Report on the validation of Seychelles Industrial longline data.

J. Dorizo[1], A. Fonteneau[2] ,V. Lucas[3]

Abstract

This report presents the various validation methods applied to the logbook data for all vessels licensed to fish in the Seychelles EEZ, all Seychelles registered vessels and the statistics for the Seychelles registered vessels for the years 2000 to 2005. All the historical data was transferred into FINSS and data entry in FINSS started in June 2005. The data was verified for missing positions, positions on land, missing hooks and wrong species. The logbook positions were crosschecked with VMS data revealing 88% of correct data. Weight of fish and catches in multiple of 10 was verified and corrected or flagged. Analysis of the corrected data shows an improvement in logbook coverage and an increasing trend in catch and effort of the Seychelles flagged vessels. Bigeye tuna remain the most dominant species of the total catch from 2000 to 2004 and was replaced by Yellowfin tuna in 2005. Spatial distribution map of catch and effort is similar to that of the Taiwanese with a significant yellowfin catches in the Arabian Sea in 2005. Further analysis will be conducted to investigate the species composition and yearly economic yield of vessels.

Report on the validation of Seychelles Industrial longline data.

J. Dorizo[1] , A. Fonteneau[2] ,V. Lucas[3]

1.0 INTRODUCTION

Industrial fishing activities began in the Seychelles waters in the early 1950's with the Distant Water Fishing Nations (DWFN) longlining for tuna in the Western Indian Ocean (WIO). This was initiated by the Japanese and soon followed by the Taiwanese (1954) and the Koreans (1960). Longliners from European Union countries (Britain, France and Spain) applied for licenses to fish in the Seychelles Exclusive Economic Zone in 1993. Seychelles registered vessels (Taiwanese origin) started operating in 1999.

All longliners licensed to fish inside of the Seychelles Exclusive Economic zone is under obligation to submit a logbook to the Seychelles Fishing Authority detailing their daily catch and effort while inside of the Seychelles EEZ. Seychelles flag vessels are oblige to submit a logbook of all their daily activities even while operating outside the Seychelles EEZ.

Given that those distant water longliners vessels do not use port Victoria for transshipment as they usually transship on the high seas, landing data are often missing and logsheets are the only source of fisheries data. These data needs to be verified so as to establish whether they are credible or has been falsely reported. This problem concern logbooks from all international fleet licensed to fish inside the Seychelles EEZ but most importantly Seychelles flag vessels of Taiwanese origin, as it is the responsibility of Seychelles to ensure that the statistics that are submitted annually to IOTC are credible.

With landing data missing, alternative methods needs to be use to validate the submitted logbook data. A series of validation procedures were identified and were applied. The aim of these validations is to detect probable errors so as to correct the data wherever possible or to flag the erroneous data and eliminate them from further analysis.

This report presents the various validation method applied to the logbook data for all vessels licensed to fish in the Seychelles EEZ including all Seychelles registered vessels for the years 2000 to 2005 and the findings of these procedures. In addition the statistics for the Seychelles registered vessels for the years 2000 to 2005 are presented.

2.0 DATA ENTRY

In June 2005, SFA started to use the FINSS software developed by IOTC for longline fishery data entry. All the historical data was transferred into FINSS. It must be noted that 38 out of 1952 (2%) trips transferred were eliminated during transfer, as they were trips that overlap with other trips that was entered using the previous software NEWTUNA. Hence 1581 duplicate sets out of 75656 (2%) sets were also eliminated.

Since the previous software used had few error checking procedures, we started to verify all the data now in FINNS working from 2005 backward. To date we have verified and almost completed correction of all data entered from the year 2000 to 2005.

3.0 VERIFICATION PROCEDURES

3.1 Geographical positions

The data was verified for missing positions that may have been omitted by mistake. The missing positions were corrected whenever they were available on log sheet and the mistake was due to data entry error. In cases where the positions were not reported on the logbook, the record was flagged.

In addition, it was noticed that a lot of records had positions as 0 degrees and 0 minutes latitude. A visual examination of such data reveals that the software previously used did not allow the minutes to be recorded, which is important to distinguish the quadrant where such position is. For these records, the minutes available on logbook were updated.

Furthermore, verification of the historical data that were transferred revealed that some positions were on land. These data were corrected or flagged.

