Strength and uncertainties in the results of the TTT software, used to estimate statistics of purse seiners (catch and catch at size) and on the ways to improve these results

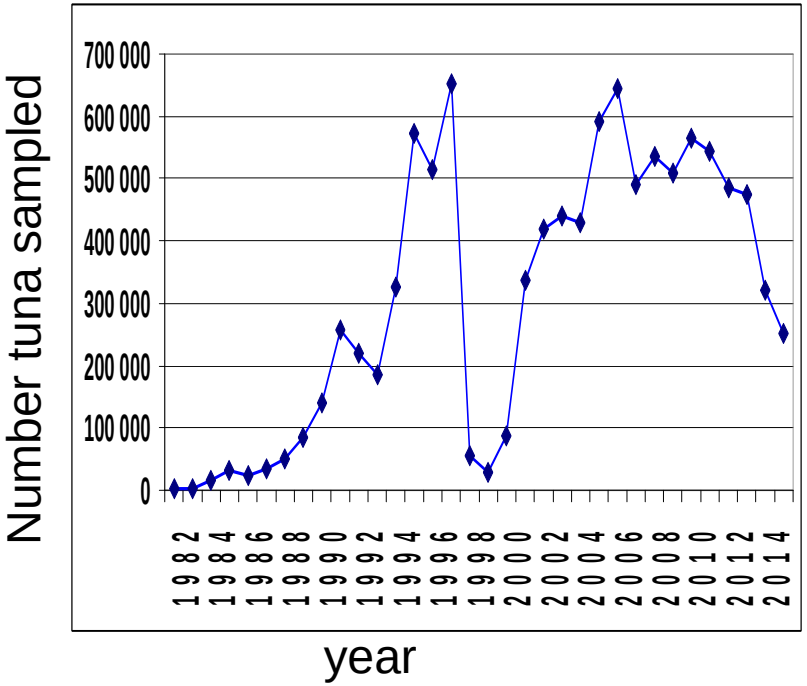
by Alain Fonteneau*

- This 2018 meeting of the IOTC TT WG: 2 papers submitted on this question by Spanish and French scientists on this complex question
- These 2 papers contains are discussing various important questions, but they are often quite weak and misleading in their presentation and conclusion
- And as a result these 2 papers are providing conflicting results & conclusions, and without a clear scientific analysis of the real characteristics and problems of the PS statistics handled by the TTT software, and on the best ways to solve these present problems
- This short presentation will try to develop a simplified & clear scientific overview of the strong qualities but clear potential problems of the past TTT results of the various concerned fleets in the Indian Ocean,
- It will also make various recommendations allowing to reduce the past and future uncertainties and errors in the PS statistics handled by TTT

* Alain Fonteneau, retired IRD scientist. Alain.fonteneau@ird.fr

EU & Seychelles CE & CAS PS statistics handled by the TTT software: a very strong data base since its beginning in 1981

- Nearly always 100% of log books collected at all landings ports of purse seiners
- Very good agreement between total tuna catches in log books and in commercial data
- Full VMS validation of all fishing positions: since 2002 for France, more recently for Spain
- Large scale multispecies sampling of sizes and of species composition, in Victoria but also in various other ports, and very large numbers of tunas sampled each year: **11 millions of tunas sampled since 1982** (small numbers sampled in the 1998-2000 period): these multispecies samples are used in the data processing of each the concerned PS fleet



Numbers of tunas sampled yearly

-This sampling has been targeting most landings of PS, independently of their fishing strata & they cover very well most fishing strata, even at small scale (as it was shown by the SMS analysis of Fonteneau and Lucas 2015 based on small 5° and 3 month strata)

- A strong data processing software,the TTT software, based on an in depth statistical analysis of the 1991-1996 data, was established by EU & Seychelles scientists and routinely used since 1998.
- This data set and its data processing method are of course questionable today because of mulktple changes in, the fisheries, but they remain today and by far, the best ones in the Indian Ocean tuna statistics.

But these TTT results are clearly facing multiple various uncertainties and potential biases that have been discussed by scientists in recent years and in the two 2018 documents by Herrera & al et by Duparc et al., such as:

1) Potential overestimation by TTT of large BET catches in some recent years (Fonteneau et al 2009)
2) Excessive smoothing of the TTT results, and especially in the Atlantic, potential bias in the species composition of each flag, due to the large strata used, where the species composition and sizes are assumed to be homogeneous (Fonteneau & Lucas 2015);
3) Past use by TTT during the 1982-2015 period of 3 length weight relationships (the Atlantic SKJ L/W was used) that were clearly highly inadequate.
4) TTT errors in the Spanish & Seychelles landings of large YFT and inverse errors in the estimated catches of small BET/SKJ/YFT, a potential problem shown by Herrera & Baez 2018.
5) Minor TTT bias in the catches of large tunas, due to the catches of small tunas identified in the log books, but sold to local markets or discarded (a problem shown by Herrera & Baez 2018)
6) Systematic smoothing by the TTT data processing of the species composition & CAS of all the individual vessels (a problem shown by Herrera & Baez 2018)

