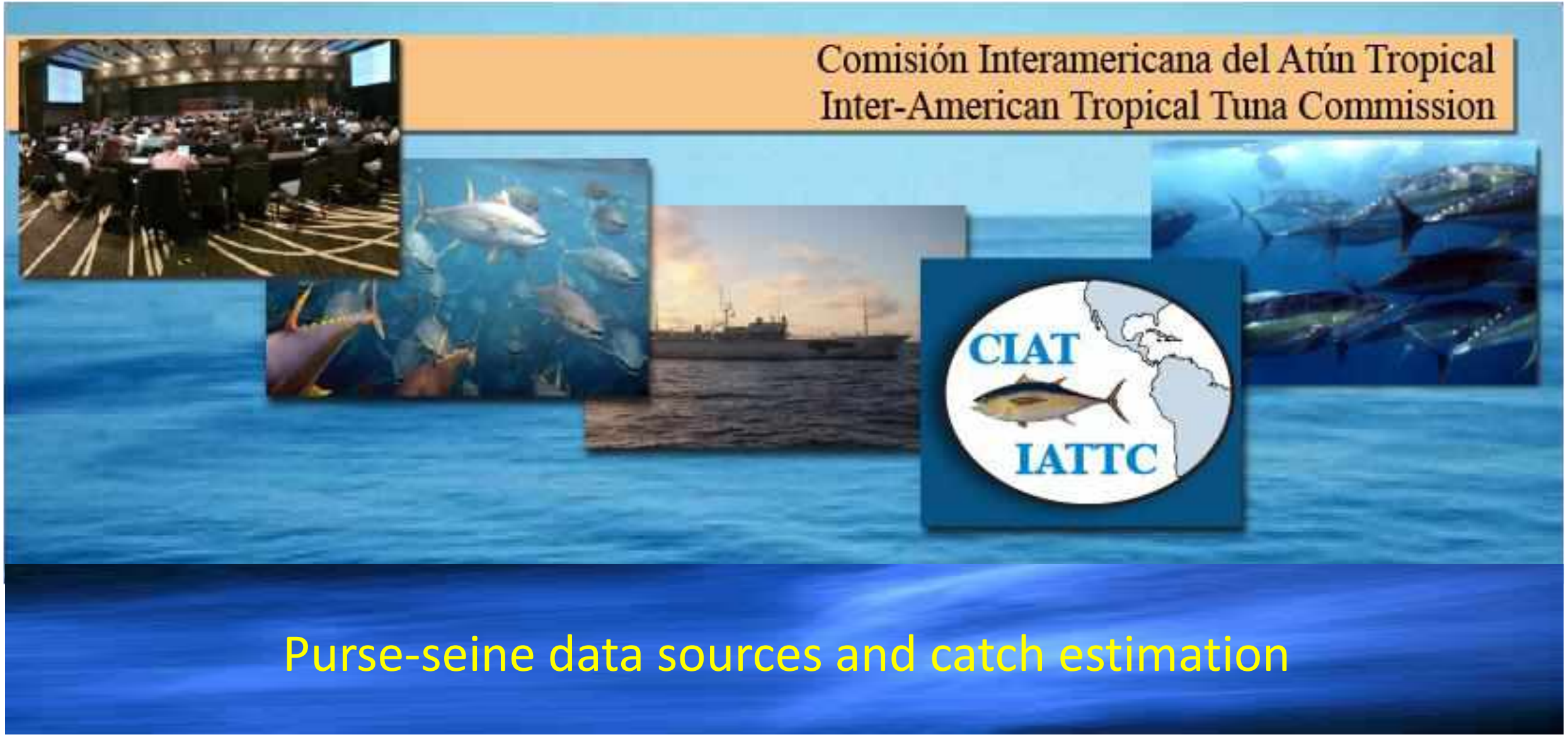


Purse-seine data sources and catch estimation

Data Review Workshop, La Jolla, California, USA, 2-6 October 2023

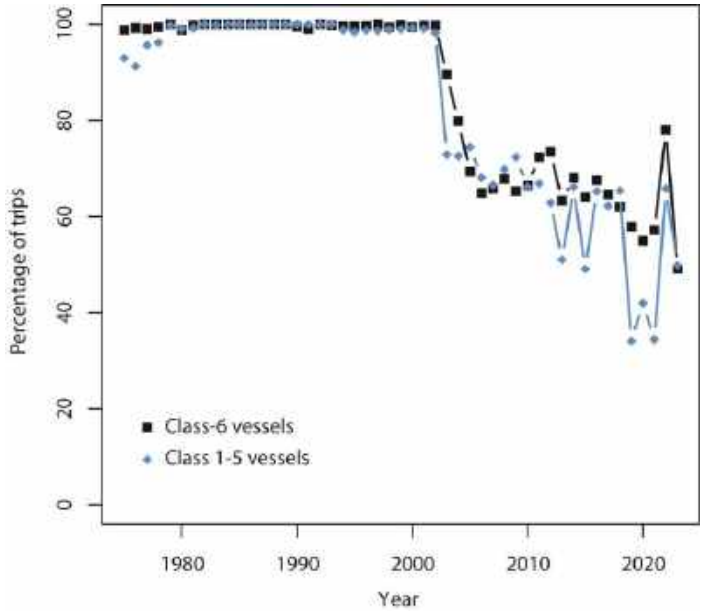
Presentation overview

- Data sources
  - Canner/processor
  - Logbook
  - Observer
  - Port-sampling

- Estimated quantities
  - Catch
    - General methodology
    - Estimation for strata with no port-sampling data (‘substitution’)
    - Matching fleet totals
  - Length composition

