMP\_NBS\_STRATUM (7,441 records). GAP\_PRODUCTS.SIZECOMP has 32 unique observed taxa (7,446 records).

Of the shared 7,441 sizecomp records, there are 18 mismatched records with at least a 0.5% difference due to differences in the population abundance estimates outlined in the biomass section.

**AGECOMP:** There are 4 unique species with age composition data present in both tables. 2022 was excluded in case of delays in uploading age data. Of the shared 1,007 agecomp records, there are 228 records that are in the production tables that are not in the agecomp tables. These records are -9 AGE inidividuals (there are lengths for those sexes but no specimen information. Maybe years where there are no collected age data should be removed?)

**Eastern Bering Sea**

**CPUE**: There are 72 unique taxa included in HAEHNR.CPUE\_EBS\_PLUSNW (1,048,630 records). GAP\_PRODUCTS.CPUE has 31,413 unique observed taxa (13,371,741 records).

Of the shared 1,068,552 records, there are 19,925 mismatched records. 19,922 records are zero-filled records for instances where a taxon was split during the time series (i.e., rock soles unid 🡪 NRS/SRS). The remaining three records were from 2022 and were removed from RACEBASE.CATCH.

**BIOMASS:** There are 90 unique taxa included in *HAEHNR.BIOMASS\_EBS\_PLUSNW* (41,672 records), six of which have not been observed (SPECIES\_CODE 21340, 23020, 23030, 23060, 23061, 23071). GAP\_PRODUCTS.BIOMASS has 901 unique observed taxa (423,470 records).

Of the shared 39480 records, there are 635 mismatched records. 628 records are from zero-filled records for instances where a SPECIES\_CODE was introduced after the start of the time series (e.g., SPECIES\_CODE 232 was added in 1984), 3 records were strata with mismatched CPUE values, and 4 records had missing count data.

**SIZECOMP:** There are 28 unique taxa included in *HAEHNR.BIOMASS\_EBS\_PLUSNW* (204,624 records). GAP\_PRODUCTS.SIZECOMP has 79 unique observed taxa (216,717).

Of the shared 205,414 size composition records, there are 3924 mismatched records with at least a 0.5% difference. Of those records, 193 are due to differences in the population abundance estimates outlined in the biomass section.

**AGECOMP:** Age-length keys are created by pooling specimen data within a region, so the age-length key for the EBS\_STANDARD (all strata except 82, 90) region will be slightly different from that of the EBS\_PLUSNW (including strata 82 and 90) region. To facilitate a useful comparison, the EBS\_PLUSNW region will be compared to GAP\_PRODUCTS.AGECOMP subsetted to years >= 1987 and inclusive of all strata. 2021-2022 data are excluded in case of any lags in data uploading.

GAP\_PRODUCTS.AGECOMP contains 19 unique species (101,844 records) and HAENHR.AGECOMP\_EBS\_PLUSNW\_STRATUM contains 28 unique species (202,872 records).

Of the shared 96,719 records, there are 10,351 records that are expected to be different because of the inclusion of specimen data from hauls with negative performance codes. There are no mismatches between datasets.

**Aleutian Islands**

**CPUE**: There are 76 unique taxa included in AI.CPUE (44,182 records), three of which have not been observed (SPECIES\_CODE 150, 21220, 21397). GAP\_PRODUCTS.CPUE has 1,413 unique observed taxa (8,478,000 records).

Of the shared 438001 records, there are 15,507 mismatched records. 14,181 records are zero-filled records that are not in the historical databases due to the splitting of rock soles into NRS and SRS and dark/dusky rockfish unid into dark and dusky rockfish. In 1304 records, there was an observed catch record in the historical tables but that catch record does not exist in RACEBASE.CATCH anymore. These are records that were vouchered and subsequently re-identified as another SPECIES\_CODE value. For the remaining 22 records, those corresponding catch records in RACEBASE.CATCH were modified but not updated to the AI.CPUE table.

**BIOMASS**: There are 76 unique taxa included in AI.BIOMASS\_STRATUM (52,128 records), three of which have not been observed (SPECIES\_CODE 150, 21220, 21397). GAP\_PRODUCTS.BIOMASS has 1,413 unique observed taxa (1,001,884 records).

