

**NEW SAMPLING AND DATA PROCESSING STRATEGY FOR ESTIMATING THE  
COMPOSITION OF CATCHES BY SPECIES AND SIZES IN THE EUROPEAN  
PURSE SEINE TROPICAL TUNA FISHERIES**

by

R. Pianet<sup>1</sup>, P. Pallarés <sup>2</sup> and Ch. Petit<sup>3</sup>

<sup>1</sup> ORSTOM-HEA BP 5045. 34032 - Montpellier. Cedex 01. France

<sup>2</sup> Instituto Español de Oceanografía. Corazón de María ,8. 28002 - Madrid. España

<sup>3</sup> ORSTOM-HEA BP 5045. 34032 - Montpellier. Cedex 01. France

## **Abstract**

The difficulties associated with tropical tuna sampling strategies, particularly those resulting from the large-scale adoption of the method of fishing on logs, have caused the Indian Ocean tuna scientists concern since the beginning of the purse seine fishery, as well as the ICCAT SCRS (Standing Committee for Research and Statistics of the International Commission for the Conservation of Atlantic Tunas) in recent years. Upon the recommendation of this Committee, in 1996 Spain and France embarked on a European Union project aiming to study the consequences of this new fishing tactic on the estimated catch by species and by size, and to implement a new sampling strategy so as to improve the accuracy of the statistics.

Many specific analyses have been necessary because of the complexity of the problem. In this regard, close attention has been given to the influence of different factors (place, time, fishing associations, flags, vessel class, etc.) on the breakdown of catches by species and sizes, the sampling coverage rate, the sample size, the unit of sampling, etc. In this document the results of the different analyses conducted are described and discussed, and a new sampling system for both east Atlantic and Indian Ocean tropical tuna purse seine catch sampling is proposed. This document is mainly focussed on Indian Ocean, while another document (SCRS/97/28) - presented at the last year's ICCAT meeting – was aimed to the Atlantic Ocean.

## **Résumé**

Les difficultés soulevées par l'échantillonnage des senneurs tropicaux - et plus particulièrement celles qui résultent de la stratégie d'utilisation à grande échelle de la pêche sur épaves – ont posé problème aux chercheurs de l'océan Indien dès le début de la pêcherie, aussi bien qu'aux scientifiques travaillant dans le cadre du SCRS de l'ICCAT (Comité Permanent pour la Recherche et les Statistiques de la Commission pour la Conservation des Thonidés de l'Atlantique) dans les années récentes. Sur la recommandation de ce Comité, la France et l'Espagne ont proposé en 1996 pour financement par l'Union Européenne un projet d'étude ayant pour objectif d'étudier les conséquences de cette nouvelle tactique de pêche sur les estimations de la composition en tailles et en espèce des captures, ainsi que de proposer une nouvelle stratégie d'échantillonnage dans le but d'améliorer la précision de ces estimations.

De nombreuses analyses spécifiques ont été nécessaires en raison de la complexité du problème. À cet effet, une attention particulière a été portée à l'influence de différents facteurs (zone, période, association, pavillon, catégorie de taille des thoniers, etc.) sur la répartition par espèce et en taille des prises, le taux de couverture de l'échantillonnage, l'unité d'échantillonnage, etc.. La méthodologie ainsi que les résultats des principales analyses sont décrits et discutés dans ce document, et une nouvelle procédure d'échantillonnage des prises des senneurs tropicaux est proposée pour les deux océans. Ce document est essentiellement centré sur l'océan Indien, tandis qu'un document semblable (SCRS/97/28) présenté l'année dernière à l'ICCAT était lui centré sur l'océan Atlantique.

## 1- INTRODUCTION

The multi-species nature of tropical tuna surface fisheries gives rise to a series of difficulties at the time of estimating basic catch by species and catch by size statistics. The Working Group on Juvenile Tropical Tunas (meeting at Brest in 1984) observed the existence of bias in captains' reports on catches by species in the Atlantic. It consequently designed a procedure (Anon., 1984) for the correction of the breakdown by species entered in logbooks based on the multi-species sampling system (Bard and Vendeville, 1985) which had, since 1980, been used in the sampling of the catches of the surface fleets (purse seine and, to a lesser extent, Ghana baitboats) of the east Atlantic. The method thus developed has provided more accurate statistics on the estimate of the species composition those fisheries.

Since 1980 purse seine catches have represented the principal component of both Atlantic and Indian Ocean catches of yellowfin and skipjack, and a considerable proportion (from 15 to 20%) of those of bigeye. A substantial proportion of the yellowfin and virtually all the bigeye in those catches are juveniles. As the bias observed in the reports delivered by captains on catches by species affected, above all, specimens of small sizes, it is necessary to correct these biased data in order to decrease the degree of uncertainty in the final estimates.

This problem of bigeye and yellowfin misidentification in the landing statistics was discussed first by the ICCAT SCRS in 1975 (Fonteneau 1976). Following these discussion a multispecies sampling scheme was established in the Atlantic on the landing of all purse seiners. Following this implementation, the ICCAT's Working group on Juvenile Tropical Tunas held in Brest, 1984, analysed the results of this new sampling scheme and fully confirmed that there was large bias in the species composition of both the log books and in the landing commercial figures. This first multispecies sampling was developed in the Atlantic in 1979; it was a size sampling done in proportion of the weight of each species, with a stratification of sampling by large size categories (well coded in the logbooks)

Up to 1990, the traditional fishing strategy had been followed, but since 1991 fishing on log schools developed considerably (Ariz et al., 1993, 1996; Bard et al., 1985; Fonteneau, 1993; Pallarés et al., 1995). This phenomenon was soon identified in the Indian Ocean, where a specific sampling strategy allowing to separate log and free schools catches was developed (Hallier, 1985, 1991) in order to improve the estimate of species and size composition of the purse seine catch. This leads to some divergence in the processing methodology between Atlantic and Indian Ocean fisheries, both dominated by the European Union fleets, as it was clearly confirmed at the ORSTOM workshop held in Paris (Pianet, 1995).

In order to preserve the quality of these data bases as well as to standardize the procedures in both oceans, Spain and France - the countries most closely involved in the Atlantic and Indian Ocean purse seine fisheries - presented a joint project to the European Union (Pallarés and Norstrom, 1997) to undertake the analysis and design of a new sampling and processing system for surface fleet tropical tuna catches. Within the framework of that project, named "Analysis of the Tropical Tuna Multi-species Sampling Scheme", the appropriate studies were performed for the development of a new sampling and data processing system for the purse seine fleet catches in the Indian and east Atlantic oceans.

This paper will essentially deals with the results of this program usable in the Indian Ocean.

## **2- MATERIAL AND METHODS**

The studies were carried out on the basis of the data recorded on the east Atlantic and Indian Ocean purse seine fleet catches and size distributions between 1991 and 1995.

They extended to the French and Spanish fleets and also the "NEI" (not elsewhere included) fleet which includes vessels of the aforementioned nationalities, in their origins, fishing under flags of convenience. The data came from logbooks providing detailed information on catches by set, and efforts. Information on types of fishery association – already collected on a routine basis - was added to the usual data on position, species, commercial category, etc. entered in logbooks, thus allowing analyses to distinguish log and free schools sets.

Although the problems associated with the European Union purse seine fleets are shared by other surface fleets (as Tema based baitboats in the Atlantic or Japanese purse seiners in the Indian Ocean) the analysis was restricted to them for reasons of availability and quality of data. The resulting sampling method and data processing system may, nonetheless, easily be adapted for use in relation to those other similarly affected fisheries.

The analyses for determining the sample size variability according to species composition and/or size throughout the landing and other analyses of the accuracy of the estimators were undertaken on the basis of two experience of an intensive sampling (one in both oceans), involving the systematic sampling of the whole well, through the taking of samples of equal size throughout the landing. The data corresponding to a super-sampling process undertaken at the ports of Abidjan (in March 1997, Pallarés et al., 1997) and Victoria (April 1997) were studied for that purpose.

In view of the complex problems affecting the sampling of this fishery (multi-species nature, simultaneous estimate of catch by species and by size, different types of fishing associations, marked seasonality, etc.), many different analyses, with a specific methodology for each type of problem, were necessary. The methodology used for dealing with each of those problems is described below.

### **2.1- SCHOOL TYPOLOGY**

The type of school is the major factor to be considered for of both sizes and species composition analysis. A difficulty faced in this area is the question of how to process catch and sampling data which have no association code (this problem applying mainly to the Atlantic ocean). Different methods were developed to answer that problem.

#### **2.1.1- Assignment of association codes drawn from logbook sighting codes**

Sets in which sighting cue codes (birds, etc.) were entered, but without any association codes were assessed to be free schools. As following studies confirmed that the breakdown of those sets by species and size was closer to free schools than to schools associated with objects, a two-step procedure was developed: 1) assignment of a newly created logbook code to sets; 2) assignment of codes to the samples corresponding to those sets.

#### **2.1.2- Creation of a school typology**

A school typology was created (Fonteneau and Pallarés, 1997), using two variables, which made it possible automatically to assign the association code to sets which did not carry that information. Only the samples in which the type of association was clearly identified entered into consideration.

Samples with codes drawn from logbook sighting indices were excluded, as well as sets made over submarine mountains.

The variables were the following:

- the average weight of sample, taking all species into account;
- a diversity index, specifically the Shannon-Weaver index which reflects all the species present and their relative abundance.

The analyses were undertaken in the light of the two variables jointly and in that of each separately. In the latter case the accumulated frequencies were observed in order to determine whether the factors were decisive. The average weights were calculated according to the classical size-weight relationships accepted by SCRS for each species.

Analysis was done for three areas in the Atlantic ocean and five areas in the Indian Ocean, although set type was already included in analysis for the latter.

### **2.1.3- Classification of association codes in two categories: log and free school**

A more general question than that of assigning codes to the unknowns is that of classifying, as far as possible, the secondary and unknown groups of associations (marine mammals, whale-shark, under a vessel, unknown and sighting cue codes, ...) within the two broad groups of reference: logs and free.

As it has been previously said, schools with a sighting code are, on average, closer to free schools than schools under objects. Nonetheless, they form a highly heterogeneous category comprising a significant portion of free schools but also, without doubt, schools under objects and others. A certain number of mistakes are therefore made when schools with a sighting code are automatically classified as free schools. This, among other things, is what has been revealed by the two analyses described below.

In these two analyses work was conducted on the species composition with a component for yellowfin size factors. The analyses thus relate to four variables, namely the proportions of large and small yellowfin, bigeye and skipjack, the sharing between small and large yellowfin being established in such a way that a maximum of information might be obtained on the variation in size according to the association.

- Box-plots facilitated a description of the seven association types by the four variables entering into the analysis, showing the variance within each groups;
- Descriptive discriminant factorial analysis allowed to display the seven association types according to the two combinations of variables which best separate them.
- Decisional discriminant factorial analysis has yet to be undertaken. The plan is to carry out a discriminant factorial analysis in relation to the two broad groups of reference (logs and free school) in view to obtain the most discriminant linear combination variable of the source variables. Of the possible departure variables, the Shannon index and that of average weights described above would be included. The unknown schools and perhaps the sighting code schools would then be assigned to log or free school according to the respective variations of the value of the variable created with respect to the mean values of the two reference groups.

## **2.2- SUPER-SAMPLING**

### **2.2.1- Variability throughout landing**

The analyses were conducted on the basis of the data gathered in the super-sampling process in Abidjan. Size variability was studied by means of ANOVAs of average sizes. Small (<10 k) and large (>10 k) specimens of yellowfin and bigeye were dealt with separately, given the clearly bimodal nature of their distribution and the different treatment given to specimens, according to size (large or small), in the loading and unloading of wells.

The variability of the catch throughout landing was analyzed using the method developed by Snedecor and Irwin (1933), also known as the Brandt and Snedecor method, for comparing the homogeneity of more than two percentages or proportions.

### **2.2.2- Sample size**

In order to check the accuracy of estimates and determine the appropriate sample size the following procedure was developed:

- the well was reconstructed (size structures of catches of all species) on the basis of the size distribution and species composition results obtained in the supersampling experience;
- in that effect, catches by species of small specimens were estimated and their size distributions calculated, while large tunas size distribution was calculated on the basis of landing data;
- multiple multi-species samples of equal size were taken at random without replacement;
- the parameters (average and variance) of the distribution parameters (average size and species proportion) were estimated;
- error thresholds were set and the corresponding sample sizes calculated.