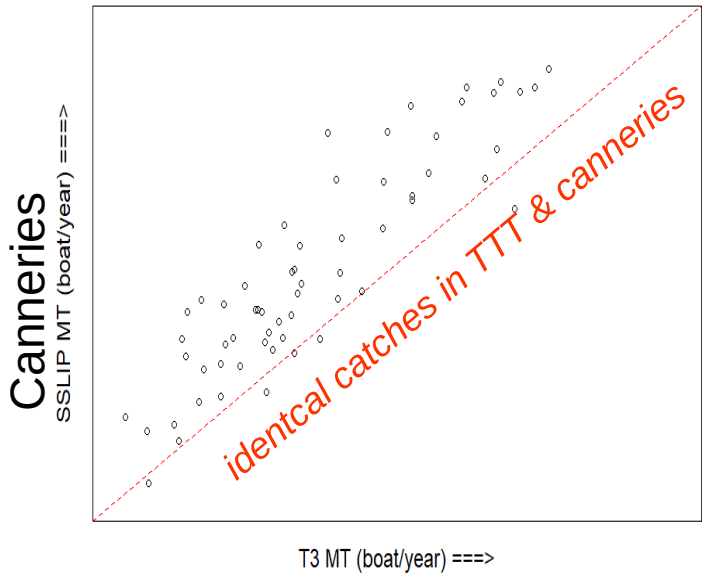
Herrera & Baez 2018: « T3 appears to largely underestimate the catches of large tunas over 10kg »

➢This statement is widely misleading because the **TTT software does not correct the catches of large YFT or large BET**; these catches are only slighthly increased or decreased by a global raising factor between the total log book and landed catches. Consequently, **all the differences between TTT and commercial catches of large YFT are only due to serious errors and bias in the studied logbooks.**

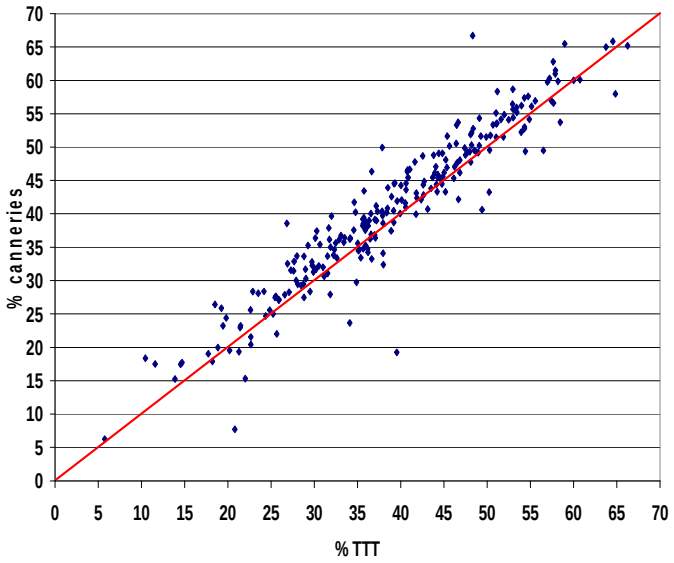
The problem in the TTT software is due to the fact that TTT is assuming that the catches of large YFT are well identified in the log books (by the skippers and the chief engineers).

➢On the opposite the analysis of the French data shows yearly catches of large YFT + 10 kg that are very similar to the level of large YFT in the cannery data.

➢**If the log book data are OK for large YFT: the TTT results will be OK & without bias**



Spain & Seychelles



TTT

France (each point: 1 boat & 1 year)