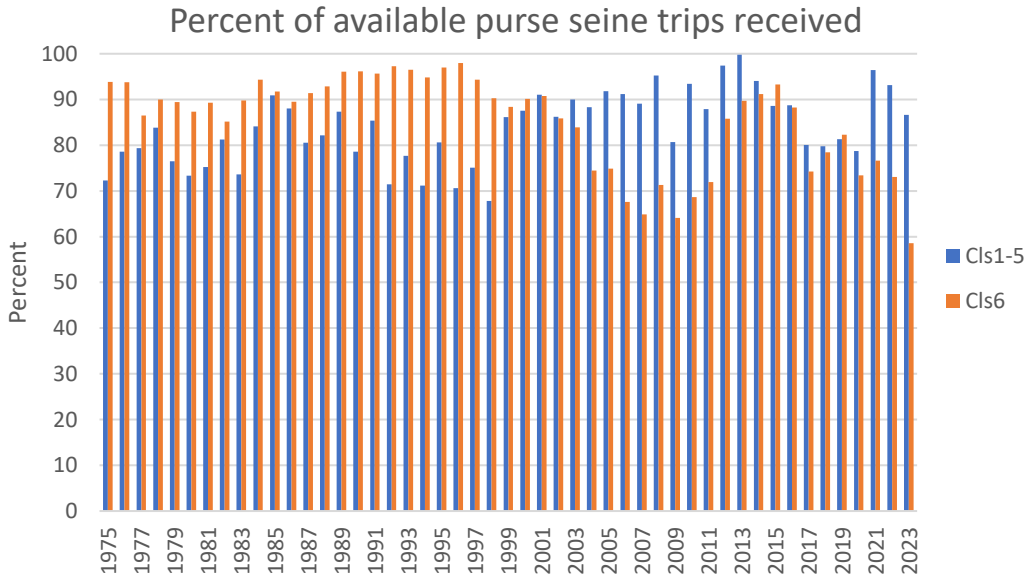
Data sources

- Canner/processor
  - Primary source for total fleet catch of tropical tunas (weight)
  - Does not include information on fishing area, set type or month/day



Canner/processor data: Percentage of trips for which catch data were reported from at least one cannery

- Logbook
  - Set-level information:
    - Retained catch (weight)
    - Purse-seine set type
    - Date and location of fishing



Data sources

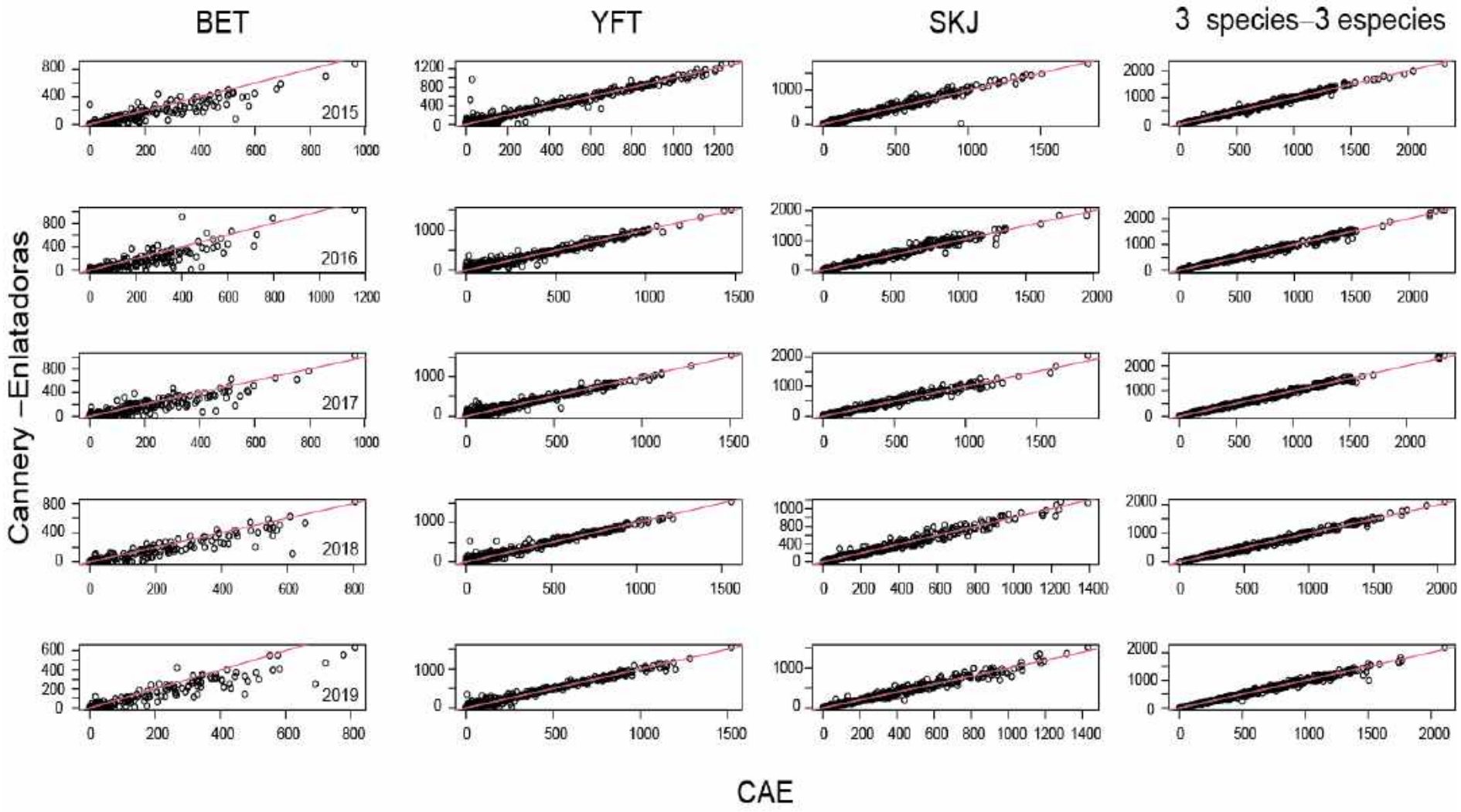
- Observer
  - Primarily large purse-seine vessels
  - Coverage at or nearly 100% since 1992
  - Detailed data:
    - Vessel fishing activities
    - Retained catch, discards (weight)
    - Fishing locations, dates, set types