Table 1 below summarises the number of records (sets) where the positions were identified as possible errors and the number of records that were flagged.

Year	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
Number of Incorrect positions	85 (0.4%)	280 (1.0%)	574 (2.0%)	238 (0.8%)	39 (0.1%)	12 (0.1%)
Flagged records	0	0	17	0	2	1

Table 1. Number of records where positions were identified as possible errors.

3.2 Missing Hooks

Next the data was crosschecked for records where hooks were missing. The missing hooks were entered where they were available on logbook. In cases where the hooks were not reported on logbooks they will be later estimated during data analysis.

Table 2 below summarise the number of records (sets) with missing hooks and the number of such records that were updated.

Year	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
Records with missing hooks	58 (0.3%)	72 (0.3%)	185 (0.7%)	517 (1.6%)	399 (1.4%)	458 (5.0%)
Updated Records	0	0	149	178	21	0

Table 2. Number of records with missing hooks

3.3 Species

Longline fishery only targets certain species of fish and only certain species are recorded as by-catch on log sheets. Keypunchers can sometimes make mistakes in the species selection during data entry. Hence the data was verified for species that were wrongly recorded and is summarised in table 3 below. The following species were entered although they were not recorded on the respective logbook. The species label as "Yellowfin and Bigeye tuna" was most of the time selected instead of Yellowfin tuna.

Species	Frequency (Number of records)
Black Escolar	2
Black Shark (Silky shark)	12
Indo-Pacific sailfish and Shortbill spearfish	1
Killer whale	1
Sawara (Barracuda & Wahoo)	1
Shortbill spearfish	9
Tunas nei	3
Wahoo	3
Yellowfin tuna and Bigeye tuna	75

Table 3. Frequency of wrongly selected species.

4.0 VALIDATION PROCEDURES

The validations procedures are based on various parameters:

4.1 Geographical Positions

Keypunchers can sometimes enter the position reported on logbooks wrongly or the positions declared on logbook can be wrong. However when VMS data are available, the positions declared on logbook can be crosschecked with the VMS data. It was observed that for 2003 SFA received 62% of VMS positions and 70% for the year 2004. For 2005 only 46% of the VMS position was available when the verification was done. Hence statistics for 2005 is still preliminary.

A FORTRAN program, DISTVMSALL.FOR, was used to validate the logbook positions using VMS data. The program calculates the difference in daily distances between the logbook positions and the VMS positions and any distances between the two positions greater than 60miles (1º) was identified as an error.

It must be noted that since the VMS data are most of the time only available for when the vessel is inside the Seychelles EEZ, positions outside the EEZ could not always be validated. Furthermore validation of positions was done for the year 2003 to 2005 only, since the SFA VMS centre became operational in 2003.

The test reveals that 88% of the daily logbook positions for the Seychelles flag vessels were consistent with VMS data with 66% of the distances less than 10miles. It was noted,

however that a significant number of positions had distances greater than 120miles (2º) or even greater than 500miles. (Figure1). It was noticed that 30% of the incorrect positions had distances between 61 to 120 miles and 24% had distances over 960miles. (Figure2). The later case was mainly due to north, south, east and west being wrongly entered.

These positions need to be corrected wherever possible and if not the records need to be flagged. The positions for all Seychelles flagged vessels have been corrected or flagged for the years 2000 to 2004 and partly for 2005. This process is still ongoing for vessels of other nationalities.

Table 4 below shows the percentages of records with wrong positions by year for the Seychelles Flag vessels.

Year	Total Records	N0. Positions with distance >60	% Positions with distance >60	% records that was flagged
2003	3590	551	15%	14%
2004	3958	423	11%	9%
2005	2801	276	10%	5%
Total	**10349**	**1250**	**12%**	10%

Table 4. Records with wrong positions

4.2 Weight of fish by Species

The longline logbook catch are most of the time reported as number and weight of fish by species. This allows the mean weight (kg) of the fish by species to be calculated and compared. The following rules were applied:

- The weight of the species should not exceed the maximum for the Indian Ocean
- The weights that are below the minimum weight are suspected as unlikely but are not impossible.

The program LECPDNB.FOR developed by Mr. A. Fonteneau was used to list out all records where the sizes of the main commercial species were either below the minimum or exceeded the maximum weight given in table 5 below.