Tunas >10kg catches: in the TTT & the canneries data

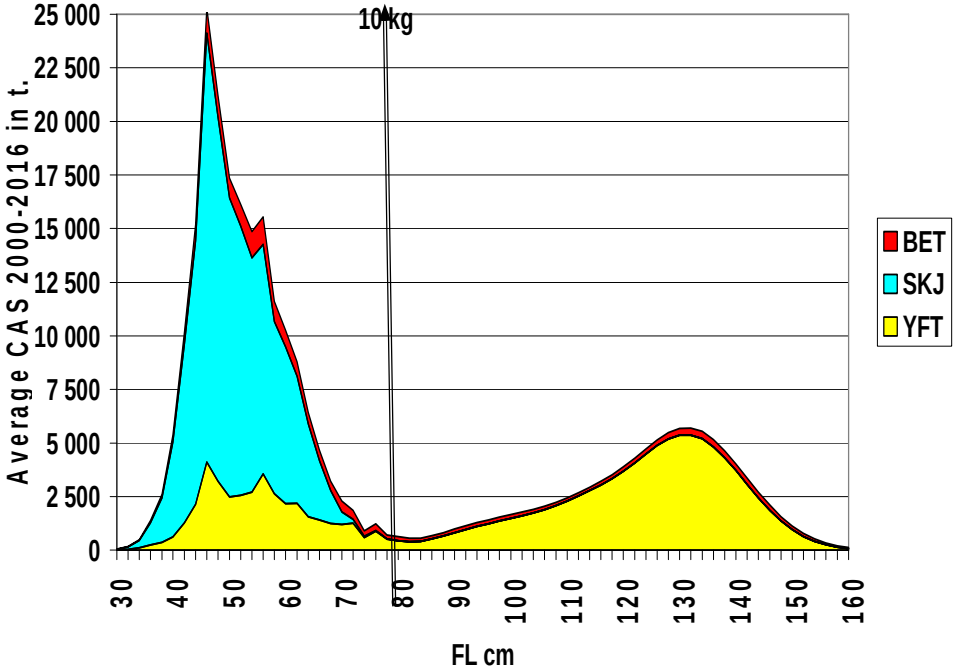
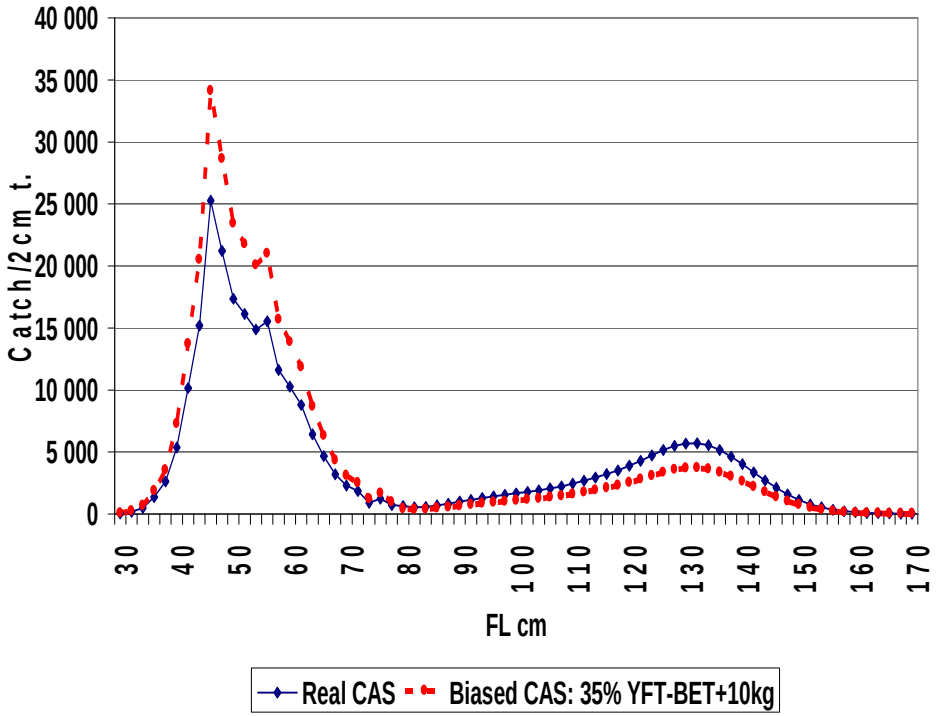


Figure showing the average CAS of IO PS for the 3 species combined, period 2000-2016: 100.000t. of large tunas >10kg and 200.000t of small tunas



In blue the same total CAS, and in red the CAS estimated with catches of large YFT overestimated of 35% (Spanish case)

Log books errors in the weight of large tunas producing inverse errors in TTT results:

Example: under the average multispecies CAS of Indian Ocean PS, the weight of large YFT is at an average level half the weight of the 3 species of small tunas (left figure)

When large errors& biases in the log books are overestimating the weight of large YFT, then the weight of the small tunas will be correspondingly underestimated (right)

