Note upon difficulties, uncertainties and potential bias in the multispecies sampling and data processing of large tunas (yellowfin, bigeye and albacore) sampled in free schools by the Indian Ocean and Atlantic purse seiners

by Alain Fonteneau, Alain Hervé, Renaud Pianet, Alicia Delgado de Molina and Viveca Nordstrom

Summary

This document analyzes the potential uncertainties and biases in the species composition estimated for the EU purse Seine free school catches in the Indian and Atlantic oceans. When it is generally estimated that the species composition of these catches dominated by tuna at large size is highly reliable (in opposition to FAD catches dominated by small tunas), this paper tend to show in both oceans various sources of uncertainties and bias in these catches by species. Most of these statistical problems could easily be corrected by the concerned scientists developing ad hoc actions, doing some additional biometry sampling and doing correction of past and present data processing programmes.

Résumé

Ce document analyse un certain nombre d'erreurs et de biais potentiels dans la composition spécifique estimée des captures sur bancs libres par les senneurs européens dans les océans Indien et Atlantique. Il est généralement admis que ces prises dominées par des thons de grand tailles sont hautement fiables (en opposition avec les captures de thons de petites tailels sous les DCP). Toutefois, ce document tend à mettre en évidence dans les 2 océans un certain nombre d'incertitudes et de biais dans ces évaluations des prises par espèce. La plupart de ces problèmes peuvent être aisément résolus en développant des actions ad hoc, en particulier des échantillonnages complémentaires de biométrie, et en corrigeant les traitements des données (passés et present).

1-Introduction

The multispecies sampling scheme developed on purse seiner landing in the Atlantic and Indian oceans has been primarily built to sample tuna catches of small sizes, because these small tunas (for instance < 3.2 kg or <10kg) often tend to be misclassified as being skipjack, when they contain significant amount of small yellowfin and small bigeye (simply because yellowfin, skipjack and bigeye <3.2 kg tend to be sold at same prices at small sizes, ICCAT TT WG, Brest 1984). Furthermore, it has been noticed since 1975 that small yellowfin also may contain a significant percentage of small bigeye that can be identified following their careful examination, and only by well trained technicians (small bigeye are never identified from small yellowfin in the log books).

However, the uncertainties in the species composition of the tuna catches are not limited to these catches of small tunas: the comparative analysis of multispecies sampling results, of log books and of landing data, tend to show that there are also serious uncertainties in the catches and sizes by species of large tunas: bigeye, yellowfin and albacore. For instance, significant quantities of large bigeye can be under-estimated by skippers in their log books, when they are sometimes caught in free schools dominated by large yellowfin, for instance simply because of their small percentages caught in the sets: an average of 10% of large bigeye caught in the total catches of large tunas over 10kg during the 1991-2008 period, in both the Atlantic and Indian ocean, see figure 1.

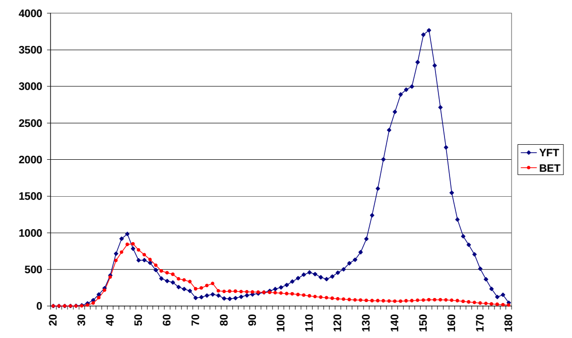

Figure 1a: Total average catch at size of yellowfin and bigeye by purse seiners in the Atlantic (1999-2008). *(Catches over 1m predominantly taken on free schools)*

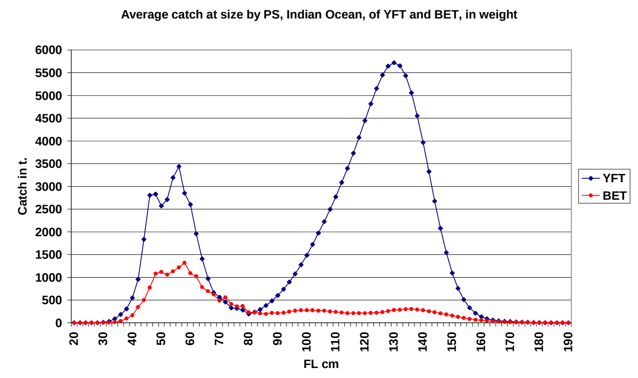

Figure 1[b]~~a~~: Total average catch at size of yellowfin and bigeye by purse seiners in the Indian ocean (1999-2008). *(Catches over 1m predominantly taken on free schools)*