Species	Minimum Weight	Maximum Weight
YFT	5	100
BET	5	140
ALB	3	50
SWO	5	300

Table 5. Minimum and Maximum size of main commercial species.

The test shows 0.2% of records with anomalies in fish weight and table 6 below shows the number of records by species that was flagged after corrections and the ranges of sizes reported that exceeded the maximum weight.

Species	Albacore	Bigeye Tuna	Yellowfin Tuna	Swordfish
Records with weight of fish < Minimum	0	12	13	6
Records with sizes of fish > Maximum	19	33	39	2
Range of sizes (kg) > Maximum	51-88	142-554	101-431	312

Table 6.Records with erroneous fish weight.

4.3 Catches in multiple of 10

A visual examination of certain logbooks revealed that in some logbook most number of fish were recorded as an exact multiple of 10. The frequency of a vessel daily catch being a multiple of 10 exceeding 30% of the total catch for that particular month gives reason to suspect that the data may have been falsely reported.

The program LECPNDB.FOR lists out all records by vessels by trip where over 30% of the records for a particular month are in multiple of 10. This test was done by vessel by month as a vessel may make false report of its catches only for a certain part of the trip.

In consequence 702 fishing days (23months) representing 0.5% of the total records (fishing days) for the years 2000 to 2005 were flagged for this reason.

5.0 STATISTICS

The corrected data was analysed and compared to previously published data. It must be noted that after the fisheries data for the years 2000 to 2003 was published, more logbooks were received and entered. The catch data that was previously published for the years 2000 to 2002 were processed weight while the catch data that has been revised for the whole period under study has been converted to round weight using the IOTC conversion factors. Therefore an increase in catches for the years 2000 to 2002 is expected. Readers should also note that the statistics presented here has not been extrapolated to take into account missing logbooks.

5.1 LOGBOOK RECORDS

Table 7 shows the percentage of logbook records received by SFA for the years 2000 to 2005. From the year 2003 there was a significant improvement in logbook returns at SFA.

Year	Logbook returns (%)
2000	39
2001	57
2002	42
2003	94
2004	82
2005	88

Table 7. Logbook Coverage

5.2 EFFORTS

Fishing effort

Fishing effort (Vessel active, fishing days and number of hooks) is presented in table 8 and figure 3. On average 20 vessels were active per year.

The revised effort shows an increasing trend in fishing effort reaching a peak of 18 million hooks in 2005. A sharp increase can also be observed between 2003 and 2004.

Likewise the number of fishing days increased from 1486 days in 2000 to a peak of 5770 fishing days in 2005 (Figure 4).

The average number of hooks per set remains more or less stable over the years.

Year	Vessel active	Effort (1000 Hooks)	Fishing Days	Average Hooks per Set
2000	11	4,325	1486	2910
2001	14	5,566	1746	3188
2002	18	8,161	2668	3059
2003	25	9,300	2850	3263
2004	28	17,612	5664	3109
2005	26	18,224	5770	3158
Average	20	10,531	3,364	3115

Table 8. Fishing effort for the years 2000 to 2005

5.3 CATCHES AND CATCHES PER UNIT EFFORT

Total Catches

The evolution of the catches is presented in table 9 and figure 3. The total revised catches have increased steadily over the past six years. In 2005 a record catch of 12,562 Mt was reported. The mean catch per set ranges between 1.03 Mt to 2.180 Mt.

Year	Catch (Mt)		CPUE (Mt/1000 hooks)		Mean Catch Per set
	Revised Catch	Previous Catch	Revised CPUE	Previous CPUE	Revised Catch
2000	1,530	459	0.354	0.234	1.03
2001	2,610	1,185	0.469	0.513	1.50
2002	3,843	913	0.471	0.547	1.44
2003	4,793	4,638	0.515	0.513	1.70
2004	9,891	9,989	0.562	0.549	1.75
2005	12,562	12,778	0.689	0.689	2.18

Table 9. Evolution of catches

Catches by Species

From the year 2000 to 2004 bigeye tuna remains the most dominant species caught by the Seychelles registered longliners. However in 2005 Yellowfin tuna (51%) replaced bigeye (39%) as the most caught species. It must be noted that in 2001, other species accounted for 46% (44% Albacore) of the total catch, followed by bigeye tuna 29% and yellowfin tuna 14%. Table 10 and figure 5a and 5b show the catch by species for both the revised and the previously published data.