## **2.3- UNIT OF SAMPLING**

The fishing unit is the set, and the sampling one is the wells of the vessel which can contain either part of a set or several of them. It is therefore necessary to choose between the well or the set as unit of sampling in order to determine the weighting to be used to raise the sample, well or set weight. If the biological unit we want to sample is the set, the accessible unit on the vessel which is landing is the well. If it is established that the well fairly represents the set, it may be considered that the sets as a whole may be sampled through the well and the set may thus be taken as the unit of sampling.

In that purpose, and considering that the knowledge of the date of fishing allows to identify the set, the variance which remains when the variance explained by the date of fishing is taken away is the variance between samples of the same set (coming from different wells). Depending on whether that remaining variance is or is not weak, it may or may not be considered that the well fairly represents the set.

## **2.4- INFLUENCE OF DIFFERENT FACTORS: STRATIFICATION AND SUBSTITUTION**

Knowledge of the influence of different factors on species composition and size distribution makes it possible to determine those to be taken into consideration in the definition of the strata which will underlie the breakdown calculations and also, roughly, the degree of accuracy according to time (fortnight, month, bimonth, quarter ...). Then, strata are redefined in such a way that the catch may be as homogeneous as possible within each one. Next, the degree of accuracy of the different

factors is set, and the order of substitutions established according to the respective bearings of the factors entering into the definition of the strata.

Size distribution (percentages of small, medium and large tunas) and species composition (percentages of yellowfin, bigeye and skipjack) analysis have been analyzed separately, and the interplay of the different factors - interactions, crossovers, etc. - was also studied.

The factors or combinations of factors which best explain the variance regarding the number of strata created (the number of modalities) were sought out, the potential qualitative variables being area, season, year, fleet (or country) and class of vessel. The volume of the sets of data did not allowed this issue to be addressed in a global and systematic way; nonetheless the analysis was able to determine factors to be taken into consideration with their approximate degree of accuracy with respect to time.

The substitution scheme was thus established in the following manner: the variance explained by the model under consideration, after one of the factors has been removed, is observed, and this process is repeated for all the factors of the model. The model with least influence is then selected and the missing factor addressed. That is the factor of least influence and substitutions will be made primarily according to that factor, secondarily according to the factor of least influence of the remaining three, etc.

## **2.5- DETERMINATION OF THE MOST HOMOGENEOUS AREAS**

As the spatial factor studied in the preceding section was considered to have a great influence and relatively homogeneous areas appeared in the data mapping, it was considered important that areas entering in the stratification be reevaluated in such a way that they might be as homogeneous as possible.

The first step within the definition of homogeneous area strata was to carry out a  $1^\circ \times 1^\circ$  square using Ward classification, whose principle is to define classes in such a way that variance within classes is minimum and variance between classes maximum. Several criteria prove of assistance for the purpose of deciding the number of classes to consider. The most used is the semi-spatial-square, which indicates the loss of variance between classes on moving from  $n+1$  classes to  $n$  classes. Results are interpreted through principal component analysis, which gives the components explaining maximum variance from departure variables. This work was performed for the two types of association - log and free schools - given that the best results are obtained when the associations are addressed separately.

The second step was to map the results; although the defined classes are certainly to be discontinuous, some homogeneous areas may be seen to emerge. On the strength of these results and on that of the mapping of species composition and size distribution, the homogeneous areas may be redesigned.

The third step was to seek out more exact limits and sub zones for size distributions as generally areas defined in the second step analysis do not provide a good explanation for variance when they are combined with the other factors considered (year, quarter and association) implied in the definition of areas.

The whole area or part of it (for example, a predefined area) is considered, in which we observe the total variance at a  $1^\circ \times 1^\circ$  square level. Then, the field is shared into two parts (for example, along its latitude), and the variance between the two sub zones reflects the fall in variance inside those

zones: the greater is the inter-zone variance, the more the two sub zones are homogeneous. An algorithm have been created to calculate, according to the chosen limit, the inter-zone variance and to cause the limit to vary automatically in order to find the best that will reduces to the minimum variance within the zones. The latitudinal and longitudinal limits to vary, one after the other, within the algorithm. The operation is repeated as many times as necessary, up to the point where the variance shown does not improve or the number of strata takes on greater importance.

### 3- RESULTS AND DISCUSSION

#### 3.1- SAMPLING COVERAGE RATE

The sampling coverage rate was examined in both Indian and Atlantic Ocean, according to different criteria : number of non-sampled strata, strata with less than 1 fish sampled per ton caught, or a coverage rate less than 1/250 in weight). The analysis was performed on a 5° square by month basis, for each year and school association type (log and free schools).

General information on the sampling coverage for both log and free schools from 1991 to 1996 are presented in Tab.1. In order to evaluate the quality of the sampling scheme, distributions of the sampling coverage according to 5 strata classified on the total catch by school type where plotted, using the Box and Whisker technique (Fig. 1), and the sampled weight versus total catch for all strata plotted on Fig. 2.

The analysis put in evidence that – even their number may appears to be quite high – most of the strata which were not or badly sampled are related to small catches. The coverage is good as a whole, and this all the more than it is the free schools – having the smaller variance – which are the less sampled.

Nevertheless, analysis indicated that (even if the overall sampling coverage is quite good in both ocean) it appears to be necessary – in order to avoid to get too many insufficiently or non-sampled strata – to define larger and more homogeneous strata, taking into account the school type, and having the possibility to be different for size and species composition estimates. The definition of homogeneous strata as large as possible was highly recommended.

#### 3.2- SCHOOL TYPOLOGY

The defined typology implied low species diversity figures and high average weights as the criteria for the definition of free schools and high diversity indices with low average weights for log schools. The variables were then considered separately, and results for free and logs schools for the five zones covered are reported on Fig. 3 a-e for accumulated frequencies of the diversity indices, and Fig. 4 a-e for average weights. On a general basis, this method was shown to be quite efficient to separate log and free school in the whole Indian Ocean, and the criteria of the mean weight was most of the time sufficient for discriminating them, except in the Somali and Mozambique channel areas where the diversity index is more efficient; retained values are as follow:

Area	Diversity index : limit (error)	Mean weight : limit (error)
Somalia	0,7 (8% log, 12% free)	
Western Seychelles		9 kg (9% log, 4% free)
Eastern Seychelles		8 kg (11% log, 7% free)
South Seychelles		9kg (5% log, 8% free)
Mozambique channel	0,4 (13% log, 20% free)	



Trials were also made without success to define similar criteria to separate log and free schools from the logbook information. The main reason is the uncertainty on the species composition regarding the diversity index, the use of large and overlapping weight categories.

### **3.3- CLASSIFICATION OF ASSOCIATION CODES IN TWO CATEGORIES : LOG AND FREE SCHOOL**

This study was specifically aimed on Atlantic ocean, as generally this information was directly taken in account in the Indian Ocean. Details of this study are described in the original report as well as in a document presented at the SCRS meeting of ICCAT in 1997 (Pallarés and Petit, SCRS 97/28).

Six association codes were retained for the study: log (natural or artificial), free schools (no associations except birds), marine mammal (whale, dolphin), shark whale, under a boat, and no association or unknown.

Box plot analysis shows a very significant variance in sighting code group, and even more in the unknown group, leading to the conclusion that true “unknowns” come from all association types and thus justifies their classification according to their resemblance to the two broad reference groups - log and free- although some mixing is still observed. A fairly close resemblance is also observed between code appearance group and free schools. Variance in the minor categories – “marine mammals”, “whale-shark” and “under a vessel” - is likewise significant and the “under a vessel” group comes close to the log schools.

The canonical discriminant analysis clearly shows, that the “whale-shark” and “under a vessel” associations are linked to the log group, while the “marine mammal”, “unknown” and sighting code associations are linked to the free school group. The sighting code category is clearly closer to the free schools than are the unknowns.

This confirms the need to reclassify sighting codes, but it is therefore clear that the seven associations should be grouped in the following manner:

- Logs group: object, whale-shark and under a vessel.
- Free school group: free school, marine mammal, sighting code and unknown.

The sighting code and unknown categories are assigned to the free school group pending an assignment rule based on resemblance to the two reference groups.

### **3.4- SUPER-SAMPLING RESULTS**

#### **3.4.1- Variability of sizes and species composition throughout unloading**

This analysis was mainly based on the results of the Atlantic super-sampling.

Changes recorded in the average sizes observed along the unloading of a well are reported on Fig. 5, and the results of the ANOVAs performed in Tab. 2. Average sizes of yellowfin and skipjack change significantly throughout landing, contrarily to those of bigeye. For skipjack, although the differences are significant, the variances, both between and within groups, are small.

Changes recorded in the proportion of different species observed along the unloading of a well are reported on Fig. 6. All the proportions change significantly throughout unloading, which indicates that the arrangement within the well is not homogeneous insofar as species composition is concerned, the greatest differences arising during the first part of operation. Changes in proportion of large yellowfin and bigeye may be due to sorting and subsequent unloading separately without

any mixing with small specimens. As to small specimens, yellowfin is the species which maintains the most stable proportion. The appearance of small tuna in the deepest part of the well may be due to the fact that its smaller size allows it to slip into the gaps left by other species. There is no explanation for the significant changes in proportion of the bigeye and skipjack at the beginning of landing.

### **3.4.2- Sample size**

Sample sizes of yellowfin, skipjack and bigeye corresponding to a 5% and a 10% error level in the average sizes are reported in Tab. 3, and in Tab. 4 for percentages of skipjack and bigeye at an equal error level. It may be observed that species proportions exhibit a greater variability than do sizes, and that to achieve a similar level of accuracy a substantially larger sample size is required: 1,125 fish accepting an error of 5% in the proportion of skipjack, and 281 fish for an error of 10%. If we include the proportion of any other species the sizes triple, at least, for the same error levels.

### **3.5- UNIT OF SAMPLING**

A study done in the Atlantic (which should be easily extended to Indian Ocean) from 95 sampled sets coming from different wells shows that 98% (for species composition) and 92% (size composition) of the between-well, within-set variance is explained by the fishing date, the latter going to 95% if only those samples with over 15 measured specimens are considered.

With respect to the variance existing between samples from the same well (Pallarés and Dewals, 1997), the 2% variance in species between samples of a single set is minimum and allows to consider that a well fairly represents the set insofar as species composition is concerned. Contrarily, the 8% variance in size distribution in the samples of a single set would appear rather high. However, this variance decreases to 5% when the sample size is over 15 fishes. This is again minimum with respect to the variance found inside wells through random sampling.

To conclude, and as far as a minimum sampling level is respected, we can consider that the well is effectively representative of the set, as well for the size distribution than the species composition of the catch.

### **3.6- INFLUENCE OF DIFFERENT FACTORS: STRATIFICATION AND SUBSTITUTION**

Tested factors were area (6 zones, 5° and 1° square), association (log and free schools), fleet (France and Spain), size category of the purse seiners, and time (year, quarter, bi-month and month), using GLM to test the explained variance.

#### **3.6.1- Species composition:**

In the Indian Ocean, analysis put in evidence that the main effects were association (34% of the overall variance), followed by the zone (10%) and the season (8% with quarters); the combination of association-zone-quarter can explain 42% of the variance, with only 36 strata with a minimum of 85 samples (Fig. 7). Using smaller time strata (one month or two month) give a small increase of the explained variance, (43% and 45%) at the cost of an increased sampling (respectively 50 and 90 strata with 61 and 34 samples, Fig. 8). Only using a 5° square strata rise significantly the explained variance (52%), but with a high sampling cost (274 strata with 11 samples each). The zone-quarter interaction is noticeable (Fig. 9).

Using a zone-bi-month-association model with optimal zones (see later), adding the year increase the explained variance with a reasonable number of strata. Finally, it was noticed that the

strong auto-correlation between factors made that grouping the best one leads to a less convenient result. An attempt to define new fishing seasons increased the explained variance, but was not considered as usable because of its heterogeneity (Fig. 10).

Both the fleet and size category of boats factors are negligible; this result is particularly important in term of sampling, as it would allow to process commonly both French and Spanish fleets.

On the contrary, the year factor cannot be put aside. This was not a problem as on one side, year is in practice an imposed factor in the model given that one of the goals of a sampling plan is to monitor trends in catches over the years, and on the other, adding the year to the zone-bi-month-association model increase the explained variance and decrease the number of strata.