In the same way, albacore are quite often caught with large yellowfin in free schools, but most often in very small percentages, and then widely underestimated in the log books (but very well recorded in the landing data at least for French purse seiners, because they are easily recognized and often sold at a highest market value, compared to all other tropical tunas). Albacore offers good examples of this type of problems: this species has been frequently observed in mixed species free schools, at least in some strata, but most often in small quantities.

This is why the multispecies sampling scheme has been also permanently conducted on the landing of large tunas, but giving more weight to the well maps and landing data when these data are estimated to be reliable. This analysis has been widely based on the analysis of 2 files that are now available for the Indian and Atlantic oceans:

1) the **In&Out file**: based on log books, a file comparing the species composition of each positive set taken on free schools, in the log book and corrected by the TTT software used since 1991 to correct the species composition of the log books
2) the **SPECIES file**, based on multispecies samplings, containing the “observed” species composition in weight of all the samples taken on pure free schools sets (with the exact position, date and characteristics of each sample). This SPECIES file provide a true estimate of the observed composition sample by sample (single large set or combination of homogeneous small sets), without any strata substitution.

This note will examine the peculiar difficulties and uncertainties faced in the sampling and in the data processing of the sampling of these large tunas taken on free schools, trying to estimate their potential bias (in the sampling and/or data processing), its final goal being to propose recommendations allowing to improve the quality of these results, both in the past, and also in the future data processing.

2- Large tunas sampling and data processing

The sampling and data processing schemes of the PS have been always be conducted based on a basic stratification between 2 strata of large and small tunas, the limit between these 2 categories being at 10 kg i.e. around 80 cm of fork length. The total weight of each categories is known at least in theory, from 2 quite independent sources:

(1) log book information and "Well maps" routinely prepared for each well and indirectly for most sets, at least large ones: large yellowfin, large bigeye and large albacore are easily identified set by set by all skippers, but with some potential bias and contradiction between log book and "well maps" information. All the final species composition is estimated for each set and on a set by set basis based on the declared catch at size by species and by size categories (table 1)

(2) landing information obtained from various commercial sources (especially for the French fleet): quantities of large tunas sold, and quantities of each species identified and weighted by stevedores during the landing process. In theory these landing information tend to be more valid than the previous ones, but the information tend to be often disconnected from the set information.

All the data processing has been stratified by types of school (here only the free schools strata has been envisaged), but also by time and area strata and size categories of the fishes (+ or -10kg): as a result, the processing method used estimates for each school a corrected species and size composition based on the multispecies sampling data.

The same information can be available for tuna catches taken in the Atlantic and Indian oceans, as similar sampling schemes have been developed in both oceans since 1980, but their full availability has been quite variable during the period 1980-2008 as a function of the landing port and years of activity. The goal of this paper will be to examine in the run, i.e. during this full period, the consistency between these various data sets and the potential bias and errors faced by this multi species sampling of free schools.

3- Potential problems, uncertainties and bias in the sampling and data processing of free schools catches

3-1- Non random selection bias

The large yellowfin and large bigeye (>10kg) that are typical of free schools tend to be selected/grabbed randomly, and measured by technicians in proportion of their numbers caught. Furthermore, these large yellowfin and large bigeye are often (but not always) well recorded in the log books and in the landing data as the market values of these 2 species has been always quite different (large bigeye being sold at a skipjack price). As large bigeye are always mixed with large yellowfin and most often in small percentages, they may tend to be under declared in the log books (being lost in the yellowfin catches), and these rare bigeye may also be facing a typical selection bias, for instance large fishes being more frequently sampled that their real low proportion in the landed catches (the Y axis scales are different for the 2 oceans because of the much higher catches in the Indian Ocean)