Year	Total Number of Trips	Trips Sampled by the IATTC	Trips Sampled by National Programs	Combined Sampling Coverage (%)
1980	532	66	45	20.9
1981	447	60	37	21.7
1982	328	48	32	24.4
1983	248	33	0	13.3
1984	331	24	11	10.6
1985	381	47	23	18.4
1986	396	94	20	28.8
1987	473	125	80	43.3
1988	503	159	33	38.2
1989	543	194	73	49.2
1990	539	223	41	49.0
1991	425	237	26	61.9
1992	427	279	140	98.1

From Joseph 1994

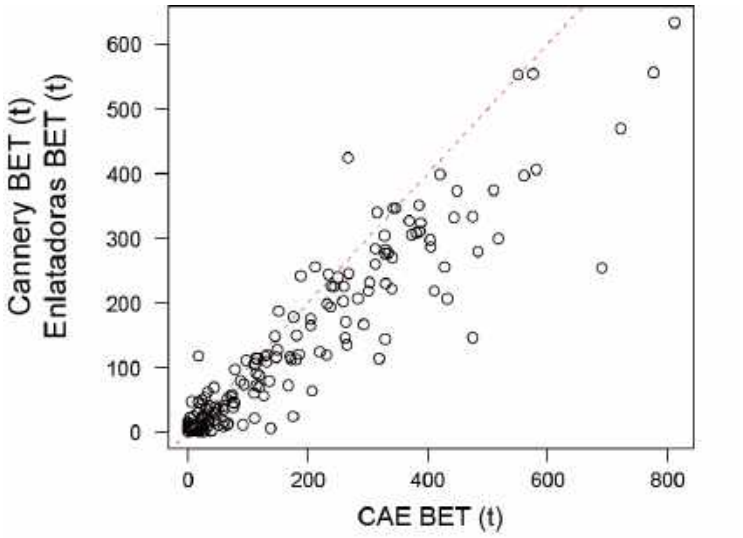
  - At-sea reports and Set summaries
    - Preliminary observer data source for set-level catch (weight)
    - Includes set type, date, 5° area
    - Superseded by full observer data once those data are processed

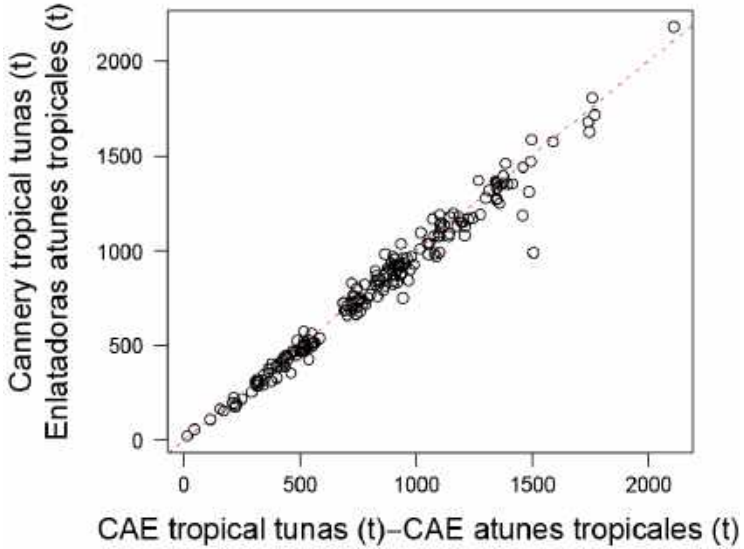
Data sources: comparing canner/processor and observer