Of the shared 51773 records, there are 2,429 mismatched records. 1756 records are zero-filled records that are not in the historical databases due to the splitting of rock soles into NRS and SRS and dark/dusky rockfish unid into dark and dusky rockfish. Of the remaining 673 records, 531 records are strata that contain mismatched CPUE values and 140 have different numbers of hauls. Specifically, when the biomass tables were initially created for year 1980 and stratum 423, 4 hauls of data were included but since then, one of those hauls were deleted from the AI.CPUE table but the biomasses were not rerun. Same story for STRATUM 313 in year 2022. The remaining 2 records were unidentified sculpin (21300, Cottidae) that were taken out of RACEBASE.CATCH, but not updated in AI.CPUE and AI.BIOMASS\_STRATUM.

**SIZECOMP:** Of the shared 162247 sizecomp records, there are 3924 mismatched records with at least a 0.5% difference. Of those records, 193 are due to differences in the population abundance estimates outlined in the biomass section.

**AGECOMP:** There are 18 unique species with age composition data present in both tables. 2022 was excluded in case of delays in uploading age data. In the AIGOA specimen data from hauls with negative performance codes whereas in gapindex, only specimen data from hauls with positive performance codes are included, which excludes 4803 records. Of the remaining 3,173 shared records, there are 810 mismatches. 320 of those mismatches are from years where no age data, 62 mismatches are from having different population estimates, 150 are from having a different number of ages, and 216 are from the historical age comps excluding unsexed individuals. The remaining 62 mismatched records come from POP in 2010 and shortraker rockfish in 2006 and cannot at this time be reconciled.

**Gulf of Alaska**

**CPUE:** There are 121 unique taxa included in GOA.CPUE (1,441,398 records). GAP\_PRODUCTS.CPUE included 1520 unique taxa, encompassing 18,439,120 records.

Of the shared 1467851 records, there are 44,749 mismatched records. 31,614 records of these records are zero-filled . In 13,088 records, there was an observed catch record in the historical tables but that catch record does not exist in RACEBASE.CATCH anymore. These are records that were vouchered and subsequently re-identified as another SPECIES\_CODE value. For the remaining 47 records, those corresponding catch records in RACEBASE.CATCH were modified but not updated to the AI.CPUE table.

**BIOMASS:** There are 1211 unique taxa included in GOA.BIOMASS\_STRATUM (108,193 records). GAP\_PRODUCTS.BIOMASS has 1,520 unique observed taxa (1,383,200 records).

Of the shared 110,110 shared records, there are mismatched records. 2827 are zero-filled records that are not in the historical databases due to the splitting of rock soles into NRS and SRS and dark/dusky rockfish unid into dark and dusky rock fish. 1893 records are strata that contain mismatched CPUE values, 345 have different number of hauls (Stratum 20 and 111 in 1984 and stratum 21 in 1987).

**SIZECOMP:** Of the shared 337,053 records, there are5388 mismatched records or records with at least a 1% difference, comprised of 180 unique combinations of YEAR/STRATUM/SPECIES\_CODE. 179 of those combinations had mismatches present in the stratum abundance estimates between the two tables. The mismatch of the outstanding combination could not be resolved and consists of 16 records (YEAR 1999, STRATUM 122, SPECIES\_CODE 31052).

**AGECOMP:** 28 unique species are included in GOA.AGECOMP\_TOTAL (X records) and 29 unique species (prowfish being the extra species) in GAP\_PRODUCTS.AGECOMP (X records). The historical age compositions were created with specimen data from negative-performing hauls. The gapindex package by default only uses specimen data from well-performing hauls so there will be species/year combinations where the two tables will not be comparable. Of the 15152 shared age composition records, 6340 records fall into this category. Of the remaining 8812 remaining shared records, there are 2320 mismatched records. Of those 2320 mismatched records, 461 records technically do not have age data for a given year/species combination so they should not have compositional data. In these examples, historically age data from non-abundance hauls were used to inform the agecomp calculation for that year/species. In gapindex process, age data from only hauls that were used in the biomass calculation (i.e., ABUNDANCE\_HAUL == “Y”) are included in the agecomp calculation. An additional 1077 records have different population estimates to begin with, 238 are from having a different number of ages, and 505 are from the historical age comps excluding unsexed individuals. The remaining 39 mismatched records come from blackspotted rockfish in 2015 and cannot at this time be reconciled.