Year	Yellowfin Tuna (Mt)		Bigeye Tuna (Mt)		Swordfish (Mt)		Albacore (Mt)		Others (Mt)	
	Revised	Previous	Revised	Previous	Revised	Previous	Revised	Previous	Revised	Previous
2000	421	153	808	190	111	41	0	0	191	75
2001	342	236	715	360	358	189	666	154	529	246
2002	499	122	1,594	400	693	191	581	49	476	151
2003	871	1,025	2,300	2,502	725	527	563	356	334	228
2004	3,032	3,108	5,214	5,350	1115	982	53	52	477	497
2005	6,387	6,517	4,883	4,999	827	770	117	117	349	375

Table 10. Catch by species of revised and previously published data.

Catch per Unit Effort.

The Mean Catch per unit of effort has increase steadily over the years as illustrated in figure 7. The CPUE increase from 0.35Mt/1000 hooks in the year 2000 to 0.46Mt/1000 hooks in the year 2001. It then remained more or less stable between the years 2001 and 2004. It reaches a peak of 0.69Mt/1000 hooks in 2005.

5.4 SPATIAL DISTRIBUTION OF CATCHES AND EFFORT.

The spatial distribution of catches and effort are illustrated in figure 8 and 9 respectively. The maps show the mean catch and effort per year for the period under study. It can be observed that bigeye tuna is the most dominant species. A significant yellowfin catches can be observed in the Arabian Sea which was reported during the period of February to May 2005. It must be noted that the fishing zone is similar to the Taiwanese.

Table 11 below shows the total yellowfin catch and effort for the Arabian Sea by month for the year 2005. It can be observed that in April the average yellowfin catch per vessel per day and the Yellowfin CPUE was the highest.

Month	Yellowfin Catches (Mt)	Effort (1000 Hooks)	YFT CPUE (Mt/1000 hooks)	Average YFT Catch per vessel per day
Feb	275	163	1.692	6.26
Mar	887	595	1.492	5.41
Apr	1,100	503	2.188	7.43
May	1,127	700	1.610	5.50
Jun	84	117	0.716	2.53
Grand Total	**3,473**	163	**1.672**	**5.85**

6.0 Further Analysis

Once all data verification, validation and correction have been completed two further analysis can be carried out to further investigate the species composition of the logbooks and the possible economic yield of a vessel.

6.1 Species composition

The species composition declared on logbook can be uncertain. The species composition can be compared to the trend of the Japanese fleet or the Taiwanese fleet according to the ecosystem of the fishing zone (Longhurst 1998). To this effect the program LLCHECKS.FOR compare for each vessel, the species composition on a quarterly basis by the longhurst zones to the Japanese and Taiwanese fleet species composition in that zone. If the catch reveals a species composition never before observed in either fleet, the data need to be investigated for possible false declaration of species composition.

6.2 Yearly Economic yield

The absolute economic yield (mt/fishing day or in dollars/1000 hooks) of a vessel is also a good index to identify suspicious logbook. Certain vessel can under declare their catches for tax purposes or for other reasons. Other vessel can overestimate their catches (e.g. Bigeye), so as to commercially launder their catches that exceed their quota or their illegal catches caught from other oceans.

The program RENTABLL.FOR estimates the economic yield of each vessel per year in according to the annual mean prices by species on the Japanese market. This can be compared to the annual estimated profit break-even point of a vessel to see if the vessel keeps running at a loss but keep on fishing.

7.0 Conclusion

In conclusion, the Seychelles Fishing Authority has taken some steps in order to publish and disseminate reliable statistics. Various validation processes has been identified and tested to improve the quality of the data for the longline fishery. It must be noted that these validation and correction is time consuming. However now that SFA is using FINSS for data entry, some error checking are available in FINNS such as checking positions that are on land. This will help reduce the amount of corrections that need to be made during verification.

Statistics of the corrected data reveals that both catch and effort of the Seychelles registered vessels have increased steadily over the years under study, with bigeye tuna being the most dominant species of the catches. The year 2005 was an exceptional year for the Seychelles longline fleet with a record total catch, CPUE and yellowfin catches.