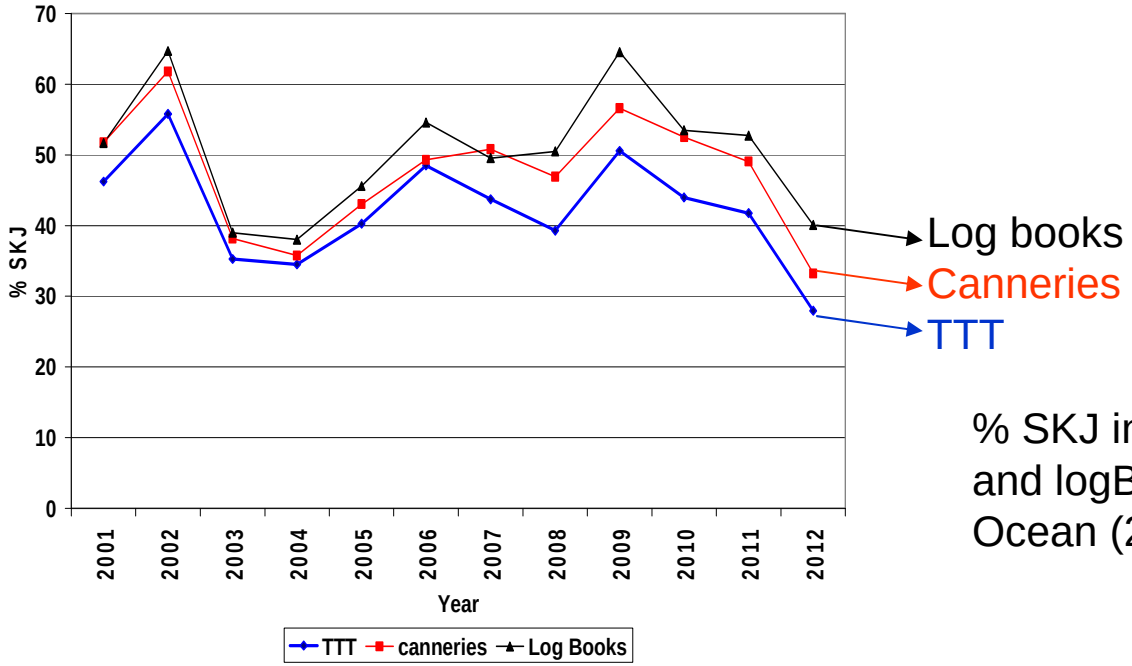
Skipjack and small sizes tunas catches

Herrera & Baez 2018 «T3 appears to largely overestimate the catches of skipjack tuna throughout the time series »

This conclusion is widely invalid: this overestimation of SKJ catches is not really due to a TTT bias: it is simply the consequence of the overestimated catches of large YFT in the Spanish log books (see previous figure right).

It should also be noted that an opposite result is observed in the TTT French species composition: SKJ catches estimated by TTT are always significantly lower than in the log books and in the commercial landings

These differences could be explianed by errors in the species identification by canneries or by unknown TTT biases (in the sampling or in the data processing?). Skipjack catches estimated by canneries remain questionable?.