From IATTC Document SAC-14-10

Data sources: comparing canner/processor and observer

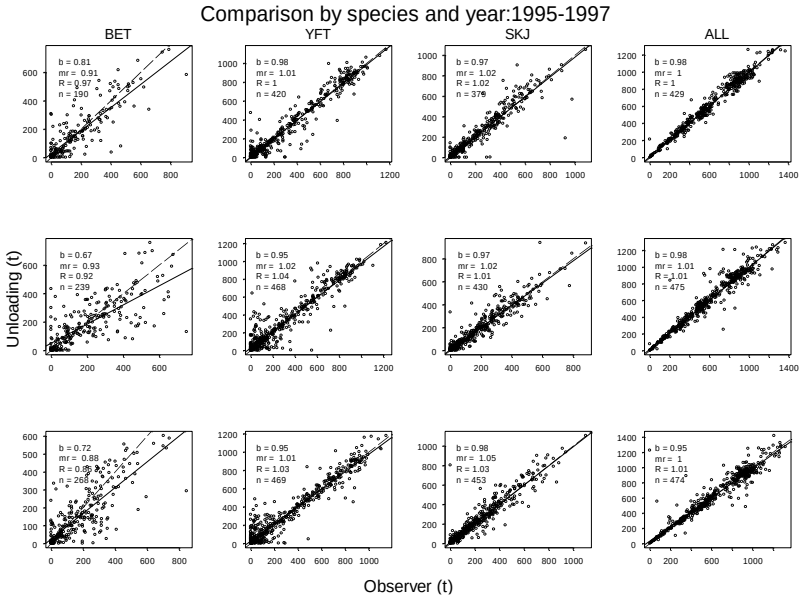


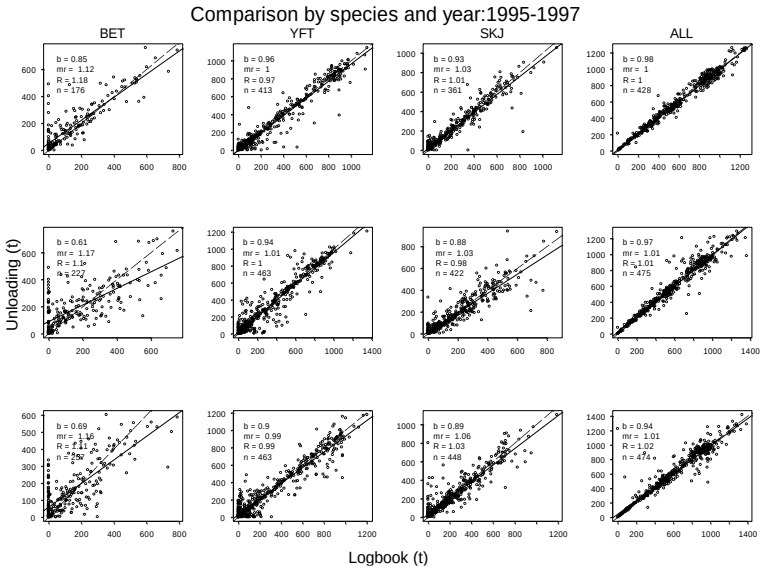


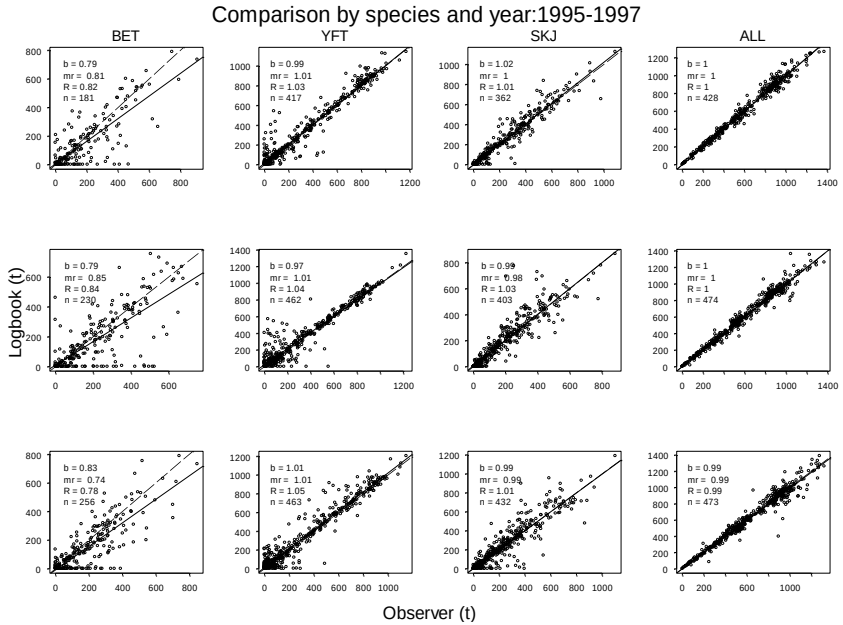
Year (number of trips)	BET GLM		Tropical tuna GLM	
	Estimated slope	% Dev	Estimated slope	% Dev
2015 (232)	0.64 (s.e. = 0.104; p < 0.01)	54%	0.96 (s.e. = 0.005; p < 0.01)	99%
2016 (221)	0.64 (s.e. = 0.049; p < 0.01)	62%	0.96 (s.e. = 0.007; p < 0.01)	98%
2017 (250)	0.79 (s.e. = 0.083; p < 0.01)	58%	0.97 (s.e. = 0.004; p < 0.01)	99%
2018 (230)	0.74 (s.e. = 0.053; p < 0.01)	69%	0.99 (s.e. = 0.008; p < 0.01)	99%
2019 (201)	0.75 (s.e. = 0.061; p < 0.01)	69%	0.96 (s.e. = 0.005; p < 0.01)	98%

- Comparing canner/processor, observer, by trip:
  - Similar for total tropical tuna catch (BET+YFT+SKJ)
  - Differ for species catch