Finally, the zone-quarter-association model appears to be satisfactory, and more efficient in the Indian Ocean than in the Atlantic. If necessary, it may be possible to take narrower areas should the case arise, or use bi-month instead of quarters. At all events, samples will be more substantial under the new plan, and the number of strata sampled will also be greater for a cost similar to the present one.

### **3.6.2- Size distribution**

A preliminary study was made using three size categories (<10 kg, 10-30 kg and >30 kg), using the same variables than in the species composition analysis. Again, the three most important factors were, in decreasing order of influence, association, zone and season (Fig. 11), explaining 49% of the variance with only 11 strata; nevertheless, the year factor cannot be excluded - another reason being the strong year-quarter interaction, while other interactions remain weak (Fig. 12) – rising the explained variance to 57%.

If in both cases the models can be refined with a better spatial (5° square) and/or seasonal (1 or 2 months) resolution, this is adversely affected by a large increase of the strata number not enough compensated by a decrease in the minimum number of samples by strata.

Finally, the best model appears to be the zone-association-year one, explaining 57% of the overall variance with 150 strata with a minimum of 20 samples.

### **3.6.3- Substitutions**

The variance explained by the model from which the factor to be tested has been removed was observed - the greater variance giving the less influence of that factor in the model – in order to define the best substitution rules for both species and sizes composition. Results are presented in Tab. 5, and the usable simplified version in Tab. 6.

It has been found, with respect both to, that the year is the factor of least influence, closely followed by the quarter. The area factor (within each association) has much more influence.

### **3.7- DETERMINATION OF NEW AREAS**

In a first step, a 1° square classification by type of association (log and free school) was achieved, considering three classes for both log and free schools, using 8 variables (3 size classes for yellowfin, 2 for skipjack and 3 for bigeye). Classification maps obtained were relatively heterogeneous – and consequently difficult to interpret – but exhibiting some borders effects (islands and coastal areas).

On a general basis, this new stratification enhance the explained variance and decrease the strata number, for both log and free schools and whatever the model used, but it was not possible to define homogeneous zones from these results. On a general basis, the previously used zones – designed from the knowledge of the fishery – appeared to be efficient, even if the new one allowed mainly to decrease the number of strata with 3 zones instead of 5.

The analysis of the seasonality allows to define three typical seasons (yellowfin and skipjack dominating, and an intermediate with equal proportions), defining 7 fishing seasons, but in a less efficient way than a traditional bi-month. So, the previously defined zone-association-bi-month remains valuable, even if the number of strata is quite high. A better sharing of these old zones will probably enhance the situation.

Results were globally very similar for both sizes and species composition.

In a second step, the performance of the new redefined areas was studied in the light of the breakdown and classification reports from the first stage, and compared to the numerous other available data - distribution maps by school type and species composition, sizes distribution according to different time and space resolution - in order to choose the new homogeneous zones to be used.

For the Indian Ocean, original zones were considered as relatively good and in accordance with the different results and observations available. Five spatial strata were retained for both sizes and species composition; as it could be expected, this does not change significantly the overall explained variance, but diminish somewhat the number of strata to sample. Finally, the Zone-Quarter-Year-Association model – which explains 59% of the variance with 132 strata - has been retained.

These zones have been slightly modified recently to take into account the new development of the purse seine fishery in the eastern part of the Indian Ocean. In fact, this does not affect the zones as defined in the traditional fishing area which was under study. Map of these zones in the Indian Ocean are reported on Fig. 13 and in Tab.5.

#### **4- CONCLUSION : A NEW SAMPLING SCHEME FOR THE LANDINGS OF TROPICAL TUNA PURSE SEINE FLEETS**

Results of the numerous analyses performed have permitted a new system for the sampling and processing of landings of Indian Ocean tropical tuna, fished by the purse seine method, to be defined. If the improvement in the accuracy of estimates of the composition of catches by species and by size is not very important, it allows to decrease notably the number of samples to realize.

##### **4.1- DEFINITION**

This is a stratified, multi-species sampling scheme enhancing the basic structure of simultaneous sampling of sizes and species conducted hitherto in the Indian Ocean, but incorporating some changes. The most important innovation lies in the fact that measurements of sizes and counting of species are combined in the selection of sample. This combined method gives rise to a substantial improvement as regards the sample effort, given that:

- a number of fishes of all the species (particularly yellowfin and bigeye) sufficient to obtain the size distribution is ensured;
- the effort of sampling skipjack far above the optimum sample size is saved;

- the sample size obtained is sufficient to yield estimates of species composition of an acceptable degree of accuracy.

The stratification variables to be considered are time, area and type of association

#### 4.2- DEFINED STRATA

**Type of association** : only two strata shall be taken into account, free schools and log schools. The remainder of associated systems shall be assigned to one type or the other according to algorithms defined on the basis of the ACP analyses conducted. Sets without an association code shall likewise be classified according to the algorithms defined in relation to school typology.

**Time strata** : the reference shall be the quarter.

**Area strata** : 5 zones were defined for both free and log schools (Tab. 5 and Fig. 13).

**Weight categories** : in the case of sets with a mixture of species and sizes (sets on logs and on free schools of skipjack) it is recommended that the sampling of large and small specimens be undertaken separately, if there is an estimate of the weight of the large specimens in the well (logbook or well plan). In exceptional cases where there is no estimate of large tuna either in the logbook or in the well plan, but they are found in the sampling process, an attempt shall be made to estimate the total number of large specimens landed. Should that not prove possible, they shall be included in the multi-species sample in accordance with the old system.

**Unit of sampling** : the sampling unit is the set. Samplings shall always be weighted to that unit, irrespective of whether it tallies with the weight of the well or not. If the well contains several sets, the weighting should be done to their total catch.

**Sample size** : this shall depend on the type of association and the species. For free schools of yellowfin it shall be 50 to 100 specimens per sample, as it was similarly the case hitherto. In the case of log schools and free schools of skipjack it shall be 300 to 500 specimens.

This apparently non-random stratification of species within the well, demonstrated through the super-sampling experience, could prove a source of bias in the estimate of the species composition of the well. It is therefore recommended that sampling be conducted in two stages, during the first and the second half of the landing, and that the size of each sample be set at 300 specimens for the first stage and 200 for the second. Thus, should it not be possible to gain access to the well for the second sampling, the previously established sample size may in any event be obtained.

**Sampling coverage** : the number of samples per stratum shall be between 15 and 30, bearing in mind that only one sample is chosen in order to estimate species composition and size distribution. These classifications represent a synthesis of the defined areas for estimates of composition by species and by size. Those areas have been defined in such a way as to ensure coverage of each and every one of the defined area strata for the processing both of sizes and of samples.

**Type of association and weight categories** : the sample selection shall be performed in different manners depending on the type of association and the weight category.

Free schools of yellowfin. Given their single species nature and uniformity of size, these shall be sampled in a random manner, irrespective of the moment in landing.

Free schools of skipjack and log schools. These shall be arranged by weight: small (<10 k) and large (>10 k). The details of captures of large specimens shall be obtained from logbooks and/or

well plans or, in exceptional cases, shall be estimated direct during landing. Large specimens shall be measured according to the procedure for free schools of yellowfin. Small specimens shall be measured in accordance with the new procedure designed.

**Sample selection** : the procedure designed is as follows:

- a traditional multi-species and size sampling process, wherein all the species in the sample are measured, is begun;
- then, when 30 specimens of skipjack (the optimum sample size for the breakdown of that species by sizes) have been measured, the remaining specimens of the species in the sample will not be measured but merely counted. The specimens of yellowfin and bigeye, on the other hand, will be measured throughout the whole of the sampling process. The proportion of skipjack in the sample shall be obtained by multiplying the total number of skipjack in the sample (specimens measured + specimens simply counted) by the average weight of the specimens measured, divided by the total weight of the sample. Should the small tunas (auxis and black skipjack) in the sample be numerous, they shall be processed in the same way as skipjack.

#### **4.3- METHOD FOR THE CORRECTION OF CATCH BY SPECIES IN LOGBOOKS.**

The process shall not be the same for the two types of association. In the case of free schools, correction shall be undertaken by weight category (<10 kg, 10 - 30 kg, >30 kg). Log catches shall, on the other hand, be corrected without stratification by weight, given that reports on catches by weight category in logbooks tend to be unbiased in the case of free schools but do show bias when the sets are on logs. In that type of set, catches of large specimens are rarely reflected in logbooks and to correct by weight category would therefore give rise to an underestimate of such specimens in this form of fishing.

In both cases the correction shall be undertaken on the basis of set weighted samples. In the case of sets taking in a mixture of species, the catch per species shall be estimated on the basis of the composition by species of the sample.

Log schools. For each time-area stratum (zone and quarter) the percentages of each of the species in the sample are calculated. These are applied to the total catch declared in the logbook to obtain a new distribution of the three species in the set.

Free schools. For each time-area stratum (zone and quarter) we shall obtain from logbooks the total catch (yellowfin+skipjack+bigeye) per weight category (>10 kg, 10 - 30 kg, >30 kg). From the weighted samples (same zone and quarter) we obtain the percentages of each three species by the same per weight categories. These are applied to each categories and we thus arrive at a new catch by species within each category.

#### **4.4- SIZE DISTRIBUTION**

Once the species composition appearing in the logbooks has been corrected, the samplings corresponding to each type of association, weighted to the set, shall be grouped according to the defined time-area strata and the catches shall be weighted. For the purpose of breaking down catches by size distribution, monthly catches on 5x5 squares shall be weighted.

#### 4.5- SUBSTITUTION OF SAMPLES

The new stratification, with its substantially lower number of strata, will provide greater sampling coverage and restrict the number of substitutions of unsampled strata. Nonetheless, hierarchical substitution criteria have been established identifying, in accordance with the analyses performed, the factors with least bearing on the variance of size distribution and species composition. **Table 6** illustrates those criteria per type of association and processing.

#### 5.- BIBLIOGRAPHY

- Ariz J., A. Delgado de Molina, A. Fonteneau, F. Gonzales Costas and P. Pallarés, 1993. Logs and tunas in the eastern Tropical Atlantic. A review of present knowledge and uncertainties. Col. Doc. Cient. ICCAT, Vol. XL(2): 421-446 (SCRS/92/039).
- Ariz J., A. Delgado de Molina, J-C. Santana, P. Pallares, R. Delgado de Molina, 1996. Tuna catches, according to the type of association, area and fishing seasons, of the Spanish purse seine fleet in the Atlantic Ocean (1990-93). Col. Doc. Cient. ICCAT, Vol. XLV(3): 130-154 (SCRS/95/058).
- Bard F.X.et P. Vendeville, 1985. Note sur l'échantillonnage plurispécifique des thons tropicaux au port d'Abidjan. Col. Doc. Cient. ICCAT, Vol. XXV: 37-45 (SCRS/85/47).
- Cochran, W.G. 1971. Técnicas de muestreo. Cía. Editorial Continental. Méjico.
- Efron, B. 1982. The Jackknife, the Bootstrap and Other Resampling Plans. Society for Industrial and Applied Mathematics. Philadelphia, Penn.
- Fonteneau, A., 1976. Note sur les problèmes d'identification du bigeye dans les statistiques de pêche. Col. Doc. Cient. ICCAT, Vol. V(1): 168-171.
- Fonteneau, A., 1993. Pêche thonière et objets flottants: situation mondiale et perspectives. Col. Doc. Cient. ICCAT, Vol. XL(2): 459-472.
- Fonteneau A. and P. Pallarés, 1998. Desarrollo de una estructura tipo para bancos libres y bancos asociados a objetos. Col. Doc. Cient. ICCAT, Vol. XLVIII (2): 221- 229 (SCRS/97/27).
- Hallier J-P., 1985. Purse seining on debris-associated schools in the Western Indian Ocean. Expert consultation on the stock assessment of tuna in the Indian Ocean, Colombo, Sri Lanka (28 Nov-2 Dec 1985), IPTP Collective Volume, Vol. 1, TWS/85/30, pp 150-156.
- Hallier J.-P., 1991. Tuna fishing on log associated schools in the Western Indian Ocean: an aggregation behaviour. IPTP Collective Volume, Vol4; TWS/90/66, pp 325-342.
- Hallier J.-P. and J.I. Parajua, 1992. Review of tuna fisheries on floating objects in the Indian Ocean. IATTC, February 11-14, 1992; Special report of IATTC (in press).
- Pallarés, P., A. Delgado de Molina, J.C. Santana, R. Delgado de Molina y J. Ariz, 1995. Evolución de la pesca al objeto y sobre bancos libres de la flota de cerco española en el Océano Atlántico (1990-1993). Col. Doc. Cient. ICCAT, Vol. XLIV (2): 287- 295.
- Pallarés, P. and V. Nordstrom, 1997. Análisis del esquema de muestreo multiespecífico de los túnidos tropicales: presentación del proyecto y primeros resultados. Col. Doc. Cient. ICCAT, Vol. XLVI (4): 168-174.