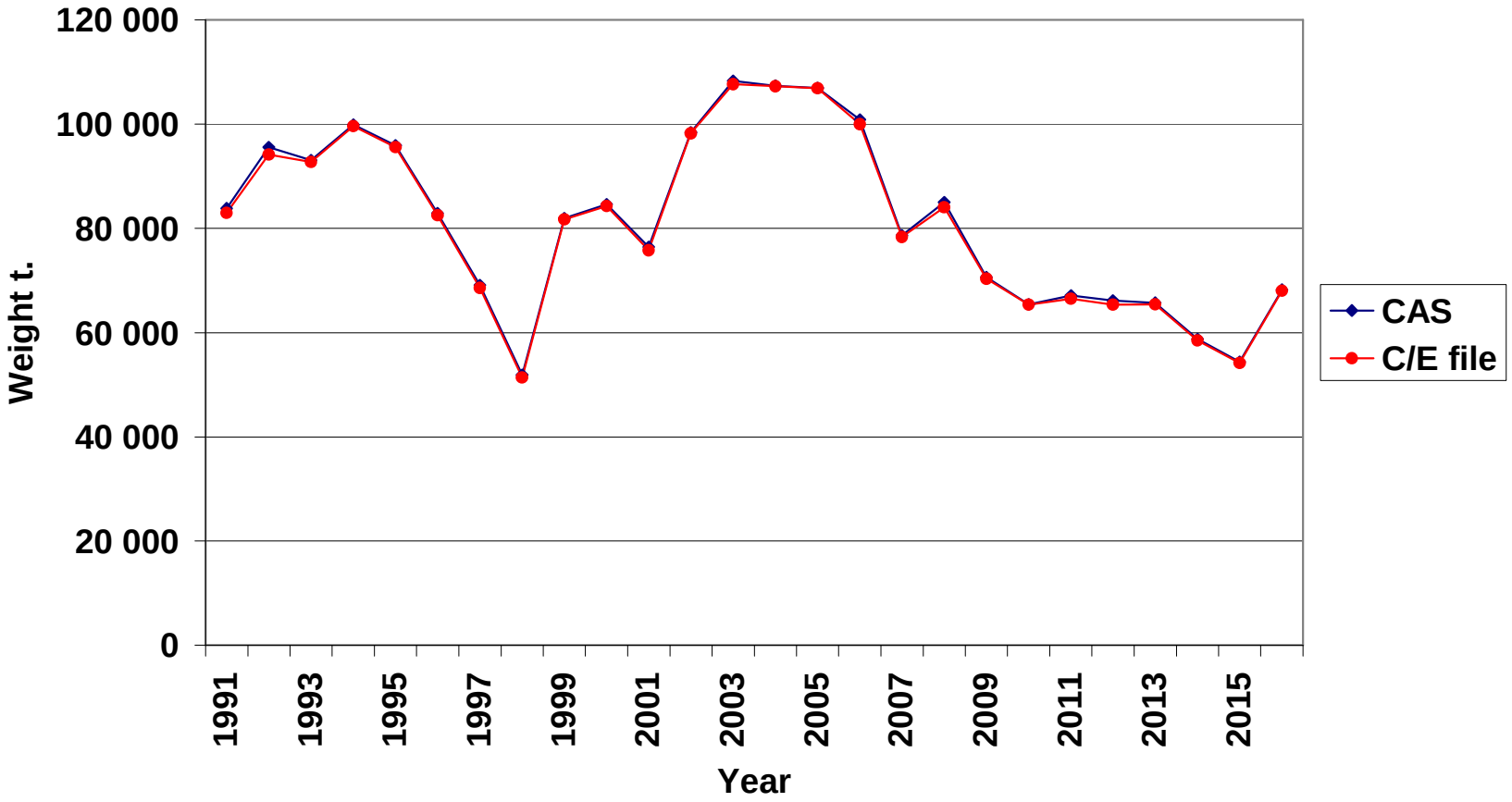


% SKJ in the TTT results, in the cannery and logBooks data for the French PS Indian Ocean (2001-2012)

Errors in the TTT catch at size results?

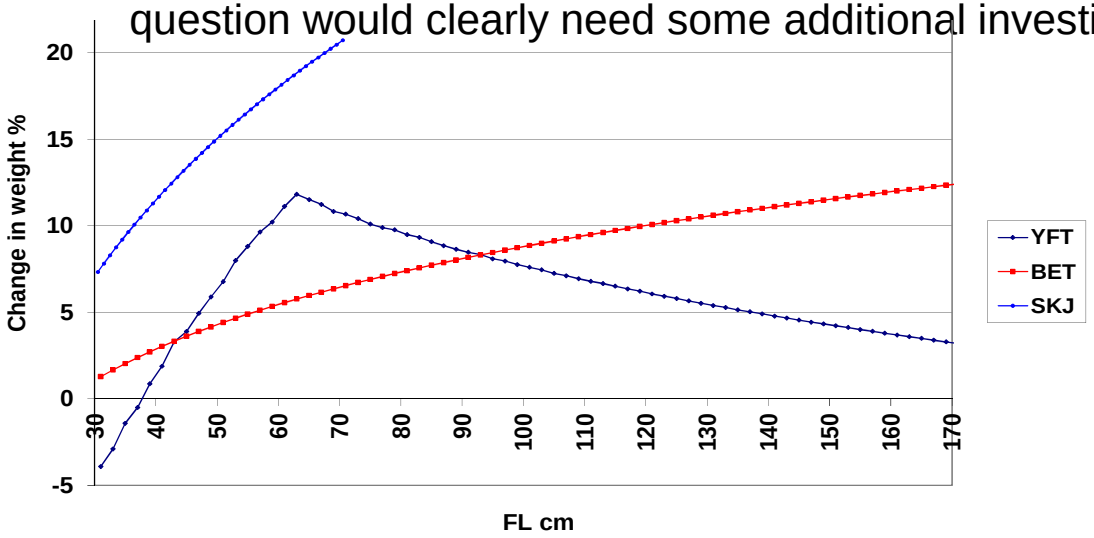
Herrera & Baez 2018 « Complete lack of catch-at-size data for some CE strata: overall: 15% of the strata for which CE exists does not have catch-at-size data»

This conclusion is totally invalid for all PS fleets: in fact there is a nearly perfect agreement in all years between the C/E and the CAS data in weight estimated by TTT, as it can be seen in the figure showing the yearly catches of French PS, 4 species, in the CAS and in the C/E files (Same result for the Spanish and Seychelles PS catches)



Questionable length weight relationship used by TTT?