Data sources: comparing canner/processor, observer, logbook



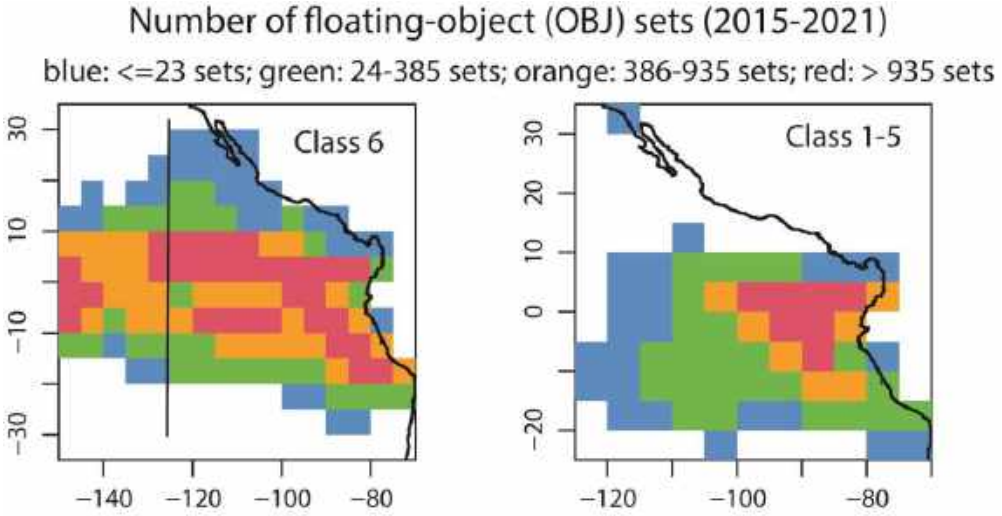


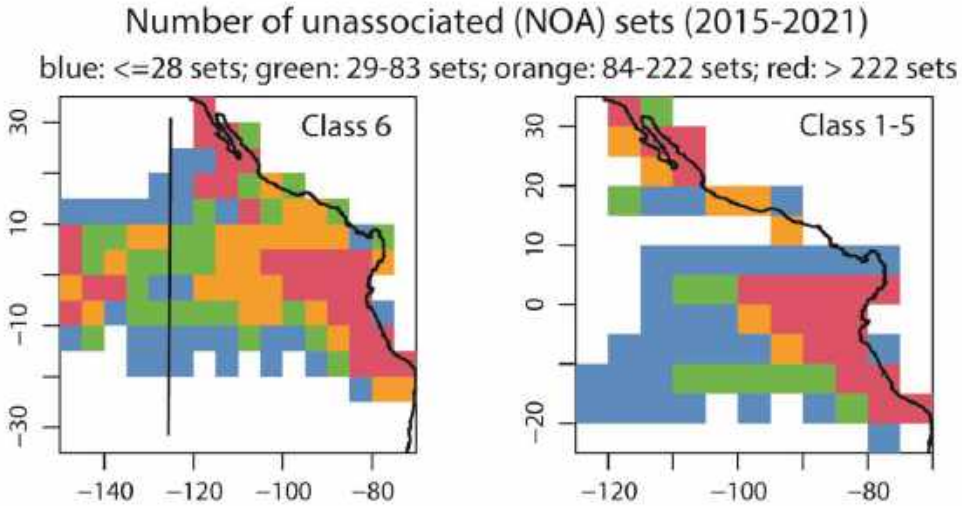


From Suter et al. 2004 unpublished

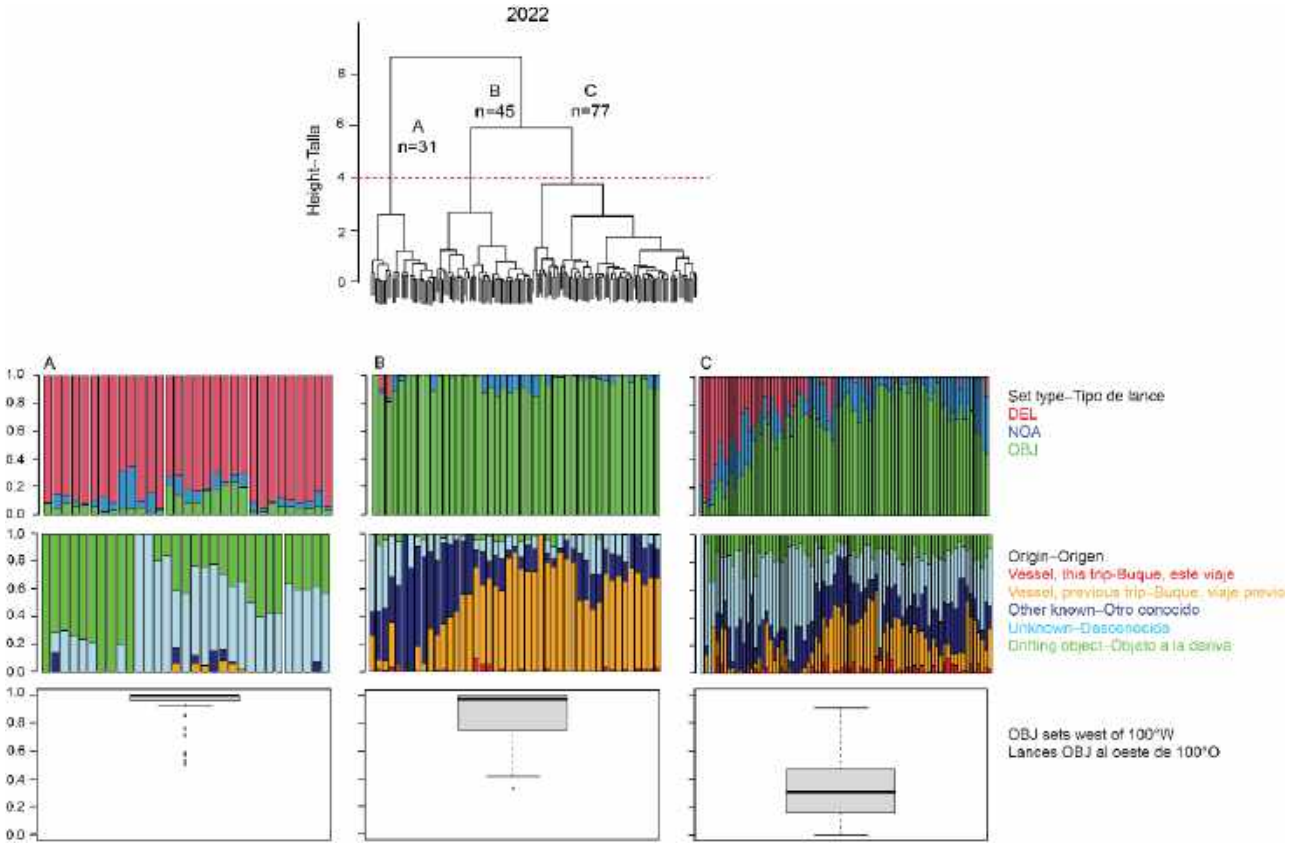
Question re: area of fishing activity, Class 6 vs. Class 1-5