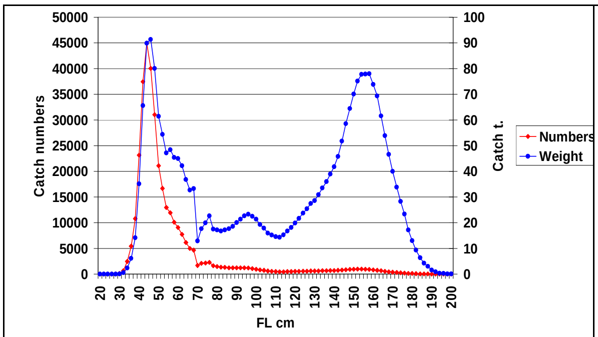

Figure 2a: Average catch at size of bigeye taken by purse seiners on free schools in the Atlantic Ocean (in tons) (1999-2008)

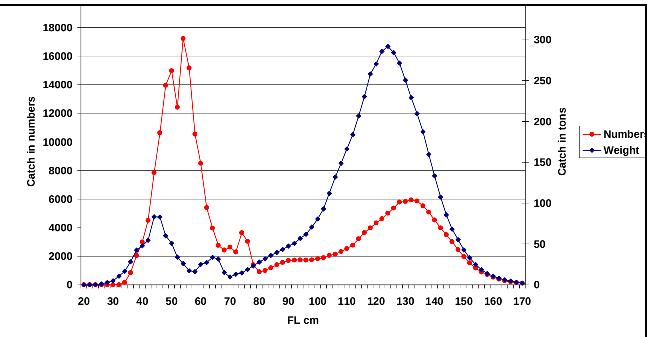

Figure 2b: Average catch at size of bigeye taken by purse seiners on free schools in the Indian Ocean (in tons)

This potential bias of the log books underestimating these rare catches of large tunas, can be estimated and corrected when the landing data provides information on the real total weights of these large bigeye. This has been often the case for part of the Atlantic and Indian Ocean purse seine landings, for instance for French purse seiners for which the SOVETCO landing data of bigeye provides the “real quantities” of large bigeye for the French fleet (an information available for each landing, but not for each set), when catches of small bigeye are + or – always widely underestimated in this landing data set. Figures 3a & 3b show the 3 data sets of bigeye quantities estimated in the Atlantic and Indian oceans:

(i) the original quantities of bigeye taken on free schools by French Purse seiners (widely dominated in weight by large bigeye in both oceans, figure 2) at their log books levels [uncorrected logbooks],

(ii) the quantities of estimated & corrected bigeye catches on free schools obtained by the present data processing [corrected logbooks], and

(iii) the official commercial quantities of bigeye sold by the French fleet [SOVETCO] (dominated by large bigeye >10kg, small bigeye being most often sold as skipjack).

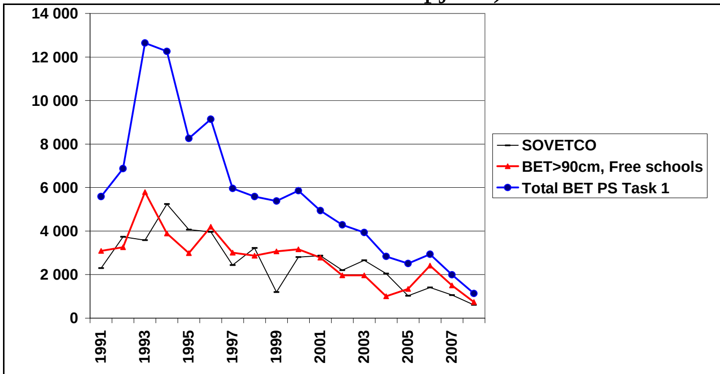

Figure 3 a: Total catches of bigeye by purse seiners in the Atlantic Ocean, in 3 data sets: uncorrected log book data, corrected log books and SOVETCO/marketing data (large bigeye)

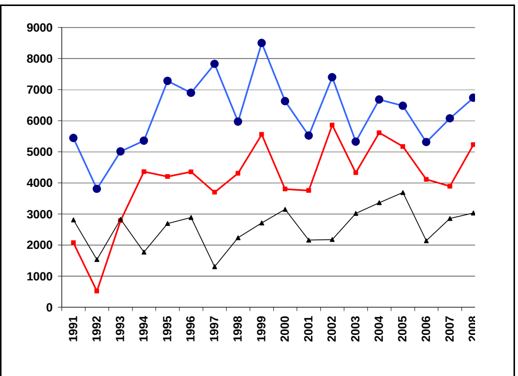

Figure 3 a: Total catches of bigeye by purse seiners in the Indian Ocean, in 3 data sets: log book data, corrected log books and SOVETCO/marketing data (large bigeye)

These results are visible in the basic data when comparing the log book and the corrected data. The levels and trends of the observed differences appear to be quite different in each ocean, and the potential reasons explaining these past changes should be better studied and understood by scientists.