8.0 List of Figures.

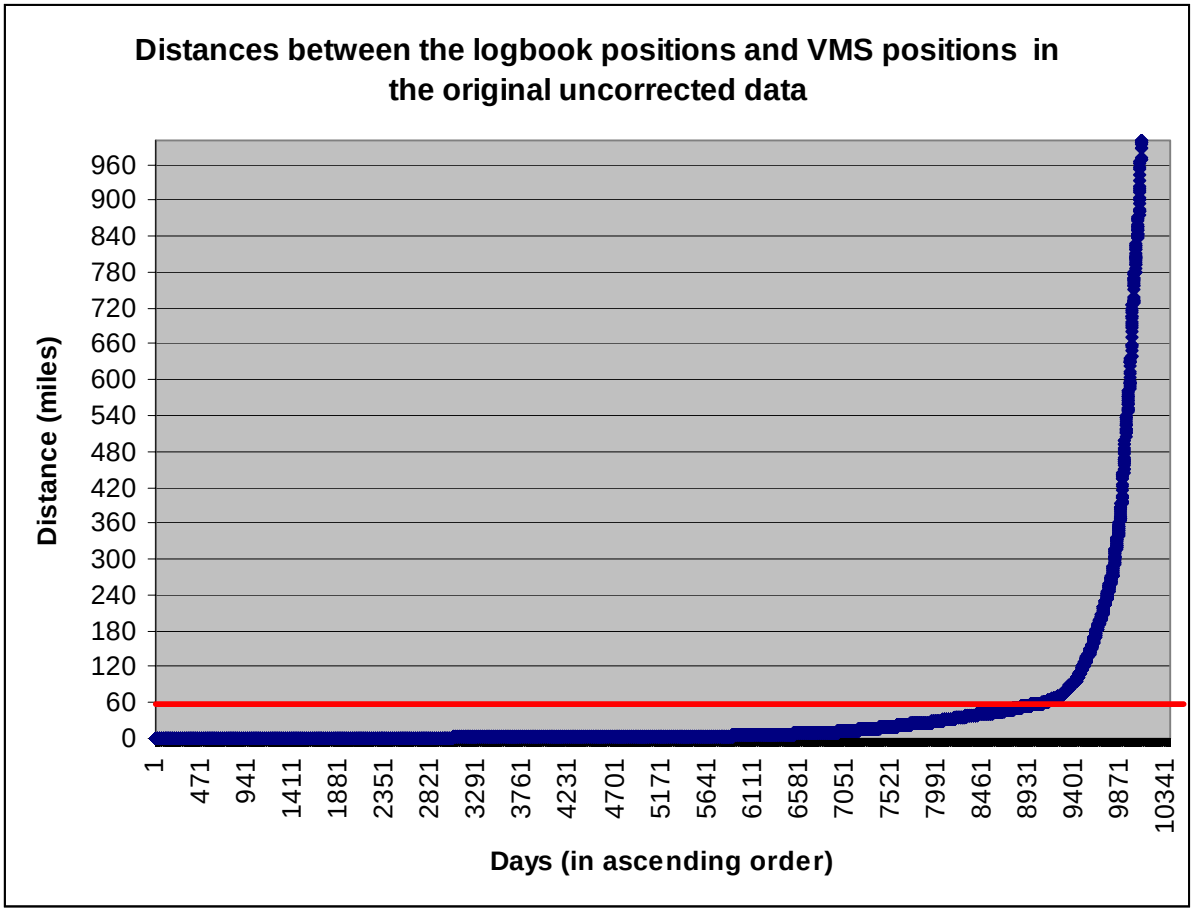

Figure 1. Distances between Logbook positions and VMS positions