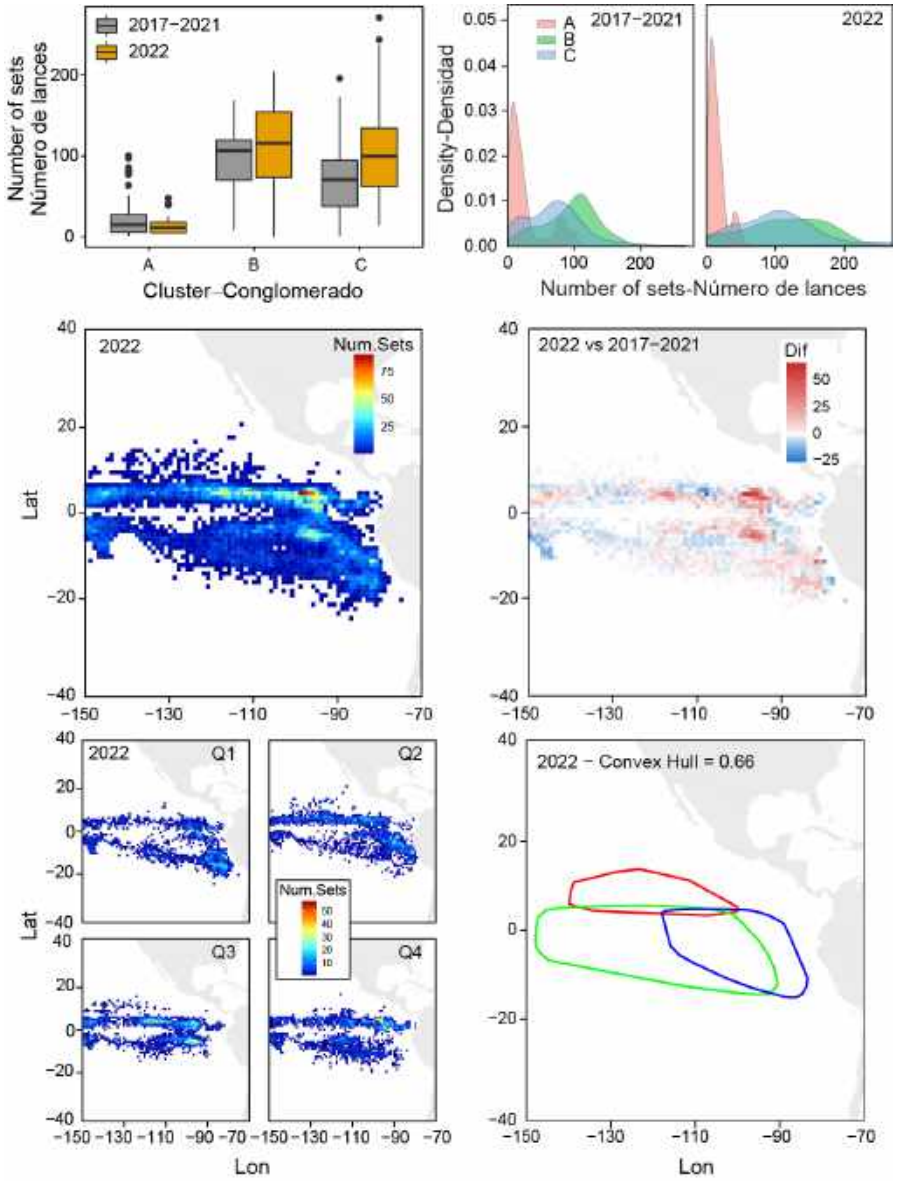
Class 1-5 purse-seine vessels tend to fish more coastally, compared to Class 6 vessels.





Question re: P-S fleet segments (Class-6 vessels)

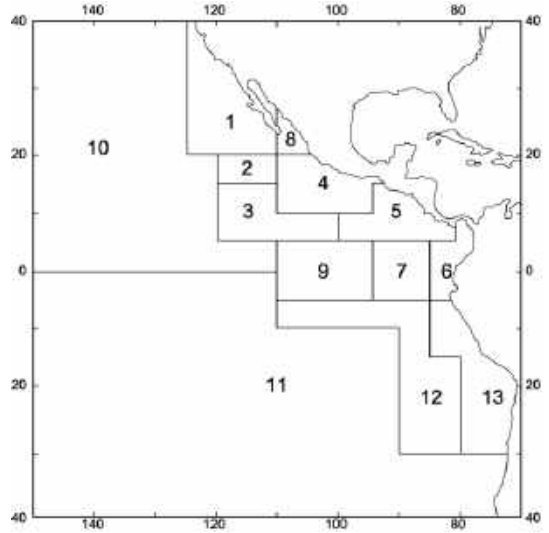




From IATTC Document FAD-07-01
Cluster analysis methods described in; ICES Journal of Marine Science (2018), 75(5), 1748–1757. doi:10.1093/icesjms/fsy046