3-2- Uncertainties in the conversion between predorsal to fork length & in the length weight relationship

All large bigeye and large yellowfin (in a range over 10-25 kg, but not at a precise limit, are measured in predorsal length, & not in fork length, in order to facilitate their measurement and to increase the numbers of large tunas measured by technicians (a measurement also allowing to reduce the bias due to the sampling of frozen and curved large tunas). All these pre dorsal length measurements of large fishes are taken by ½ cm size intervals. They are later converted into fork length, based on an historical sample of observed variability of fork length for each class of LD1. In this method, all yellowfin and bigeye landed at medium-large sizes in a range between 10 to 25 kg can be measured either in predorsal length, or in fork length, simply as a function of landing & sampling conditions. The weight of these fork length distributions is later estimated for all tuna species using a given fixed length-weight relationship (constant over time, and different between oceans).

This method has been used because the sampled fishes cannot be weighted by a routine sampling scheme. However, it should be recognized that this method is partly dependent of the validity of the 2 basic relationships, LD1-FL and length-weight, and also of changes in the condition factors (as lean or fat tunas have in the present data processing the same length weight relationship[1]). The examination of the present relationship would tend to indicate that the relationship presently used for large bigeye tend to be quite limited in numbers of large fishes sampled, especially in the Indian Ocean. However, any significant errors/weaknesses in these length-weight bigeye relationship would not subsequently have cascading impacts on the sampled species composition, but simply on the estimated numbers of bigeye landed (as the total catches of large bigeye are mainly fixed since 2005 but not before, from the log book information). During the earlier pre 2005 period, these changes in the bigeye length weight relationship should modify the subsequent species composition of free schools.

These potential errors should be better explored and estimated, covering the entire period 1980 to now, and these uncertainties should preferably be reduced, improving the quality and coverage of the 2 relationship used (LD1/FL and Length-Weight).

3-3- Data processing and estimates of quantities of large bigeye caught:

3-3-1- Overall question

In the data processing of the large bigeye data, it can be noted that there are major time heterogeneities in the data processing used:

- ✓ During the 1980-1990 period: the species composition of these large tunas (>10kg) was determined by the estimated species composition obtained by the sampling, but using questionable statistical methods (KOURPAS method by JP Hallier in the Indian Ocean and CORALPS method by Fonteneau in the Atlantic). Furthermore, in the Atlantic, all the present data before 1990 have been computerized and processed without a proper stratification between FAD and Free schools. However, this lack of information on the FAD associated schools before 1991 may not be a critical, as the analysis of log book data during the period has shown that FAD associated catches were limited during this historical period, for instance with only

[1] Condition factors : their variability should be better analysed and if necessary taken into account in the future data processing.

17% of FAD associated catches in the Atlantic during the period 1987-1990 (Ariz et al 1999).

- ✓ During the 1991- 2004 period: the species composition of these large tunas (>10kg) was determined using an improved software (TTT) and based on the estimated species composition obtained by the sampling of free and FAD schools.
- ✓ Since 2005: the species composition of these large tunas (>10kg) has been strictly fixed as a function of the estimated weights of bigeye, at least from the log books and preferably confirmed or corrected by the well maps or the landing data.

During the 2 last periods, in each set without any identified bigeye catches, the data processing method used has been estimating an unrecorded average weight of large bigeye, based on the average % of bigeye taken in each time and area strata, and for the school type and size category.

As a consequence of this heterogeneity and improvement in the data processing method used since 1980, it would appear that there is an unknown and variable uncertainty in the estimated species composition of these large tunas. This structural uncertainty should be better evaluated and reduced using a new homogeneous and optimized data processing.