- Pallarés, P. and J.-P. Hallier, 1997. "Analyse du schéma d'échantillonnage multispécifique des thonidés tropicaux" - Rapport scientifique. IEO/ORSTOM, Programme n° 95/37 réalisé avec le soutien financier de la Commission des Communautés Européennes : 64p.
- Pallarés P. and Ch. Petit, 1998. Tropical tunas : new sampling and data processing strategy for estimating the composition of catches by species and sizes. Col. Doc. Cient. ICCAT, Vol. XLVIII (2): 230- 246 (SCRS/97/28).
- Pianet R. (Ed.), 1995. Statistiques et indices des pêcheries thonières tropicales à la senne. Compte rendu de la réunion du groupe de travail ORSTOM, Paris, 2-5 juillet 1991. ORSTOM Editions, Colloques et Séminaires, 1995 : 61p.
- Snedecor, G.W. 1964. Métodos estadísticos. Cía Editorial Continental, Méjico.
- Sokal, R.R. and F.J. Rohlf. 1969. Biometry . W.H. Freeman and Co. San Francisco.
- Wild, T. 1994. Una evaluación de los procedimientos de muestreo de frecuencia de talla y del análisis subsecuente de los datos de atun aleta amarilla capturado con red de cerco en el Océno Pacífico oriental. Boletín de la Comisión interamericana del atun tropical. Vol. 21, nº 1. La Jolla, Cal.



## 6- TABLES

Year	School Type	c1	c6	c7	c2	n1	n6	n7	n2	p
91	Log	89283	25	11	11	139	53	47	47	18
	Free	75944	74	9	9	101	79	57	57	16
92	Log	103684	19	10	10	136	49	43	43	18
	Free	77461	68	15	15	105	78	61	61	15
93	Log	105705	30	10	10	139	53	49	49	19
	Free	91346	67	17	17	110	82	56	56	14
94	Log	121164	38	11	11	144	52	44	44	17
	Free	107399	72	11	11	106	78	50	50	12
95	Log	161709	19	5	5	163	43	37	37	21
	Free	72218	80	11	11	113	82	59	59	12

n1 : total number of exploited strata

c1 : total catch

n2 : % of non-sampled strata

c2 : % of non sampled catch

n6 : % of strata with less than 1 sampled fish per ton

n7 : % of strata with a sampled weight/total catch less than 1/250

c6 et c7 : % respective corresponding catches

p : % sampled catch.

Table 1. Sampling coverage used to estimate the size and species composition (by year, month and 5° square) of both French and Spanish purse seine fleets in the Indian Ocean, 1991-1996 according to the school association (Log and Free).

ANOVA							
Species	Source of var.	SS	df	MS	F	Prob.	F c. value
SKIPJACK	Between groups	172.47	6	28.75	2.2271	0.0386	2.1081
	Within groups	12236.16	948	12.91			
	Total	12408.63	954				
YELLOWFIN	Between groups	740.79	6	123.46	3.5210	0.0033	2.1888
	Within groups	3576.63	102	35.06			
	Total	4317.41	108				
BIGEYE	Between groups	588.03	6	98.00	1.4504	0.1946	2.1249
	Within groups	23311.95	345	67.57			
	Total	23899.97	351				

Table 2. Analysis of variance of sizes of skipjack and juveniles of yellowfin and bigeye. Groups corresponds to samples taken at different time of loading.

SPECIES ERROR	SAMPLE SIZE	
	5% MEAN	10% MEAN
SKJ	10	
YFT	29	7
BET	47	12

Table 3. Sample size corresponding to errors of 5% and 10% of the mean size value of yellowfin, skipjack and bigeye.

		SKJ LF	YFT LF	YFT LD1	BET LF	BET LD1	FRI
	<b>Well weight</b>	201005.89	22359.46	738.7	100153.51	164	23767.95
	<b>%</b>	57.73	6.42	0.21	28.76	0.05	6.83
<b>Sample size</b>	<b>Error</b>						
	5%	1125.18			3805.58		
	10%	281.30			951.39		
	<b>size 100</b>	57.73 +/- 9.68			28.76 +/- 8.87		
	<b>% Error</b>	16.77%			30.84%		
<b>Confidence</b>	<b>size 150</b>	57.73 +/- 7.91			28.76 +/- 7.24		
<b>interval</b>	<b>% Error</b>	13.69%			25.18%		
	<b>size 200</b>	57.73 +/- 6.85			28.76 +/- 6.27		
	<b>% Error</b>	11.86%			21.81%		

Table 4. Estimated catch by species and species composition in the well: Sample size corresponds to errors of 5% and 10% in the ratios, and Confidence limits considering different sample sizes for skipjack, and small bigeye, which are the main components of samples.

Area 1 ( <b>Somalia</b> )	12°N-50°E to 12°N-80°E 12°N-80°E to 0°N-80°E 0°N-80°E to 0°N-35°E
Area 2 ( <b>North-West Seychelles</b> )	0°S-35°E to 0°S-58°E 0°S-58°E to 7°S-58°E 7°S-58°E to 7°S-49°E 7°S-49°E to 10°S-49°E 10°S-49°E to 10°S-35°E
Area 3 ( <b>East-South Seychelles</b> )	0°S-58°E to 0°S-70°E 12°S-70°E to 12°S-49°E 7°S-49°E to 7°S-58°E
Area 4 ( <b>Mozambique channel</b> )	10°S-35°E to 10°S-49°E 10°S-49°E to 12°S-49°E 12°S-49°E to 25°S-45°E 25°S-45°E to 25°S-32°E
Area 5 ( <b>Chagos</b> )	0°S-70°E to 0°S-80°E 0°S-80°E to 12°S-80°E 12°S-80°E to 12°S-70°E 12°S-70°E to 0°S-70°E
Area 6 ( <b>La Réunion - Mauritius</b> )	South of 12°S and east of 49°E until 80°E
Area 7 ( <b>Arabian Sea</b> , including the Arabian Gulf, Aden Gulf and Red Sea)	North of 12°N

Consequently to the eastern extension of the fishery, and to take it into account, some modification were done later ; however, they do not change significantly the original ones and allows to any Indian Ocean catch to be attributed to an area.

Four areas are slightly modified

- The northern limit of **Chagos** area (now **Maldives-Chagos**) go from 0 to 5°N ;
- **La Réunion-Maurice** area (now **South Indian Ocean**) now include all the southern ocean from Madagascar to Australia ;
- the **Mozambique channel** area, now open on the south qui est ouverte au sud ;
- **Arabian sea** area extends now to 70°E instead of the Indian coast.

Three new areas are created :

- **India-Lakshadweep**, north of 5°N between 70 et 80°E ;
- **Gulf of Bengal**, north of 5°N and east of 80°E ;
- **West Indonesia**, from 12°S to 5°N and east of 80°E.

This modification create in fact to new areas to be effectively sampled (**West Indonesia**, first and fourth quarters), the two others (**India- Lakshadweep** and **Gulf of Bengal**) being there – as the **Arabian Sea** area only to « fill space ».

These areas are reported on Fig. 13.

Table 5. Areas defined to estimate both species composition and size distributions Indian Ocean tropical tuna catches.

**Substitution scheme in the Indian Ocean, Species composition**

1. Samples from the previous quarter, same year, Area and association
2. Samples from the previous year, same quarter, Area and association
3. Samples from two quarters before, same year, Area and association
4. Samples from three quarters before, same year, Area and association
5. Samples from other Areas, same year, quarter and association
6. Adding samples from last year, same Area and association
7. Adding samples from the two previous years, same Area and association
8. Adding samples from the three previous years, same Area and association
9. Adding samples from the four previous years, same Area and association
10. Adding samples from the five previous years, same Area and association
11. Adding samples from the previous year, same association

**Substitution scheme in the Indian Ocean, Size composition**

1. Samples from other areas, same year, quarter, association
2. Samples from the previous quarter, same year, Area and association
3. Samples from the previous year, same quarter, Area and association
4. Samples from the two previous quarters, same year, quarter, Area and association
5. Samples from the three previous quarters, same year, quarter, Area and association
6. Adding samples from the previous year, same Area and association
7. Adding samples from the two previous years, same Area and association
8. Adding samples from the three previous years, same Area and association
9. Adding samples from the four previous years, same Area and association
10. Adding samples from the five previous years, same Area and association
11. Adding samples from the previous year, same association

Table 6. Substitution scheme for strata without samples or with low sampling coverage in the Indian Ocean. Hierarchy is established as a function of the importance of the different factors in the variance of sizes and species composition.

**Substitution scheme in the Indian Ocean, Species composition**

1. Samples from the previous quarter, same year, Area and association
2. Samples from the following quarter, same year, Area and association
3. Samples from the previous year, same quarter, Area and association
4. Samples from the same year, Area and association
5. Samples from other Areas, same year, quarter and association
6. Adding samples from last year, same Area and association
7. Adding samples from the previous year, same association

**Substitution scheme in the Indian Ocean, Size composition**

1. Samples from other areas, same year, quarter, association
2. Samples from the previous quarter, same year, Area and association
3. Samples from the next quarter, same year, Area and association
4. Samples from the previous year, same quarter, Area and association
5. Samples from the same year, quarter and association
6. Adding samples from the previous year, same Area and association
7. Adding samples from the previous year, same association

Table 7. Substitution scheme for strata without samples or with low sampling coverage to be effectively used in the Indian Ocean.

## 7- FIGURES

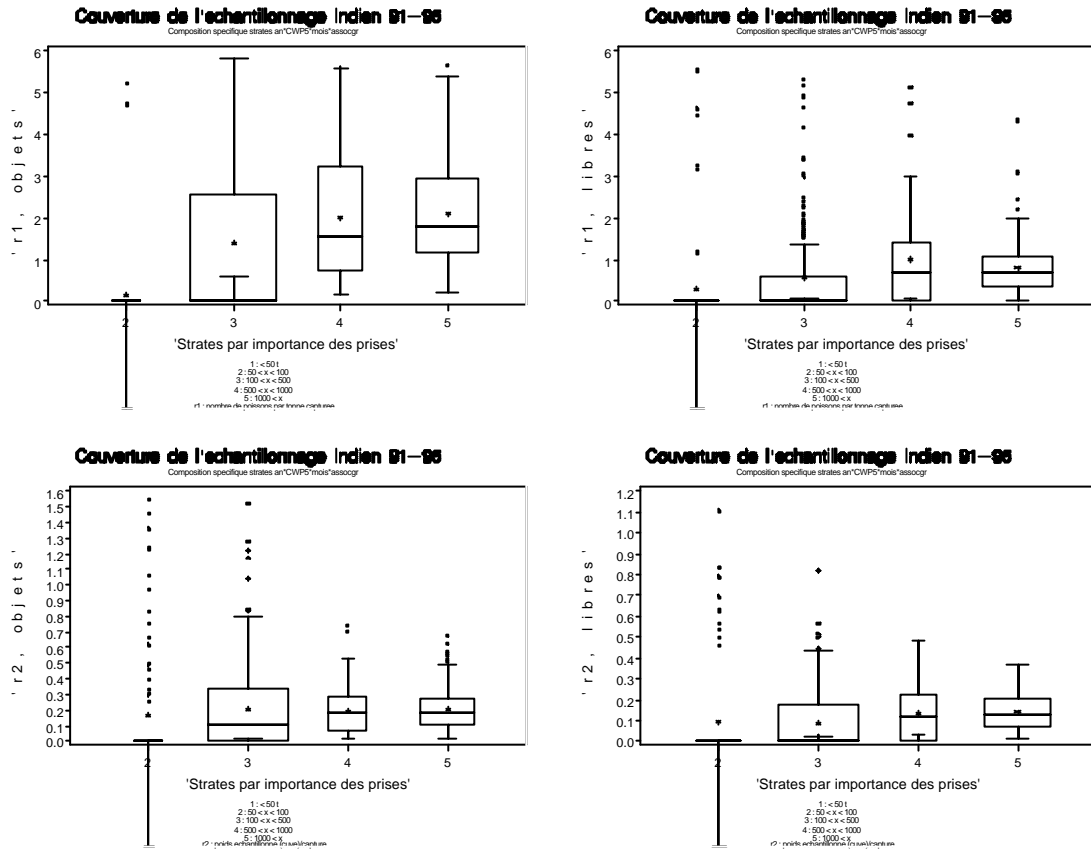


Figure 1. Sampling coverage rate, as the number of measured fishes (r1) and as sampled weight (r2) per ton caught by the French and Spanish fleets in the Indian Ocean, 1991-1995, according to the type of association (free or log schools).