Data sources

- Port-sampling
  - Two-stage protocol (vessel well; fish within a well)
  - Opportunistic, but catch in well must all be from same stratum
    - Strata: month x sampling area x set type x vessel size category
  - Well-level data:
    - Length composition (1975-1999)
      - Well sample: measurements of ~ 50 fish per species
    - Species and length composition (from 2000 onwards)
      - Counts for species composition
      - Separately, fish selected for length measurements
      - Well sample:
        - measurements of ~ 50 fish per species;
        - counts of 100+ fish for species composition.
    - Total tropical tuna catch for the well (weight)



  - Coverage
    - Annually about 50%-60% of Class-6 vessel trips, 10-20% of Class 1-5 vessel trips.
    - Coverage of wells is considerably less.

Estimated quantities: general catch estimation methodology

- Total catch by species
  - Estimates provided by strata (area x month x set type x vessel size class category).
  - For the assessments: stratum estimates summed over months, vessel categories, and possibly areas, to obtain estimates by fishery area x quarter x set type.

  - 1975-1999
    - Species catch sources: canner/processor; observer; logbook.
    - Species composition from these sources considered inaccurate.
    - Therefore, species catch is adjusted based on correction factors computed from the 2000 – 2004 species composition estimates.
    - Adjusted species totals prorated to strata using species stratum proportions from observer and logbook data.

Material from: Suter 2010 & Tomlinson (1992, 2002, 2004), as cited in IATTC Document WSBET-02-06

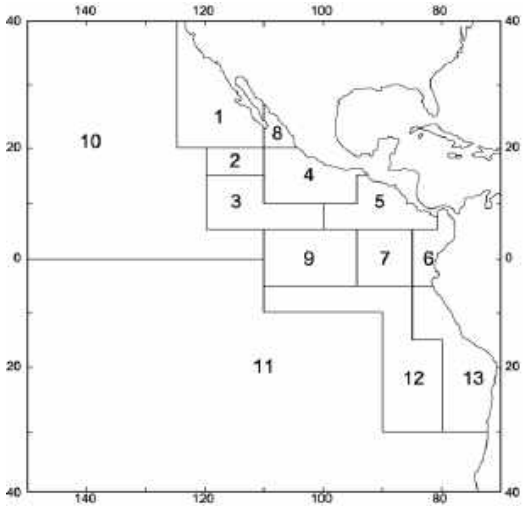
Estimated quantities: general catch estimation methodology

- Total catch by species
  - From 2000 onwards
    - Catch (BET+SKJ+YFT) sources: canner/processor; observer; logbook
    - Observer and logbook data used to prorate BET+SKJ+YFT catch to strata.
    - Port-sampling data used to estimate the species composition by stratum, considering different average weights among species $\left(\frac{w_{ij}}{m_{ij}}\right)$ and that sample species proportions $\left(\frac{n_{ij}}{n_{.j}}\right)$ are from numbers, not weight:

$$\widehat{W}_i = W\left(\frac{\sum_j^q \widehat{W}_{ij}}{\sum_j^q W_j}\right) = W\left(\frac{\sum_j^q W_j\left(\frac{\left(\frac{w_{ij}}{m_{ij}}\right)\left(\frac{n_{ij}}{n_{.j}}\right)}{\sum_i^s\left(\frac{w_{ij}}{m_{ij}}\right)\left(\frac{n_{ij}}{n_{.j}}\right)}\right)}{\sum_j^q W_j}\right)$$

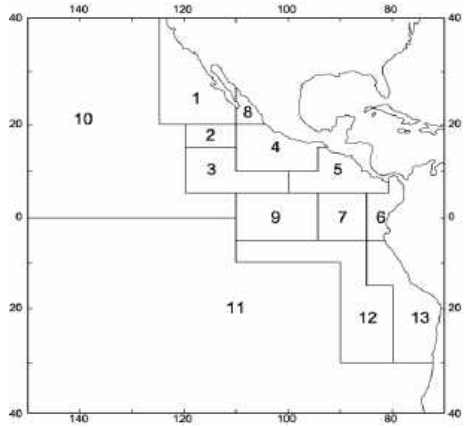
Estimated quantities: ‘substitution’ (2000 and onwards)

- Catch estimation for strata with catch but no port-sampling data
  - For estimates based on the sampling areas, there are a total of 780 possible strata (= 13 areas x 12 months x 5 set type – vessel categories).
  - On average, about 64% of these 780 strata had catch in any given year during 2010-2019.
  - Many of the strata with catch had no port-sampling data: on average, about 73% of the strata with catch, representing 24% of the annual fleet catch, had no port-sampling data.
  - Most of these strata, individually, had proportionally little catch.
  - With fewer areas (e.g. new fishery areas), there may be fewer strata with catch but no port-sampling data; however, the problem will likely persist.
  - Reasons for strata without port-sampling data include:
    - sampling does not occur in all ports where catch unloaded;
    - mixed-stratum wells not sampled;
    - sampling characteristics (sampling opportunistic, active strata not known in advance, coverage of wells relatively low).