3-3-2- Examples of the difficulties, problems and potential bias in the estimated quantities of large bigeye caught on free schools

The comparison of species composition of free schools as reported in the log book and in the SPECIES sampling most often offers striking differences, especially concerning bigeye catches. As an example, taken in the Chagos area during the 1st quarter of 2008:

- The EU log books show a quite low percentage of bigeye: 6.2 % ,
- Much lower than the estimated corrected bigeye catches submitted to the IOTC (14.9 %) for this strata.
- When the basic SPECIES file (samples) shows a lower average % of bigeye of 10.2 %

These major changes in the amount of bigeye caught are quite surprising, as these tunas at large fishes (figure 4), average weight estimated at 48 kg, when these large bigeye are easily and well identified by the skippers and in the landing data.

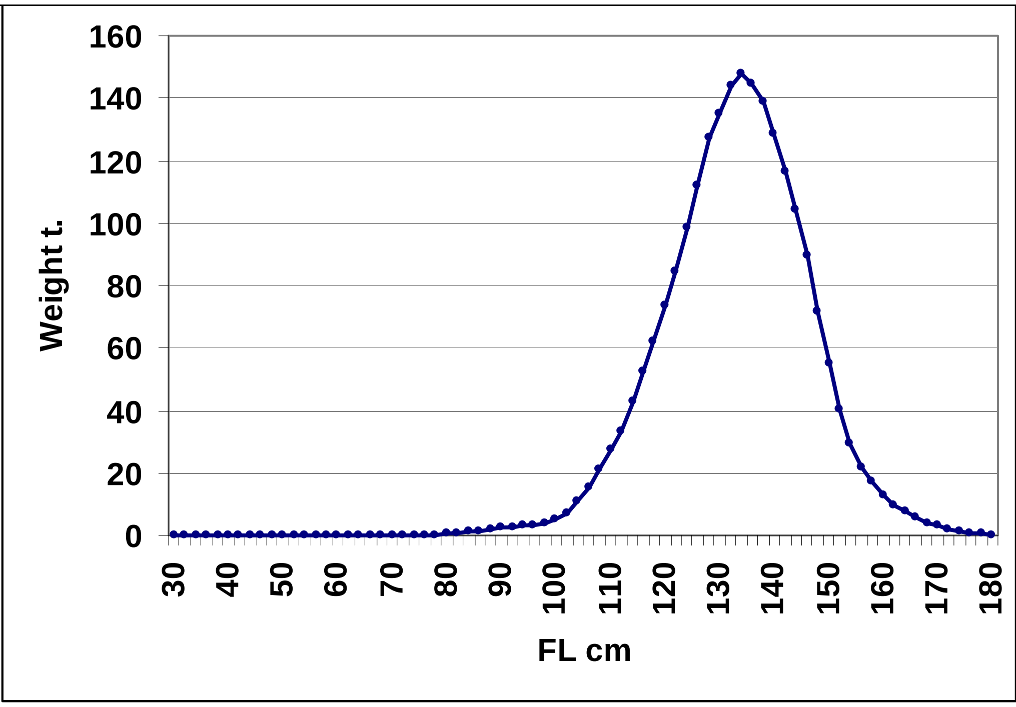

Figure 4: Catch at size of bigeye (in weight) in the studied stratum (1st quarter 2008, Chagos, free schools)

This pattern tends to be frequently observed in most/all strata. In the present data processing, the corrected catches of bigeye are obtained by a combination of 2 methods:

1) Bigeye amount kept as in the log book when mixed catches of large yellowfin and of large bigeye, or of large bigeye alone, are reported for a given set.
2) In all other types of species and size compositions declared even without large bigeye, the sets will "receive" in the data processing the "average" species composition of the strata (area, quarter, fishing mode and size category) taken from the sampled fishes.