## Couverture de l'échantillonnage Indien

Poids échantillon (cuve) en fonction de la capture 91-95

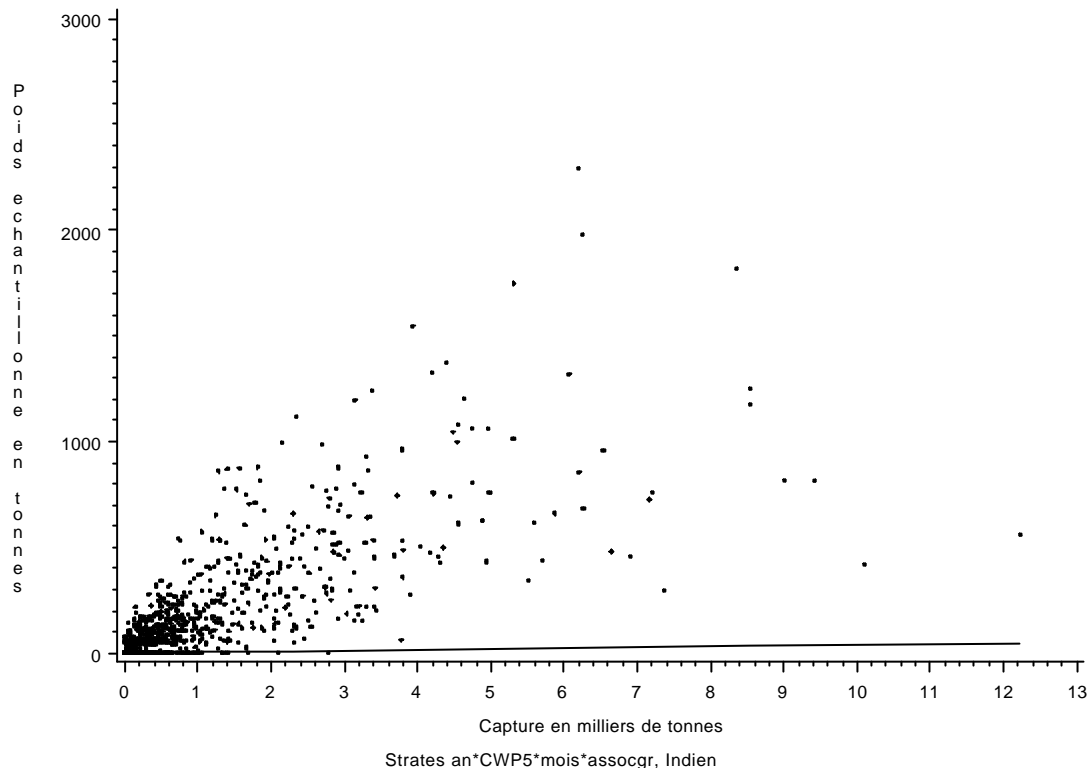
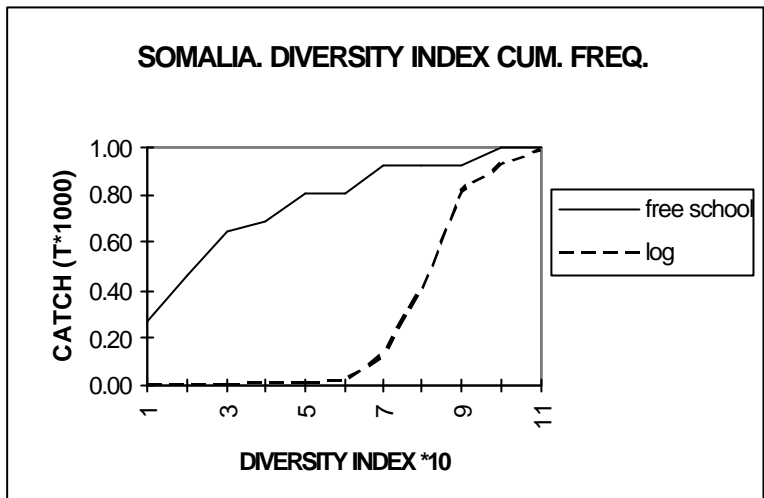
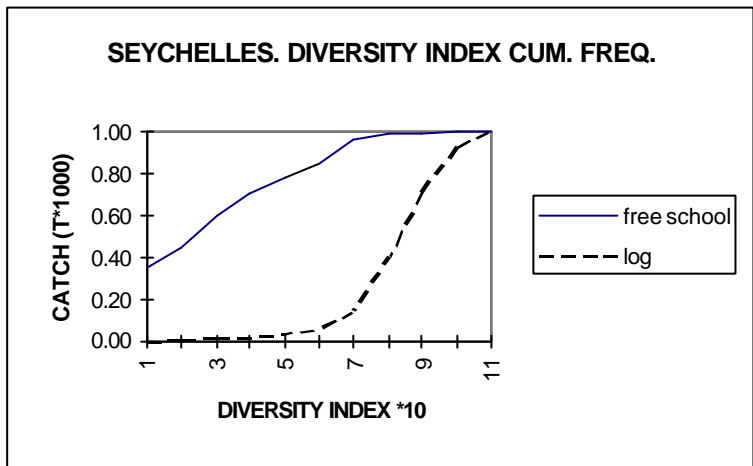


Figure 2. Sampled versus total catch weight relationship from the French and Spanish purse seiners fleets in the Indian Ocean ; strata are 5° square by month and association.

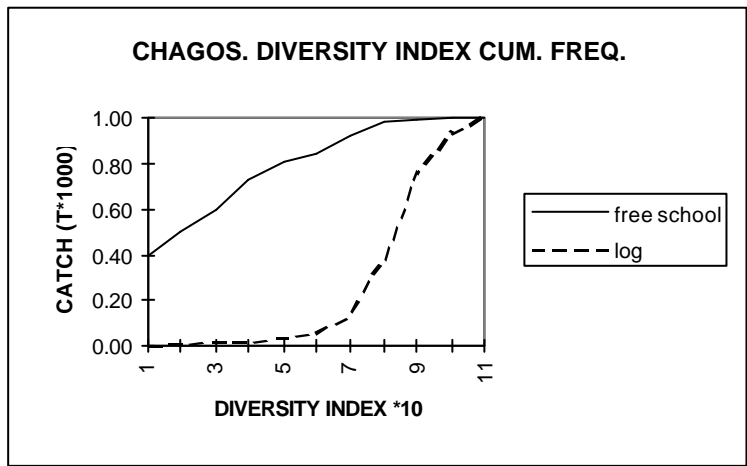
a



b

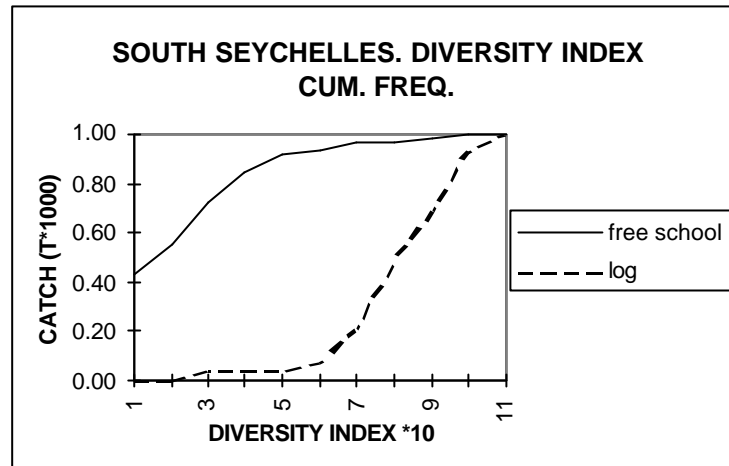


c





d



e

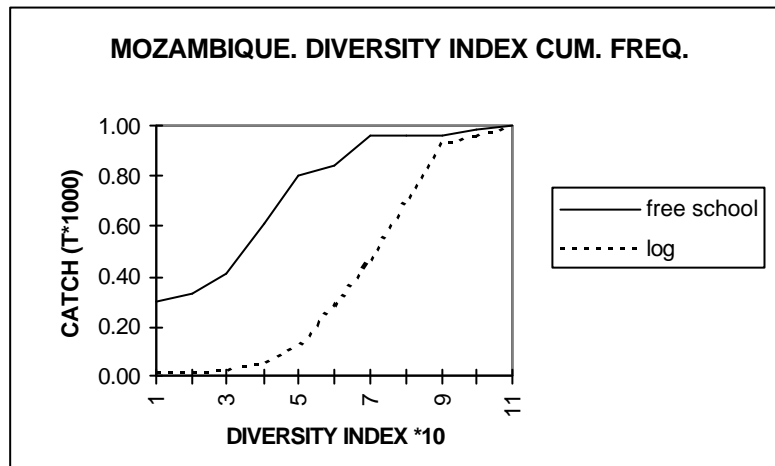
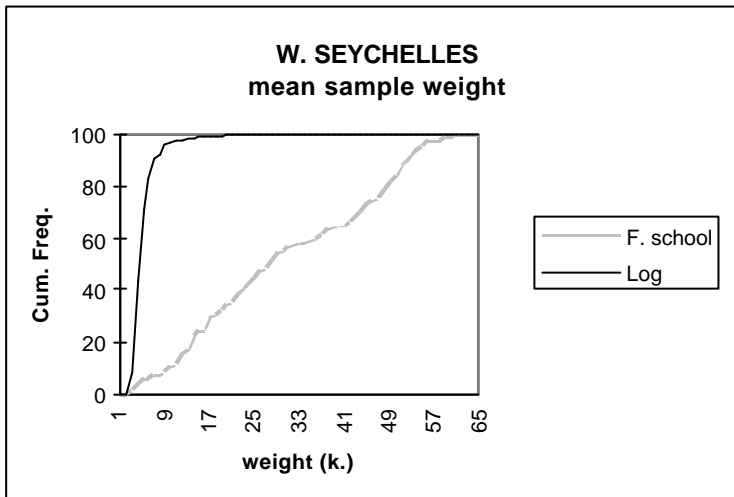
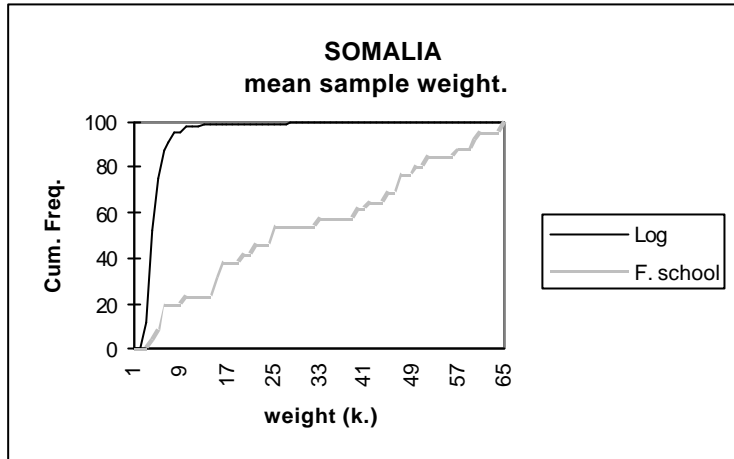
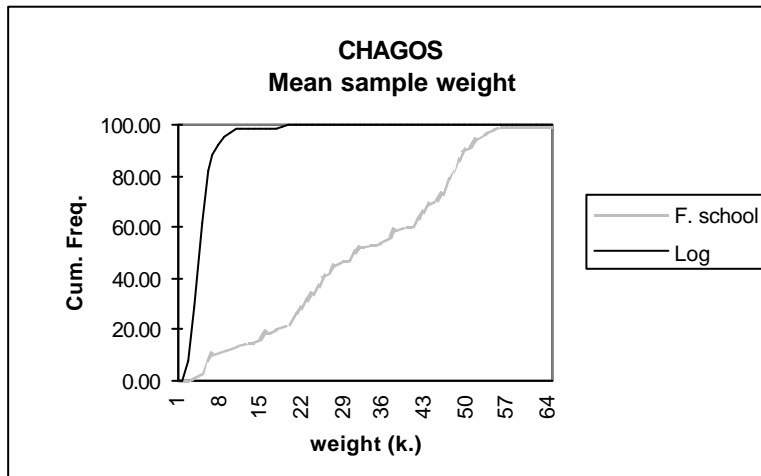


Figure 3 a-e - Cumulative frequencies of absolute value of Shannon-Weaver diversity index (\*10) computed from samples by school's type (free school and log) in five Indian Ocean areas.