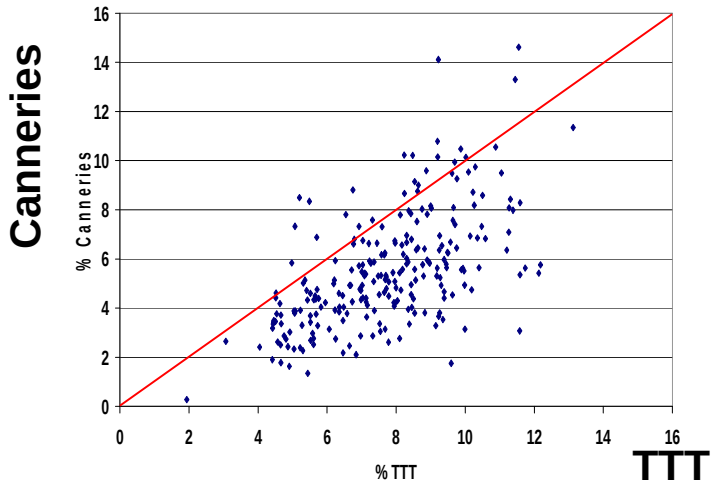
- ✓ Our multispecies sampling is based on numbers of tunas converted later to weight and the same L/W relationship have been used since 1982: any error or uncontrolled variability in these L/W have been introducing potential bias in the species composition, especially for small tuna catches: This problem has been well analyzed by Marsac et al 2017.
- ✓ TTT has been using in the IO during the 1982-2015 period 3 length weight relationships that are clearly highly inadequate, especially for SKJ, as the Atlantic SKJ built in 1981 was used by TTT until 2015, but also for BET & YFT (see Chassot et al 2016) .
- ✓ These unproper L/W relationship have been introducing unknown errors in the estimated species composition of past years: these errors should be corrected ASAP, as recommended during the 2017 IOTC meetings. This new data reprocessing is very easy and very fast!
- ✓ This problem of LW is potentially solved today by the new relationship proposed by Chassot 2016, but the new data processing of past data is still pending and this L/W question would clearly need some additional investigations



Relative changes (in %) in the weight at size of the 3 species, between the historical LW relationships used in the PS data processing (1982-2014) and the new LW relationship of Chassot & al 2016

A critical « ghost question » in the Herrera & Baez paper: quality of the species identification by canneries and by scientists?

- There is no doubt that 100% of the large YFT are well identified by scientists as well as by canneries: the weights of these large tunas should be fully consistent (as for French data) and never showing the Spanish pattern. **The weight of large YFT that have been well estimated by canneries should be fully used in TTT** for each landing
- But **serious pending questions remain concerning the perfect identification of small tunas (small YFT, SKJ & BET) in commercial data:** even if both scientists and canneries are confident of their good species identification ….
- The systematic differences in the % of BET in the cannery and TTT data are seriously questionning the cannery data: BET catches are most often much more important in the scientific samples than in the cannery data, and scientific BET samples are probably more realistic
- **It cannot be assumed today, as in the Herrera & Baez 2018 paper, that the canneries species identification of small size tunas is fully valid & better than the scientific sampling,**

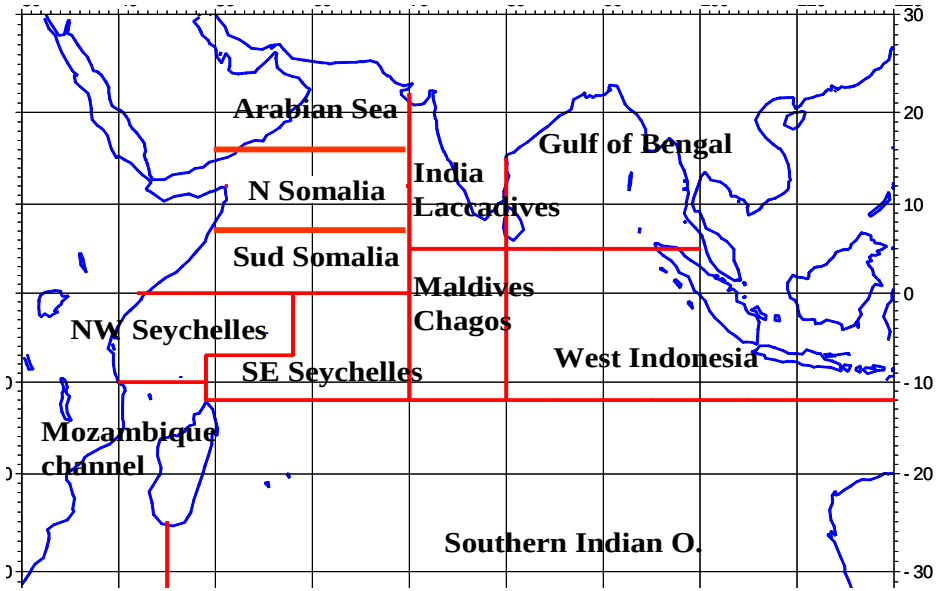


French PS: percentages of BET in the TTT results & in the canneries
(each point: 1 boat/year,
period 2000-2016)