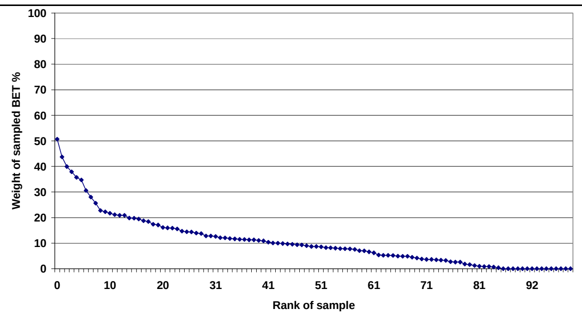

Figure 5: % of bigeye in the sampled weight of free schools sets in the Chagos area Jan and Feb 2008, sorted by declining % of bigeye (108 samples) , average=10.2%

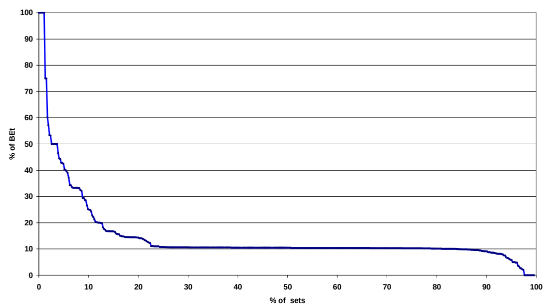

Figure 6: % of bigeye in all the free schools sets taken in the Chagos area Jan and Feb 2008, after correction of their species composition & sorted by declining % of bigeye

The first method most often corresponds to large amounts of large bigeye that are easily and well identified by the skippers, when the second method tend to compensate for the lost bigeye, small quantities not reported in the log books: the sampled average species composition of the strata is then applied to these missing bigeye: in this example an average percentage of 10% obtained from the above figure (figure 5), these 10% of estimated large bigeye being easily seen on figure 6 (raking from 23 to 90% of the sets).

The validity of this method can well be questioned: the proportion of large bigeye in the sets showing large % of bigeye appears to be consistent between the log book and the species sampling. On the opposite, all sets without reported bigeye, when they were probably containing a small percentage of bigeye, as most sampled sets in the example strata, will be assumed to carry an average 10.5 % rate of large bigeye. It may well be considered that this average rate is widely excessive, and in this case the average rate of about 3% observed for the 50 % of the samples showing the lower percentage of bigeye should preferably be used to estimated bigeye catches in the sets without reported bigeye. In this example, when a total catch of bigeye = 1150 tons has been estimated from declared yellowfin catches and a 10.5 % rate of bigeye, this bigeye catch would be reduced to about 400 t. assuming a more realistic lower rate of 3% of bigeye. And these "lost bigeye" would be converted into yellowfin catches.

Clearly this potential bias has not been well studied and it should be carefully examined and solved by concerned scientists, as it may significantly overestimate the bigeye catches taken on free schools.

Furthermore, a peculiar and rather "stupid" problem identified in the present analysis has been showing a peculiarity and an error in the present data processing of the species composition: it has been presently accepted that when a mixed catch of large bigeye (or large yellowfin) and of any quantity of albacore or skipjack was reported in the log books for this set, the albacore catch stay identical, when the catches of large bigeye and of large yellowfin is "transformed" into the average species composition of the strata (time:/area/fishing mode), simply because of this additional skipjack or albacore.

In such cases, large bigeye tend to be converted into large yellowfin, even when their original species was perfectly identified by the skipper and confirmed by the landing/sampling data. This technical error has been seldom found in the past data processing, but it may have produced yearly a loss of several hundred of tons of large bigeye. These quantities of false species correction should of course be fully estimated and corrected in future data processing.

Our recommendation is that when large bigeye tunas have been identified in the log books, and later confirmed by the species sampling and the commercial data, they should be fully "frozen" and kept in the species correction and never converted as being yellowfin.

3-4- Data processing and estimates of quantities of large albacore caught:

The albacore catches taken by the EU purse seiners are simply the total amount of this species as they have been reported in the log book, and without correction. However, the comparison of the albacore catches reported in the French log books and in the corresponding commercial landing data (a data set available for each landing) tend to indicate that the albacore catches reported in the Indian Ocean (and surprisingly not in the Atlantic) in the log books tend to be 22.5% lower than the albacore catches sold by the same fleet during the 1990-2008 period. This underestimation of albacore catches in the log book is also well shown comparing the SPECIES sampling and the log book data of the French fleet: when in a typical strata where albacore is commonly caught is taken as an example - the Chagos area 1st quarter of 2008 - albacore have been identified in 40%

of the free schools samples, when in the log book they have been reported in only 10% of the free schools positive sets.