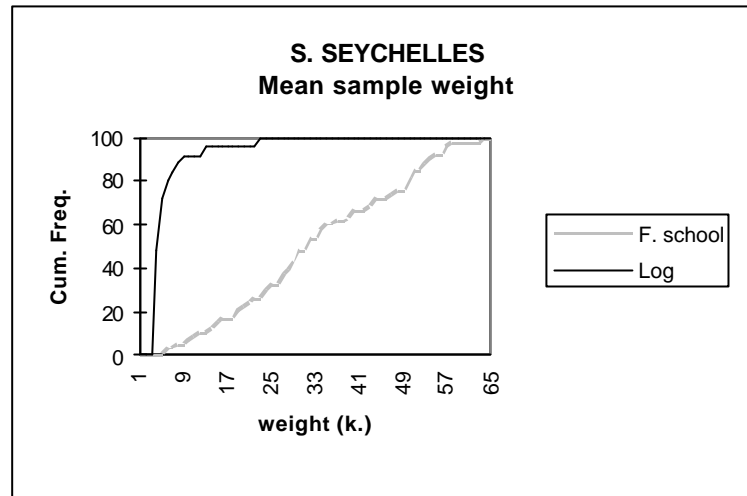
b



c



d



e

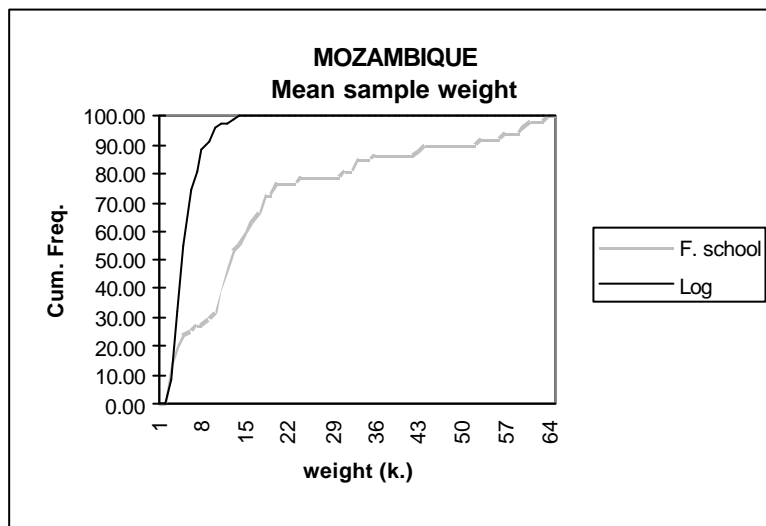


Figure 4 a-e - Cumulative frequencies of sample mean weight, all species mixed, by school's type (free school and log) in five Indian Ocean areas.

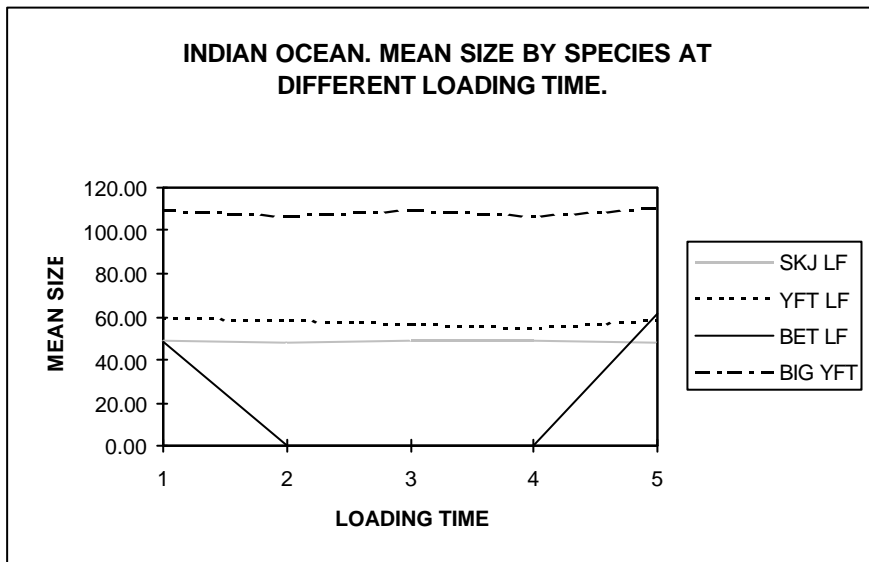
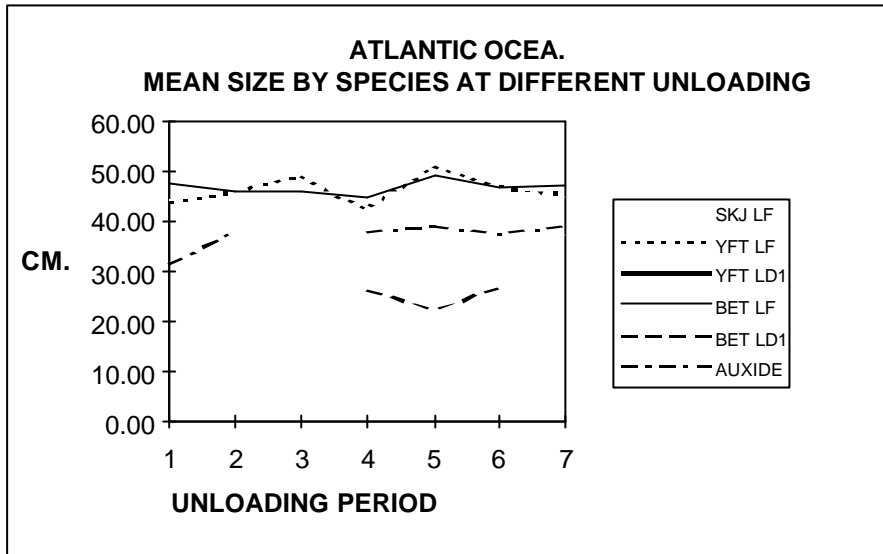


Figure 5 - Mean size trend by species along the unloading of a well, from two super-sampling experiments in both Atlantic (above) and Indian (below) Oceans.

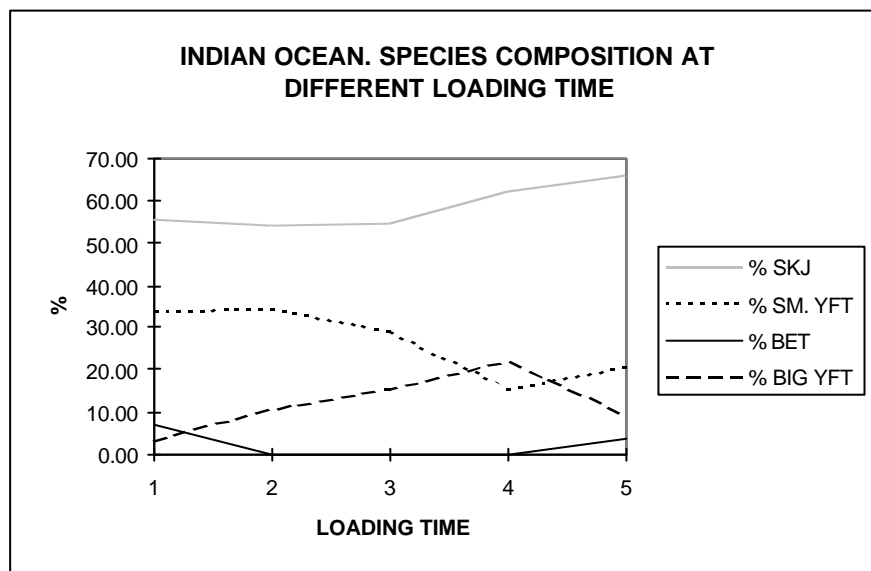
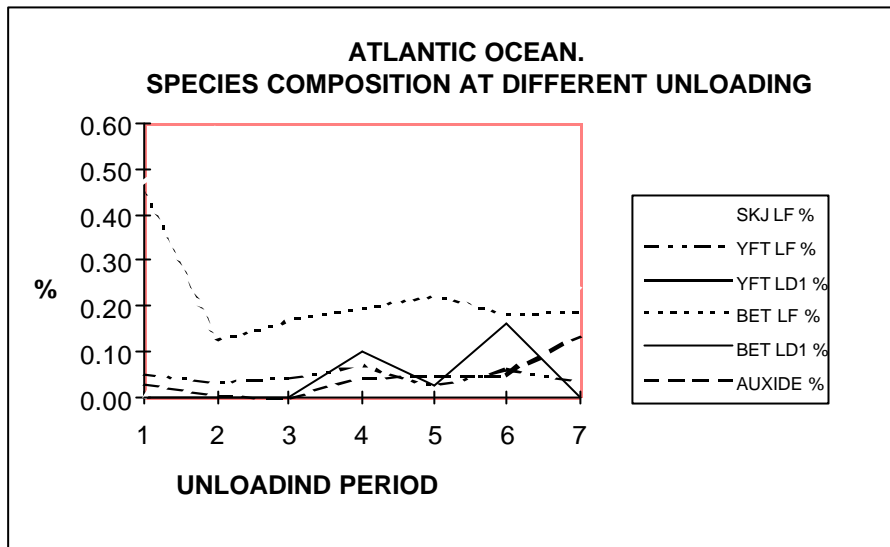


Figure 6 - Species percentage trend along the unloading of a well, from two super-sampling experiments in both Atlantic (above) and Indian (below) Oceans.

## Composition spécifique Indien Influence des différents facteurs

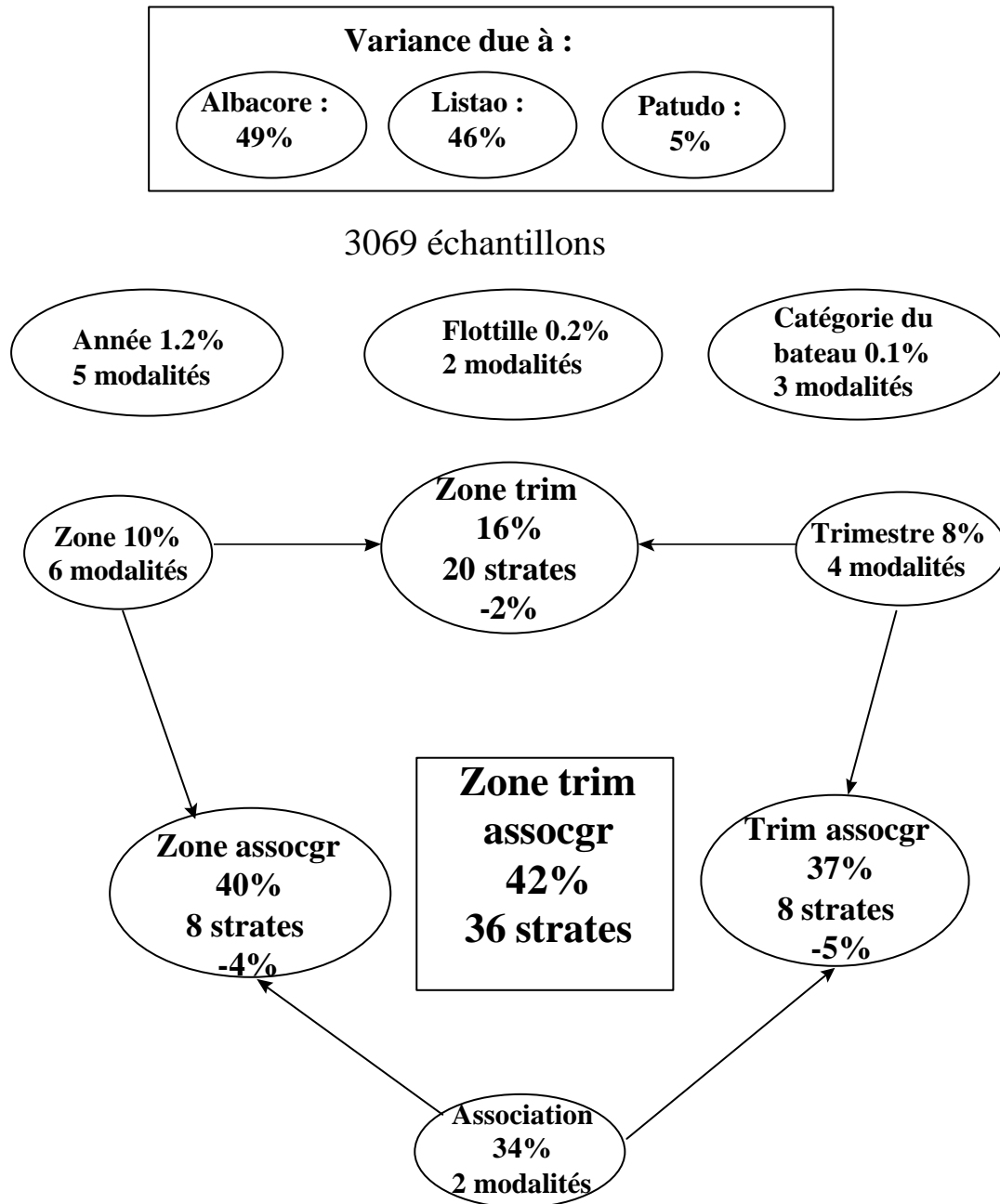


Figure 7 - Main factors affecting the species composition of samples in the Indian ocean.