BET landings have been most often widely underestimated in the cannery data

Effects of large strata used by TTT

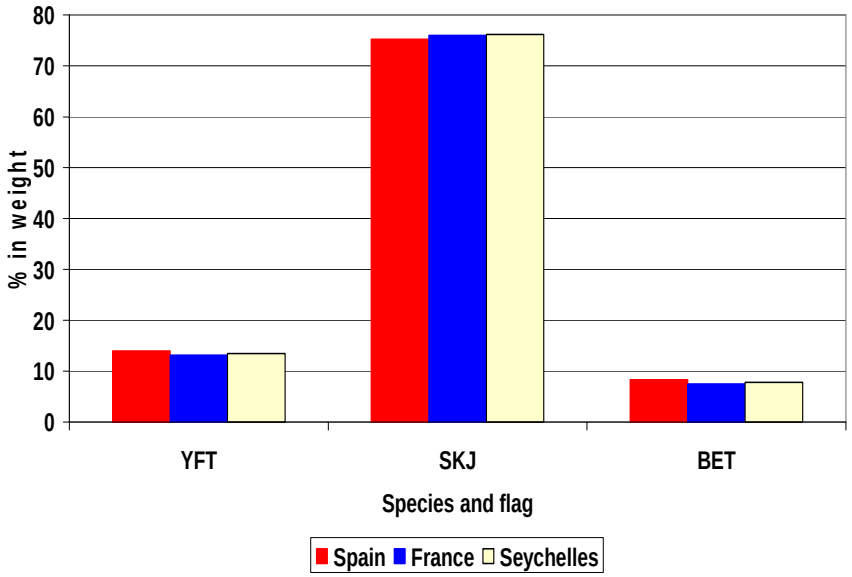
✓There no doubt that the large time-area strata used by TTT to generate its catches by species and CAS datasets have been producing wide smoothing of all its results
✓However it should be noted that in the Indian Ocean, **this smoothing by TTT does not introduce visible biases in the results**, simply because the areas used by TTT have been well selected in term of their position and sizes (the opposite was observed in the Atlantic, because of the too large and too heterogeneous FAD areas used)
✓This statistical issue has been discussed by IOTC scientists at least since 2015, following the document by Fonteneau & Lucas 2015 proposing an alternate fine scale data processing, the SMS (**S**mall **M**obile **S**trata) method, based on small mobile strata.
✓The fine scale stratification used in the SMS model reinforces the heterogeneity of the species composition and CAS, including at the level of each PS landing, when they are fishing in the same TTT strata , but on distinct concentrations with distinct species & sizes



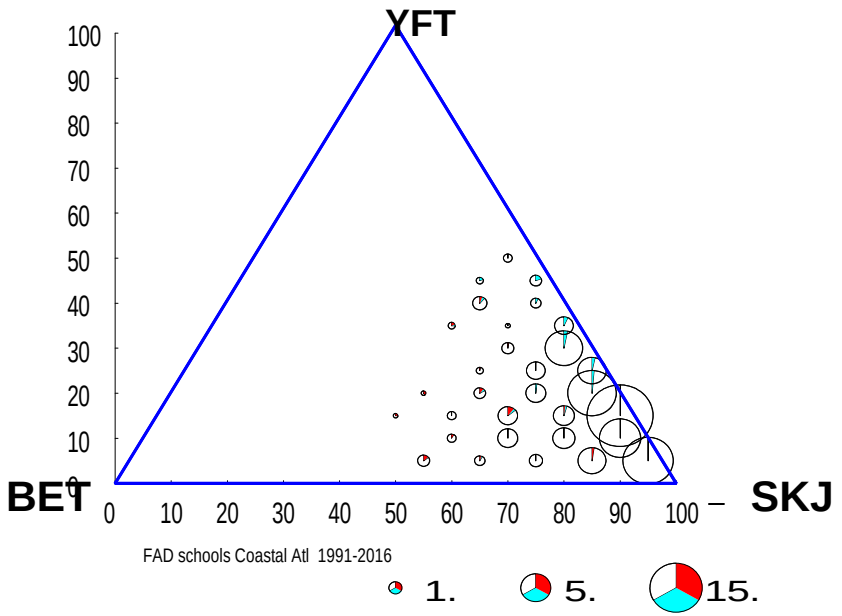
TTT geographical strata

(TTT data processing by fishing mode, area, quarter & size categories of tunas)