Estimated quantities: ‘substitution’

- In the catch estimation methodology, port-sampling data from ‘neighboring’ strata are used to estimate the species composition of catch of strata with no port-sampling data.
- A set of hierarchical rules define the choices of ‘neighbors.’
- The length data of the neighboring stratum is ‘grown’/’shrunk’ to the month of the stratum with no port-sampling data, if necessary.
- Even though the substitution rules are fixed, neighbors used as ‘substitutes’ for the same stratum may be different in different years.
- Substitution can effectively change the catch estimation ‘model’ from one year to the next, which may introduce added variability into the catch time series.



Example of some of the substitution rules for area

1	**2**	**3**	**4**	**5**	**6**	**7**
8	1	10	5	7	7	9
2	4	9	8	9	13	6
4	8	2	3	3	5	5
3	3	4	2	6	12	12
5	10	5	1	4	9	13

Estimated quantities: matching fleet totals (2000 onwards)

- There is an estimation challenge created by using different fishery definitions for each species.
  - In current sampling/estimation framework, the sum of estimated catches of the three species for the EPO (or stratum) should equal the total fleet catch of tropical tunas for the EPO (or stratum) obtained from canners (processors)/observers/logbooks.

  - When fishery definitions differ by species (and estimation is done by species), the above is not guaranteed.

Estimated quantities : matching fleet totals (2000 onwards0

- Within the current sampling/estimation framework, several possible options to address this:
  - Define sub-areas for catch estimation that can be combined to produce the fishery areas used in each assessment.
    - This approach has been used in the past.
    - It can be problematic if the sub-areas must be very small because there may be no port-sampling data available (i.e. more 'substitution' might be required).
  - Estimate for two of the three species separately
    - The EPO estimate for the third species is the total fleet catch of tropical tunas minus the EPO estimates for the other two species.
    - The EPO estimate for the third species is prorated to fisheries using the fishery species proportions computed from observer/logbook data.
    - This could produce poor fishery-level estimates of the third species if species proportions in observer/logbook data are highly accurate.
  - Other options?

Estimated quantities: length composition

- Length composition of the catch is estimated using port-sampling data
  - Estimates are numbers of fish in 1 cm length bins.
  - Estimated numbers at length are obtained by multiplying the well-level estimates of proportion at length, combined across sampled wells (*j*), by the estimated total catch in numbers for the species (*i*) in the stratum.
  - Strata without port-sampling data are not used to obtain the estimates for the stock assessments.

- 1975-1999

$$\widehat{N}_{ik} = \widehat{N}_i \hat{f}_{ik} = \left(\frac{W_i}{\overline{w}_{i.}}\right)\left(\frac{\sum_j^q \widehat{N}_{ijk}}{\sum_j^q \widehat{N}_{ij.}}\right)$$

- 2000 onwards

$$\widehat{N}_{ik} = \widehat{N}_i \, \hat{f}_{ik} = \left(\widehat{N} \, \hat{f}_i\right)\left(\frac{\sum_j^q \widehat{N}_{ijk}}{\sum_j^q \widehat{N}_{ij.}}\right)$$

Estimated quantities: other size composition estimates

- Species-specific weight composition of the catch from observer data
  - Estimates of species weight per set in 3 weight categories: small, < 2.5 kg; medium, 2.5 – 15 kg; large, > 15 kg.
  - Weight composition data are available for all Class-6 vessel sets (i.e., at high spatio-temporal resolution).
  - However, weight categories are coarse and estimated by eye; and, concerns have been expressed about the accuracy of the species amounts and identifications.

Thank you! Questions?

Estimated quantities

- Length composition estimation equations in more detail
  - 1975-1999

$$\widehat{N}_{ik} = \widehat{N}_i \hat{f}_{ik} = \left(\frac{W_i}{\overline{w}_{i.}}\right)\left(\frac{\sum_j^q \widehat{N}_{ijk}}{\sum_j^q \widehat{N}_{ij.}}\right) = \left(\frac{W_i}{\left(\frac{\sum_j^q W_{ij}}{\sum_j^q \left(\frac{W_{ij}}{\overline{w}_{ij}}\right)}\right)}\right)\left(\frac{\sum_j^q \widehat{N}_{ijk}}{\sum_j^q \widehat{N}_{ij.}}\right) = \left(\frac{W_i}{\left(\frac{\sum_j^q W_{ij}}{\sum_j^q \left(\frac{W_{ij}}{\overline{w}_{ij}}\right)}\right)}\right)\left(\frac{\sum_j^q \left(\frac{W_{ij}}{\overline{w}_{ij}}\right)\left(\frac{m_{ijk}}{m_{ij}}\right)}{\sum_j^q \left(\frac{W_{ij}}{\overline{w}_{ij}}\right)}\right)$$

  - From 2000 onwards

$$\widehat{N}_{ik} = \widehat{N}_i\, \hat{f}_{ik} = \left(\widehat{N}\, \hat{f}_i\right)\left(\frac{\sum_j^q \widehat{N}_{ijk}}{\sum_j^q \widehat{N}_{ij.}}\right) = \left[\left(\frac{W}{\left(\frac{\sum_j^q W_j}{\sum_j^q \widehat{N}_{.j.}}\right)}\right)\left(\frac{\sum_j^q \widehat{N}_{ij.}}{\sum_j^q \widehat{N}_{.j.}}\right)\right]\left(\frac{\sum_j^q \widehat{N}_{ijk}}{\sum_j^q \widehat{N}_{ij.}}\right) = W\left(\frac{\sum_j^q W_j \left(\frac{1}{\sum_i^s \left(\frac{w_{ij}}{m_{ij}}\right)\left(\frac{n_{ij}}{n_{.j}}\right)}\right)\left(\frac{n_{ij} m_{ijk}}{n_{.j}\, m_{ij}}\right)}{\sum_j^q W_j}\right)$$