It can be considered that lack of reporting is not at all surprising, or even legitimate for the skippers, simply because free schools of moderate sizes, for instance 20 tons (50% of the average set) with an average percentage of albacore in the Chagos area, 1st quarter 2008 would carry 3% of albacore, then an average weight of only 600 kg, a small amount of albacore mixed with the 19.4 tons of yellowfin/bigeye. Clearly, these ranges of small catches cannot be well identified during the fishing operations, while they are perfectly well identified when the combined albacore catches of the trip are landed and sold to the international market.

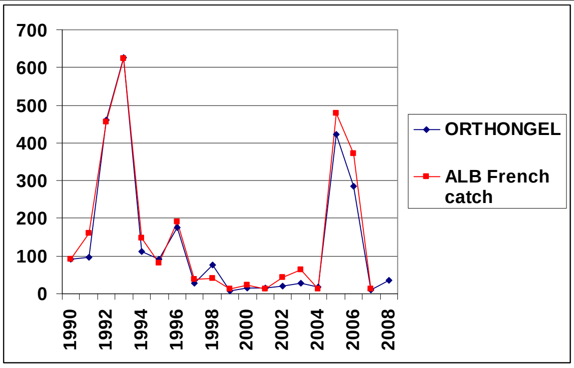

Figure 8a: Total catches of albacore by purse seiners in the Atlantic Ocean, in 2 data sets: log book data and SOVETCO/marketing data

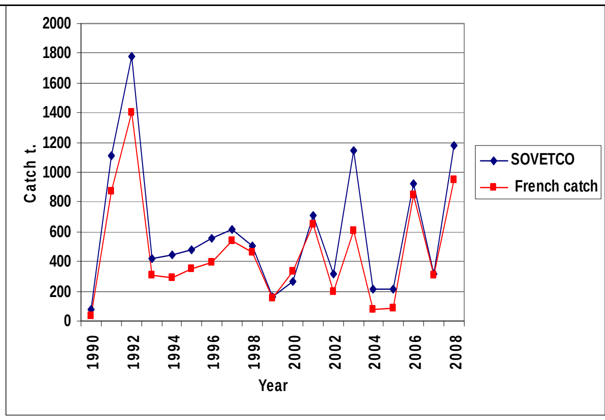

Figure 8b: Total catches of albacore by purse seiners in the Indian Ocean, in 2 data sets: log book data and SOVETCO/marketing data

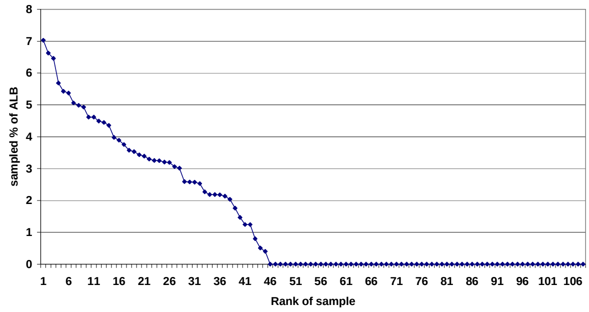

Figure 9: % of albacore in the sampled weight of free schools sets in the Chagos area Jan and Feb 2008, sorted by declining % of albacore weight (108 samples)

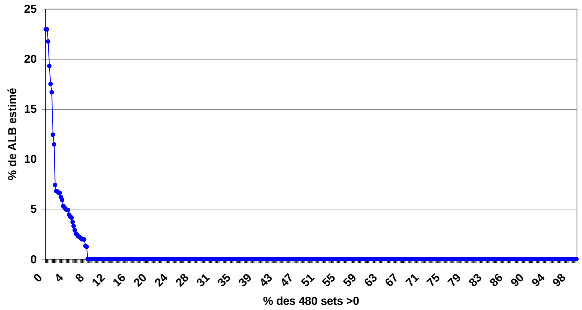

Figure 10: % of albacore in all the free schools sets taken in the Chagos area Jan and Feb 2008, sorted by declining % of albacore

As a conclusion, the small amounts of albacore caught by tropical PS, based on log book data, and declared to the RFO appears to be often slightly underestimated in the present data processing, at least in the Indian Ocean. It should be recommended that when commercial data on albacore catches have been identified, these commercial catches should be used to force and to keep “frozen” the albacore catches used for each trip in its data processing.