Typical example of excessive data smoothing by TTT



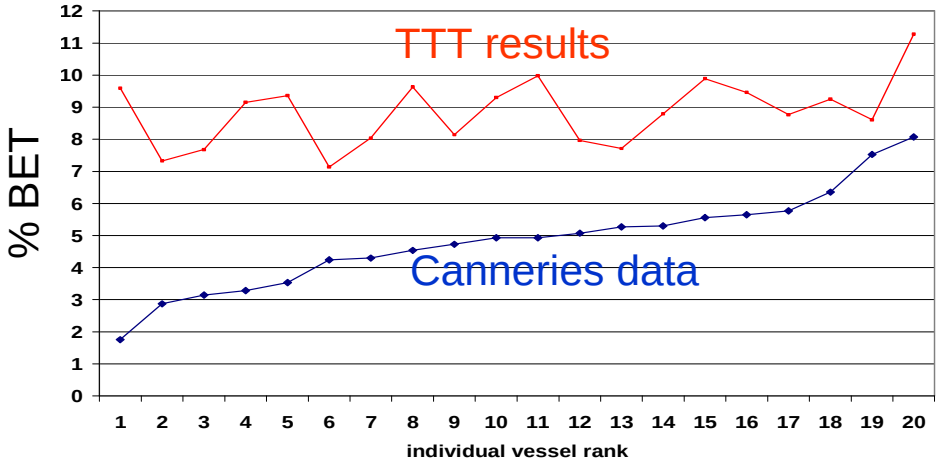
Typical TTT results:
species composition of the PS FAD catches, by flag, 3rd quarter 2002, south Somalia area (total catches 72.500 t.)



DeFinetti plot showing the frequencies of the sampled species composition in the same time & area strata (623 samples)

- The quarterly species composition of the various flags are most often nearly identical (left) for each given fishing mode in all the TTT area,
- While the samples are always showing some variance in their species composition (right), that is never visible in the TTT results of each flag or by vessel
- Then the TTT results are not biased because of its large strata, but probably widely smoothed at variable degrees depending of the variance in the species composition and of tuna sizes caught in the strata

Excessive smoothing of the TTT species composition of each PS catches



Exemple of typical percentages of BET landed by each French PS (2007): canneries (in blue) & TTT data in red (sorted by increasing % of BET in the canneries data)

Typical results observed each year:

- Much more BET in the TTT results: a more realistic average catch of BET,
- But also a very low variance of the TTT BET catches: these BET catches by TTT of each landing have been clearly smoothed by the wide strata used by TTT: then the species coomposition of the catches landed for each vessel cannot be realistic
- This potential problem has been accepted in the original TTT data processing: **TTT has been built to estimate realistic catches by species and CAS of each flag, and not the exact catches of each vessel**.
- The exact species composition of each vessel landing would need a widely different **data processing by TTT**, and probably also more sampling
- It can be noted that the SMS method and its small strata already provides a more realistic variance between boats of the tuna catches taken wthin the same TTT strata but on different tuna concentrations

Recommendations to improve the TTT results, stratified in 3 categories of importance and priority

The present data used in TTT (log books, sampling data and TTT method) are very strong, but the TTT data processing should/could be improved following the following recommandations:

*** * *** Incorporate in the TTT data processing the amounts of **large tunas** landed for each trip from the canneries data, and keep these levels of large tunas in the TTT results, without any raising factor due to small tunas.

*** * *** Improve & reduce the size and improve the selection of **time & area TTT strata**, for instance following the lines of the SMS method (small mobile strata) producing for each flag and vessel more realistic species composition & CAS of the PS catches

******* Do ASAP a **new data processing** of the TTT C/E & CAS statistics at least for recent years (2001-2015) using the corrected **L/W of Chassot** et al 2017 and based on the commercial landings of large YFT.

****** Conduct a careful investigation upon the reliability and potential biases of the **species identification done by canneries on small size tunas** (<10kg), today and in the past, this work being done on Spanish and on French commercial data.

****** Conduct more intensive L/W sampling of the 3 TT, and analyze the **time & space variability of the condition factors** at size of the 3 TT species, and study a potential use of variable L/W relationship in future TTT

***** Conduct experimental « **bin sampling** » by observers (as in the WCPFC area) on some purse seiners, especially on vessels that cannot be sampled at landing (SAPMER fleet)

***** Better evaluate the **potential bias of the today sampling** that are targeting primarily the large sets.

* All small tunas sizes of TT sold on **local markets or discarded** after landing should be included in the TTT data processing and its results (C/E and CAS)

***** Explore alternate data processing methods allowing to better estimate in future TTT the species composition of **each landing**