Composition spécifique Indienne  
Modèles a trois facteurs

3069 échantillons

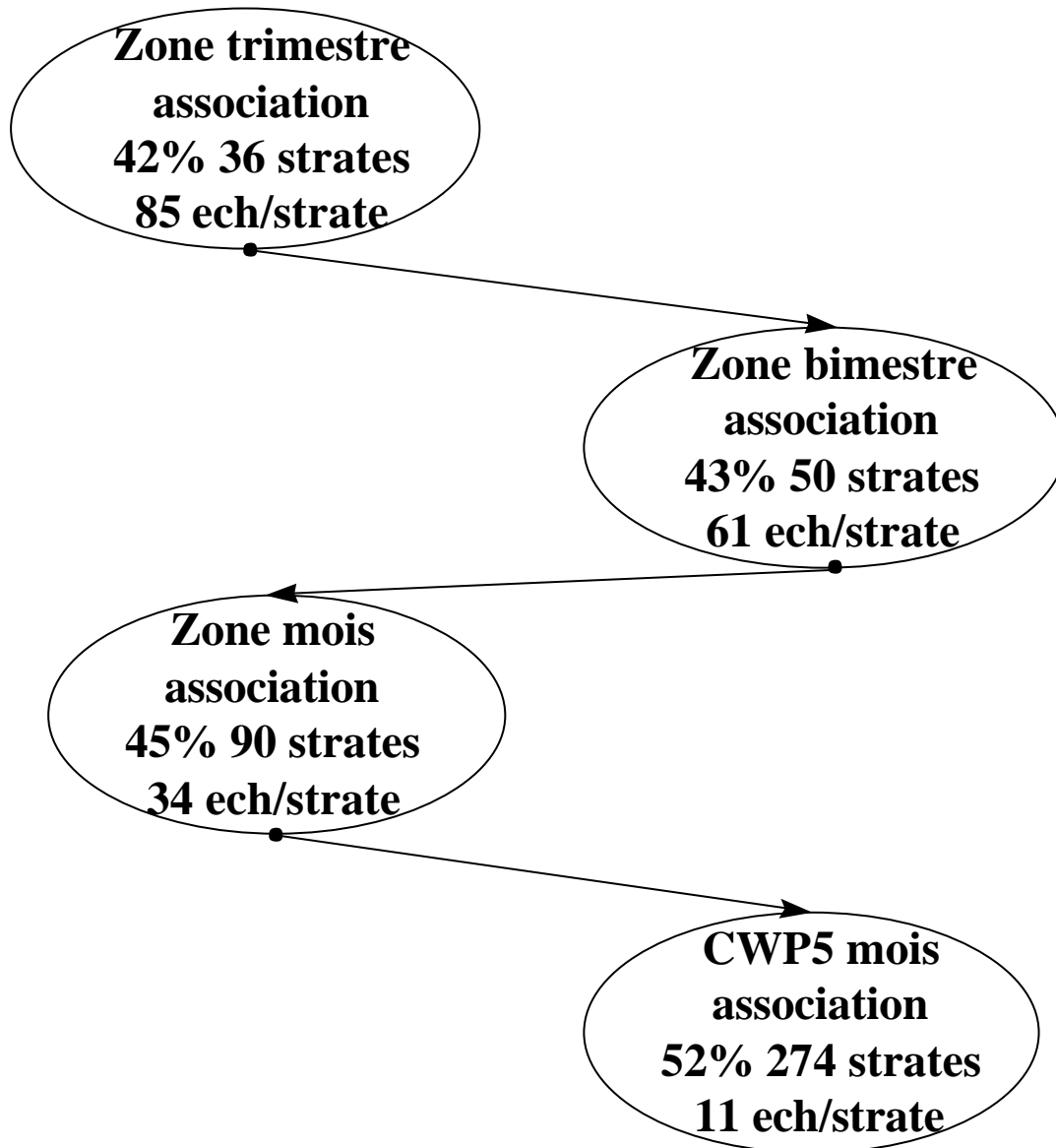


Figure 8 - Explained species composition variance from a 3 factors model.

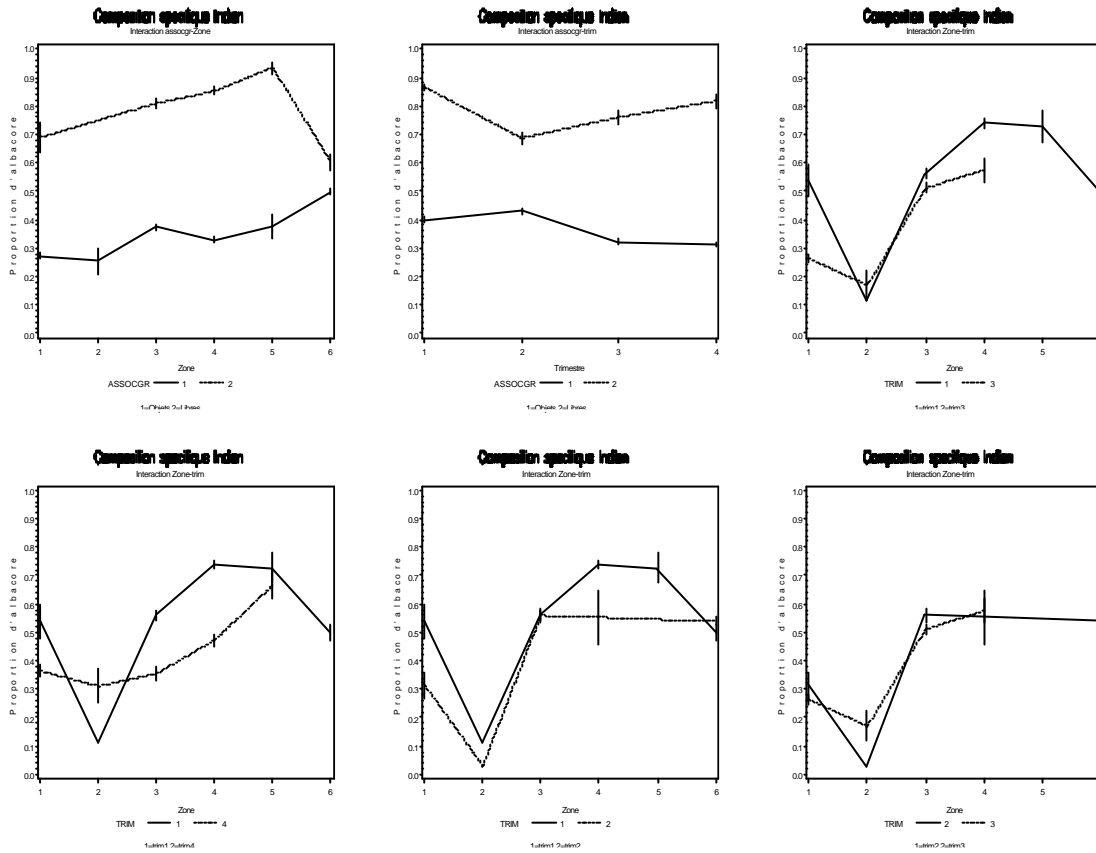


Figure 9 - Interaction between the different factors affecting the variability of the species composition in the Indian Ocean : area-association (above, left), association-quarter (above, center), area-quarter (above, right), association-fleet (below, left), area-fleet (below, center) and quarter-fleet (below, right); the variable was the yellowfin proportion in the sets.



## Nouvelles saisons Indien

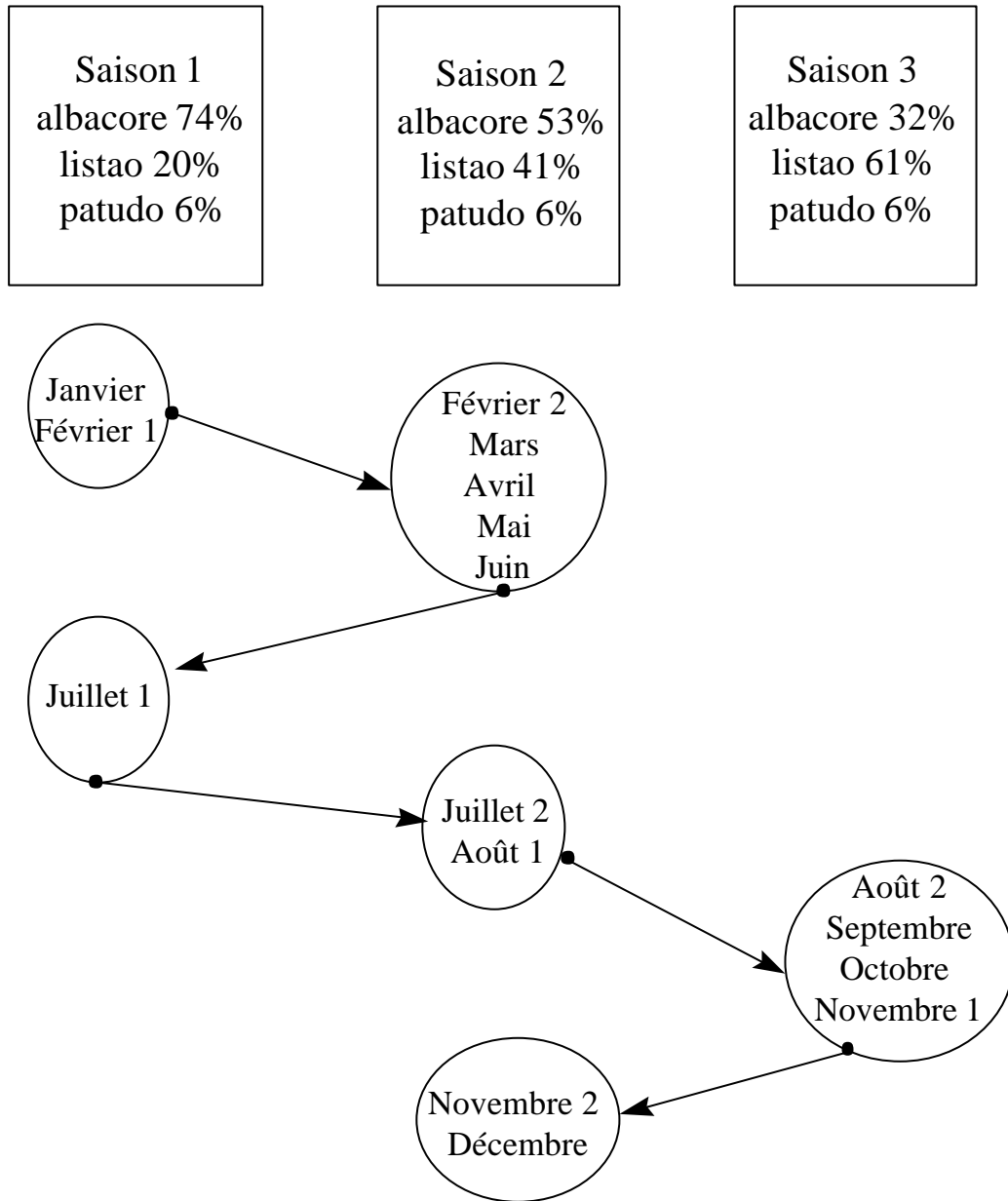


Figure 10 – Temporal strata proposed for the specific composition of the catch on log and free schools in the Indian Ocean. As for areas, those temporal strata explains a larger part of the total variance, but are heterogeneous.