4- Conclusion and recommendation

As a conclusion, there is no doubt that the multispecies sampling scheme routinely developed since 1979 in the Atlantic and Indian oceans on the EU purse seiners has been very positive in terms of its wide range of consistent results that have been very useful to simultaneously estimate the species and size composition of the PS landings. These results are of course very interesting for small tunas (for which this sampling scheme has been built since 1979), but also for large tunas that are presently caught in the free schools fishery.

However this quick review of the results and of the data processing method used indicates that this question would need more analysis and also various corrective actions allowing to improve the validity of its estimated results, targeting both the improvement of past and future data. Additional in depth analysis and comparisons of the 2 files used in the present analysis, the In&Out file and the SPECIES files, should for instance be conducted in order to identify other potential problems in the data processing.

One of the conclusion clearly emerging from this data analysis is that when given catches of large tunas (yellowfin, bigeye and albacore) have been well identified by the skipper and confirmed by the landing data, these catches should be maintained “as much as possible” in the data processing. Significant improvements in the data processing of both historical and future data could also easily be obtained, simply improving the quality and sampling basis of the LD1/FL and length weight relationship used for bigeye and albacore, and also for yellowfin.

Bibliography

Bard, F. & Vendeville, P. (1986), Note sur l'échantillonnage plurispécifique des thons tropicaux au port d'Abidjan, Col. Vol. Sci. Pap. ICCAT 25, 37-45.

Cayré, P. (1984), Procédure suivie pour la révision de la composition spécifique des statistiques thonières FISM (France, Côte d'Ivoire, Sénégal et Maroc), Col. Vol. Sci. Pap. ICCAT 21(2), 102-107.

Fonteneau, A., Chassot, E., Abascal, F. & Ortega, S. (in press), Potential bias in multispecies sampling of purse seiner catches, ICCAT Col. Vol. Sci. Pap. SCRS 2008/162.

Lechauve, J-J. (1999), AVDTH98. Acquisition et validation des données de pêche au thon tropical (84), Technical report, Institut de Recherche pour le Développement.

Pallarés, P. & Hallier, J-P. (1997), Analyse du schéma d'échantillonnage multispécifique des thonidés tropicaux (Programme n° 95/37), Technical report, IEO/ORSTOM.

Pallarés, P. & Nordström, V. (1997), Analisis del esquema de muestro multiespecifico de los tunidos tropicales: presentacion des proyecto y primeros resultados, Col. Vol. Sci. Pap. ICCAT 46(4), 168-174.

Pallarés, P. & Petit, C. (1998), Tropical tunas: new sampling and data processing strategy for estimating the composition of catches by species and sizes, Col. Vol. Sci. Pap. ICCAT 48(2), 230-246.

Pianet, R. (1999), Evolution du système de collecte et de traitement des données de la pêche thonière des senneurs européens et assimilés de 1981 à 1998, IOTC Proceedings 2, 74-96.

Pianet, R.; Pallarés, P. & Petit, C. (2000), New sampling and data processing strategy for estimating the composition of catches by species and sizes in the European purse seine tropical tuna fisheries, IOTC Proceedings 3, 104-139.

Table 1: Size categories potentially used in the log books in order to classify the catch at size by species

cat	YFT	cat	SKJ	cat	BET
1	YFT < 3 kg	1	SKJ < 1.8 kg	1	BET < 3 kg
2	YFT 3 - 10 kg	2	SKJ > 1.8 kg	2	BET 3 to 10 kg
3	YFT 10 - 30 kg	3	SKJ 1.8 to 4 kg	3	BET 10 to 30 kg
4	YFT 6 to 20 kg	4	SKJ 1.8 to 6 kg	4	BET 6 to 20 kg
5	YFT 30 to 50 kg	5	SKJ 4 to 6 kg	5	BET 30 to 50 kg
6	YFT 20 to 40 kg	6	SKJ 4 to 8 kg	6	BET 20 to 40 kg
7	YFT > 50 kg	7	SKJ 6 to 8 kg	7	BET > 50 kg
8	YFT 40 to 60 kg	8	SKJ > 8 kg	8	BET 40 to 60 kg
9	YFT unknow size?	9	SKJ unknown	9	BET unknown?
10	YFT < 10 kg			10	BET <10 kg
11	YFT > 10 kg			11	BET > 10 kg
12	YFT 10 to 30 kg			12	BET 10 to 30 kg
13	YFT > 30 kg			13	BET > 30 kg