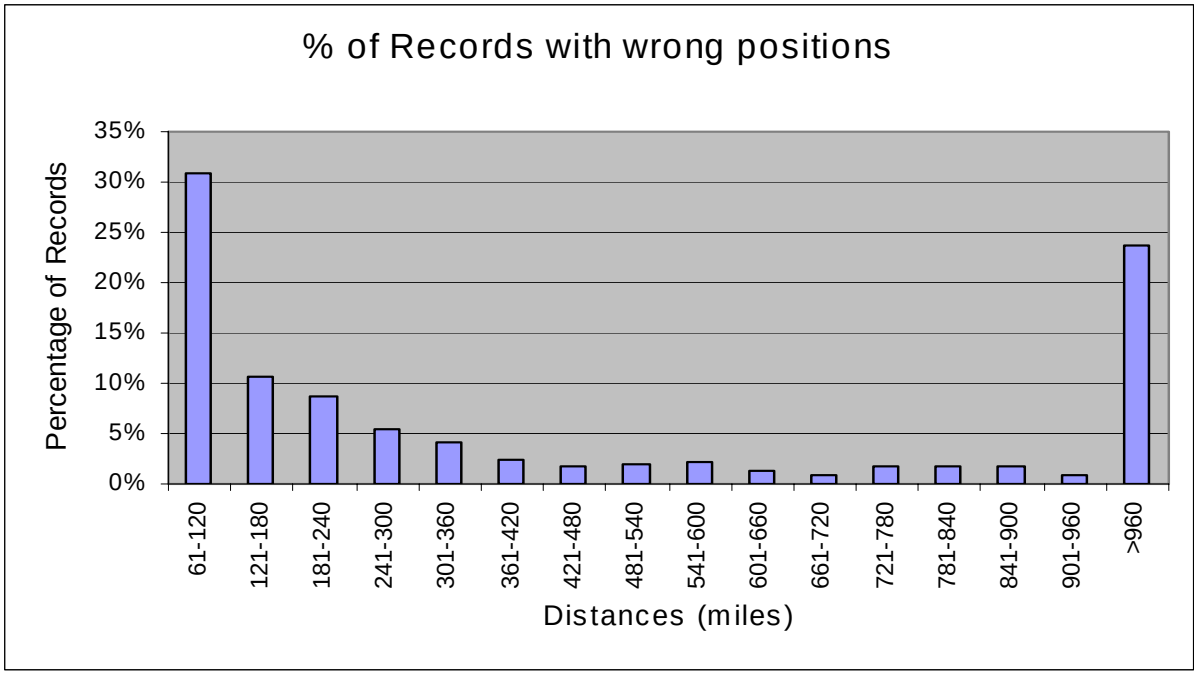

Figure 2. Percentage of records with erroneous positions

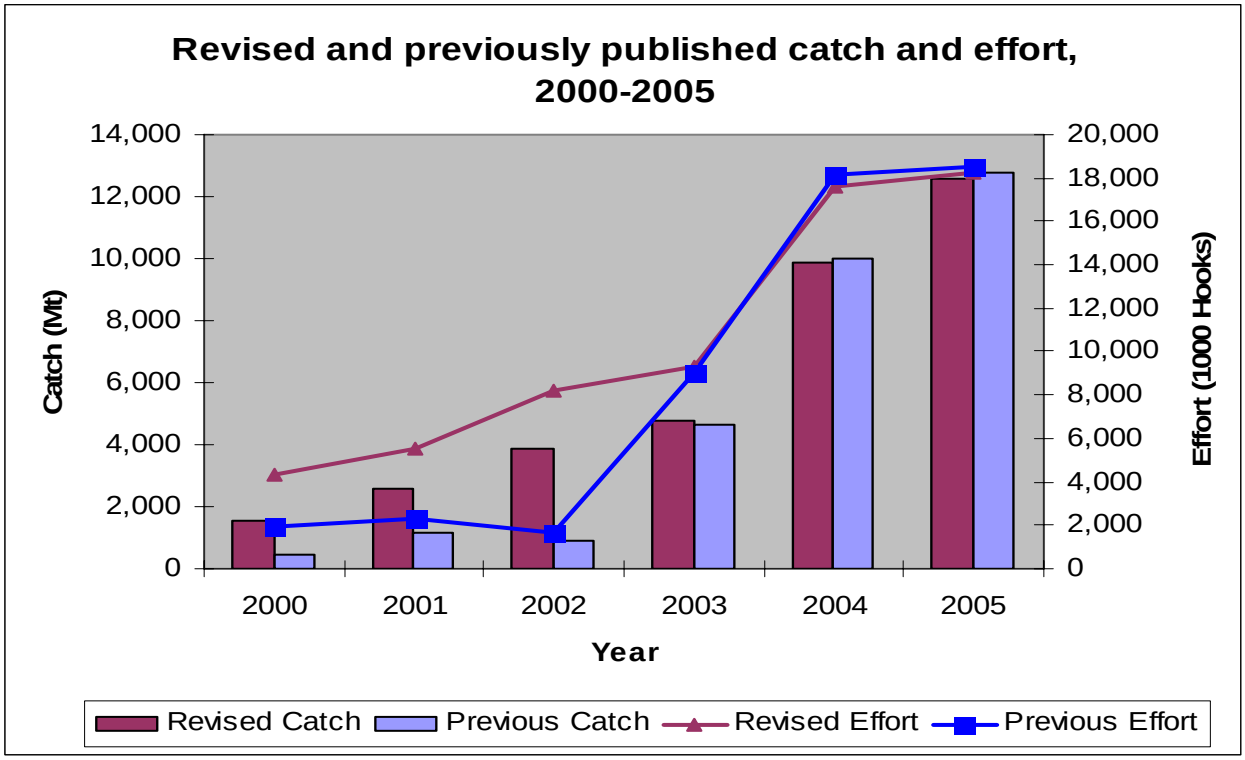