## Composition en tailles Indienne Influence des différents facteurs

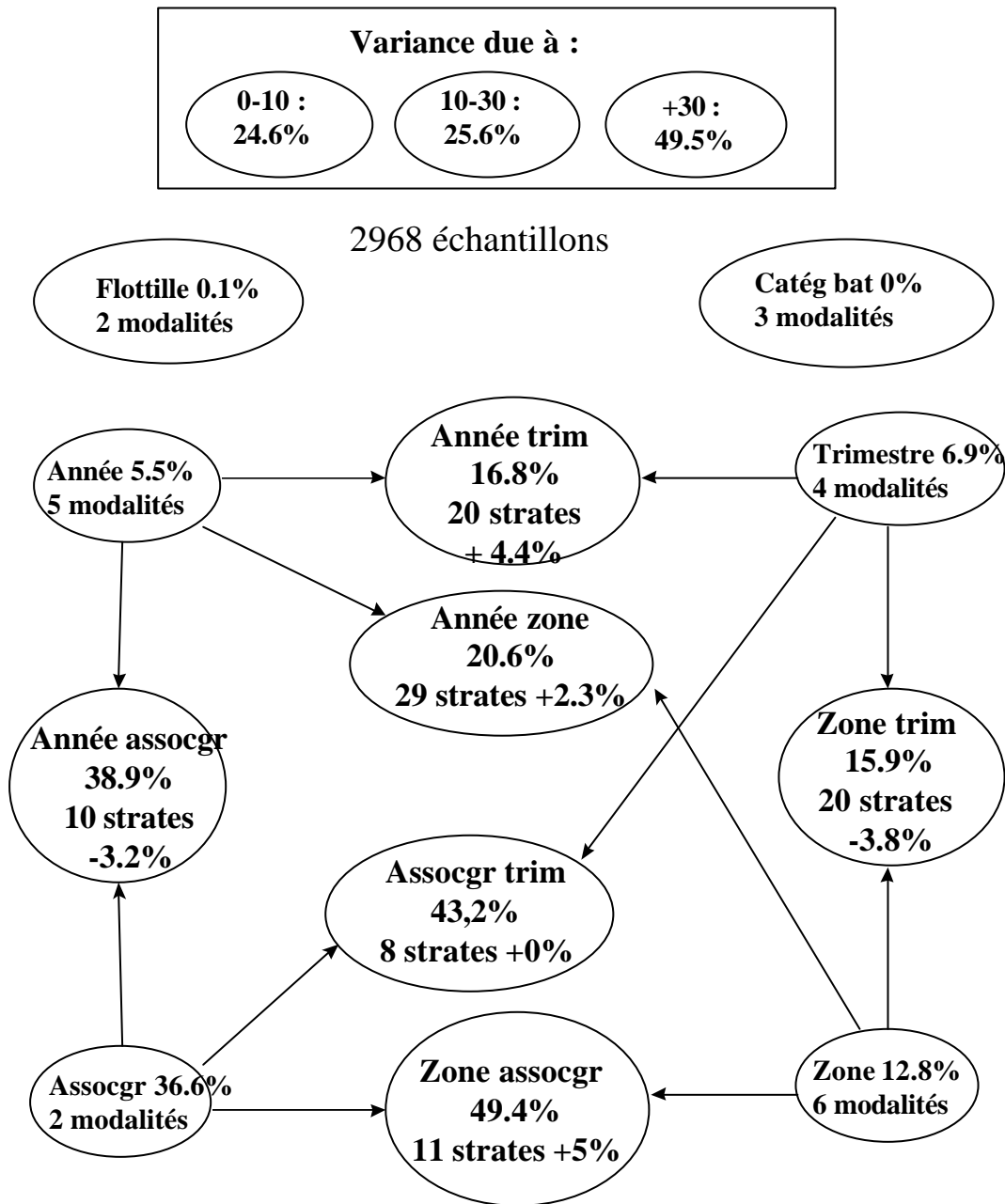


Figure 11 - Variance in the species size distribution explained by different factors and their interactions.

## Composition en tailles Indienne Albacore, Modèles à 2,3,4 facteurs

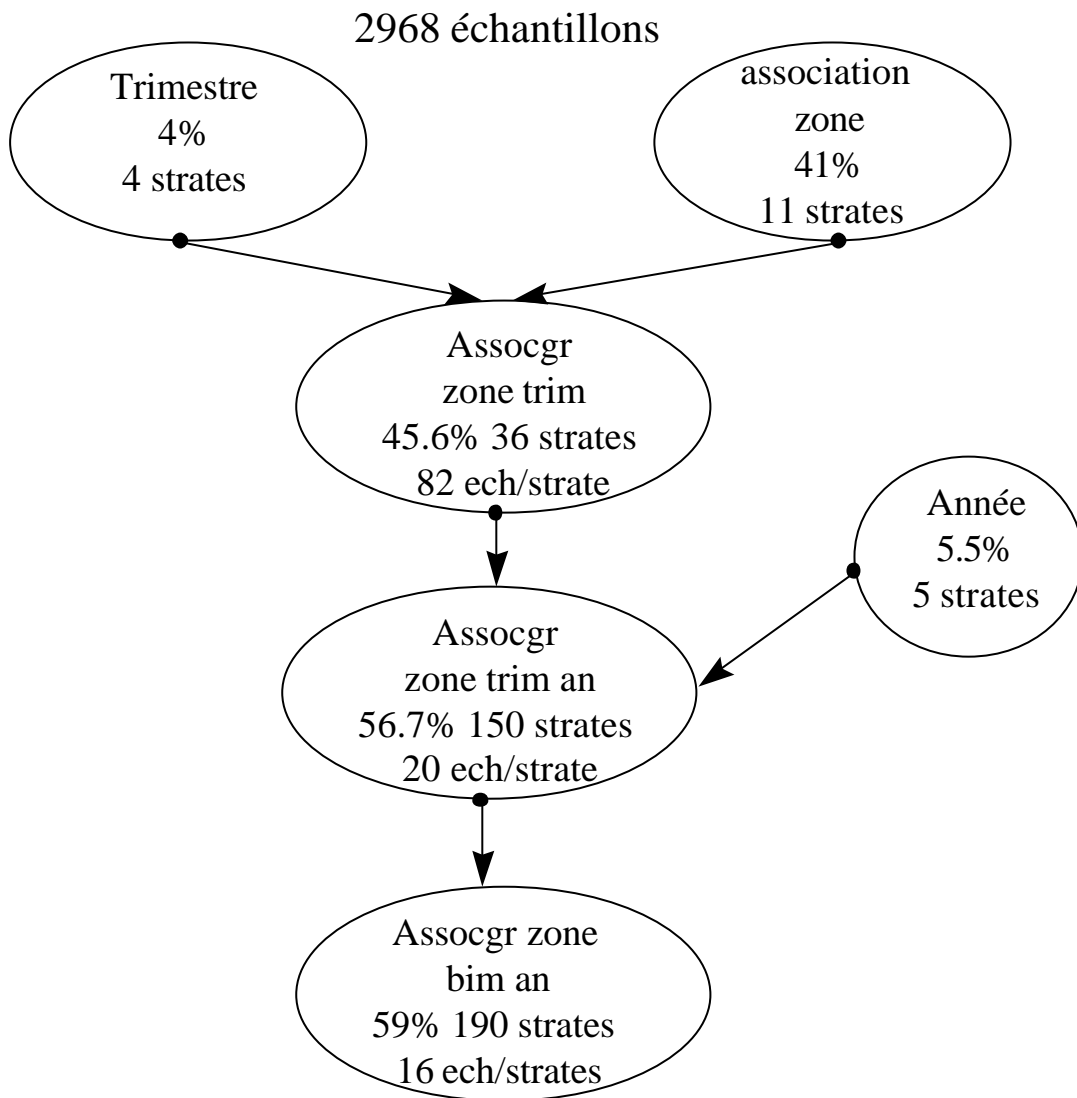


Figure 12 - Main factors affecting the species size distribution from 2, 3 and 4 factors models in the Indian ocean.

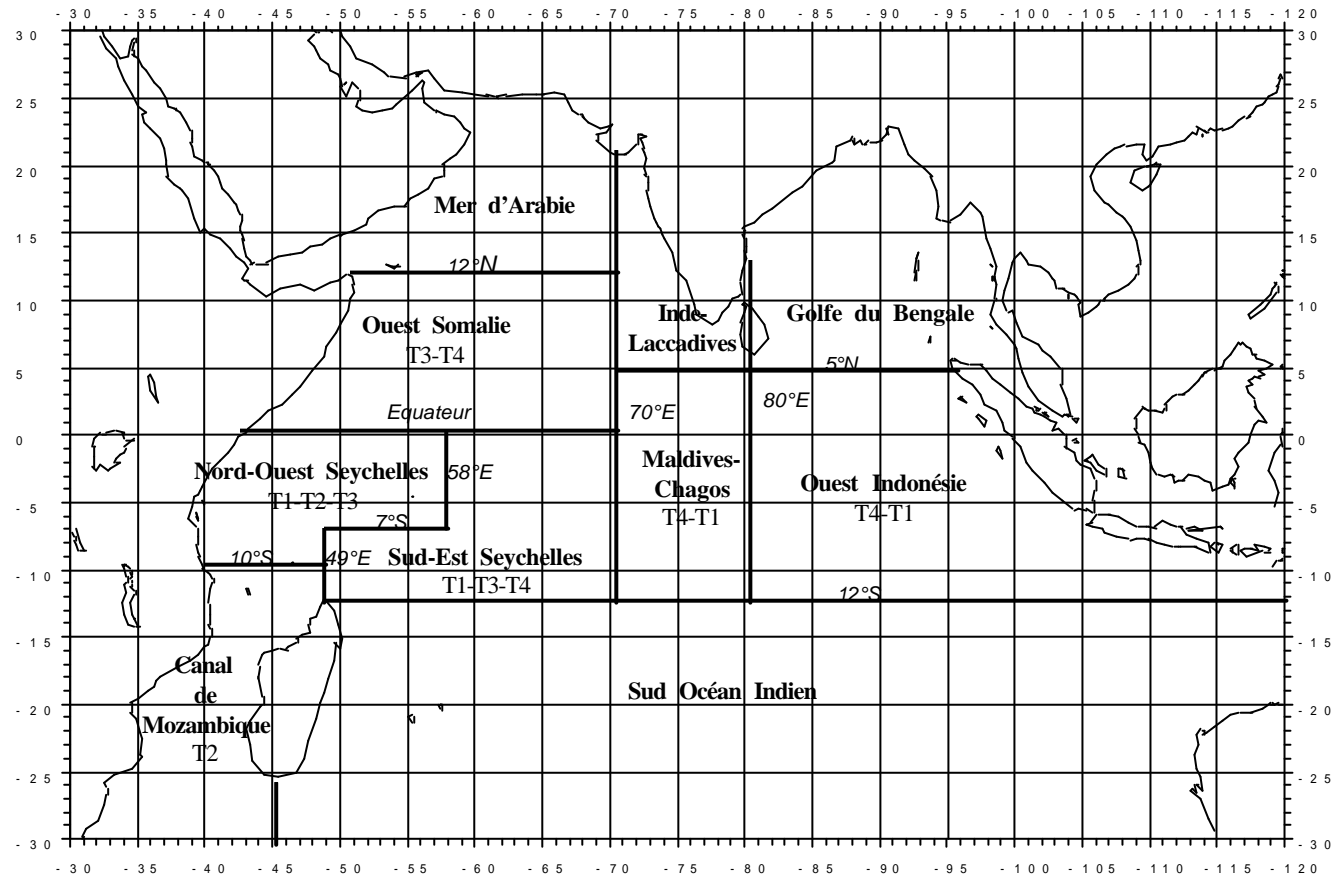


Figure 13 - Spatial strata defined in the Indian ocean to correct the species composition and to calculate the size distribution of both log and free school catches of tropical tuna.