Figure 3 Catch and effort of Revised and previously published data

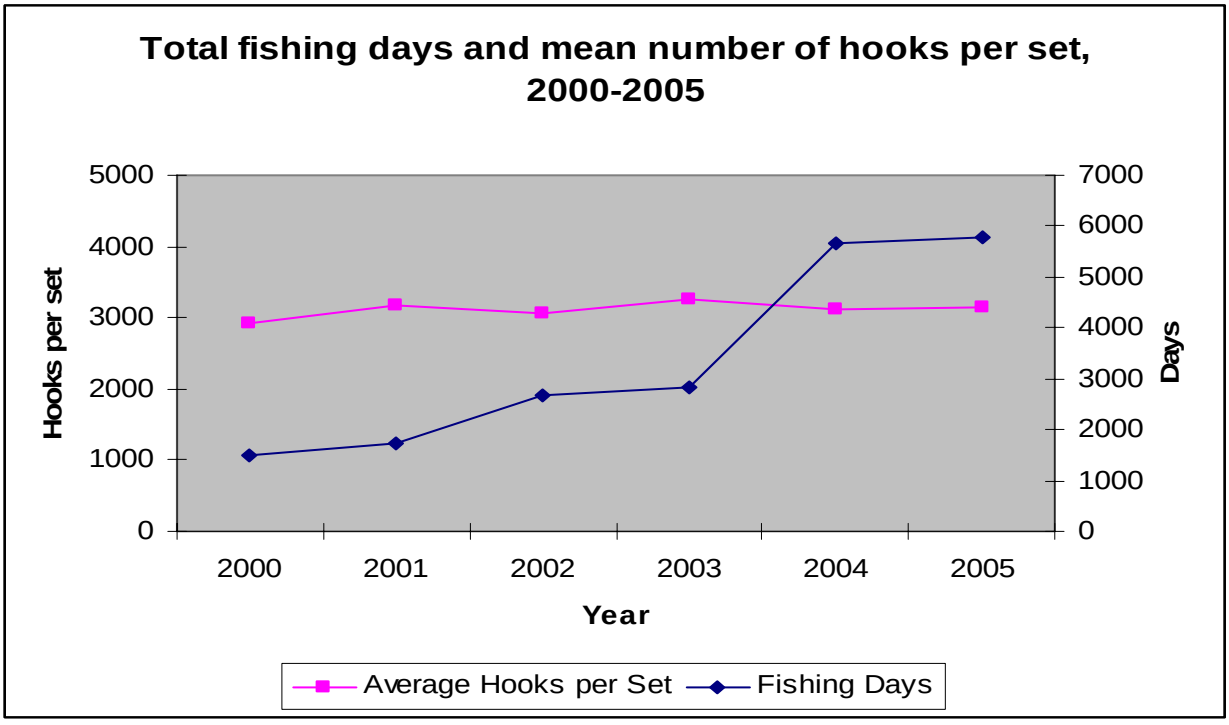

Figure 4. Total Fishing days and Mean hooks per set.

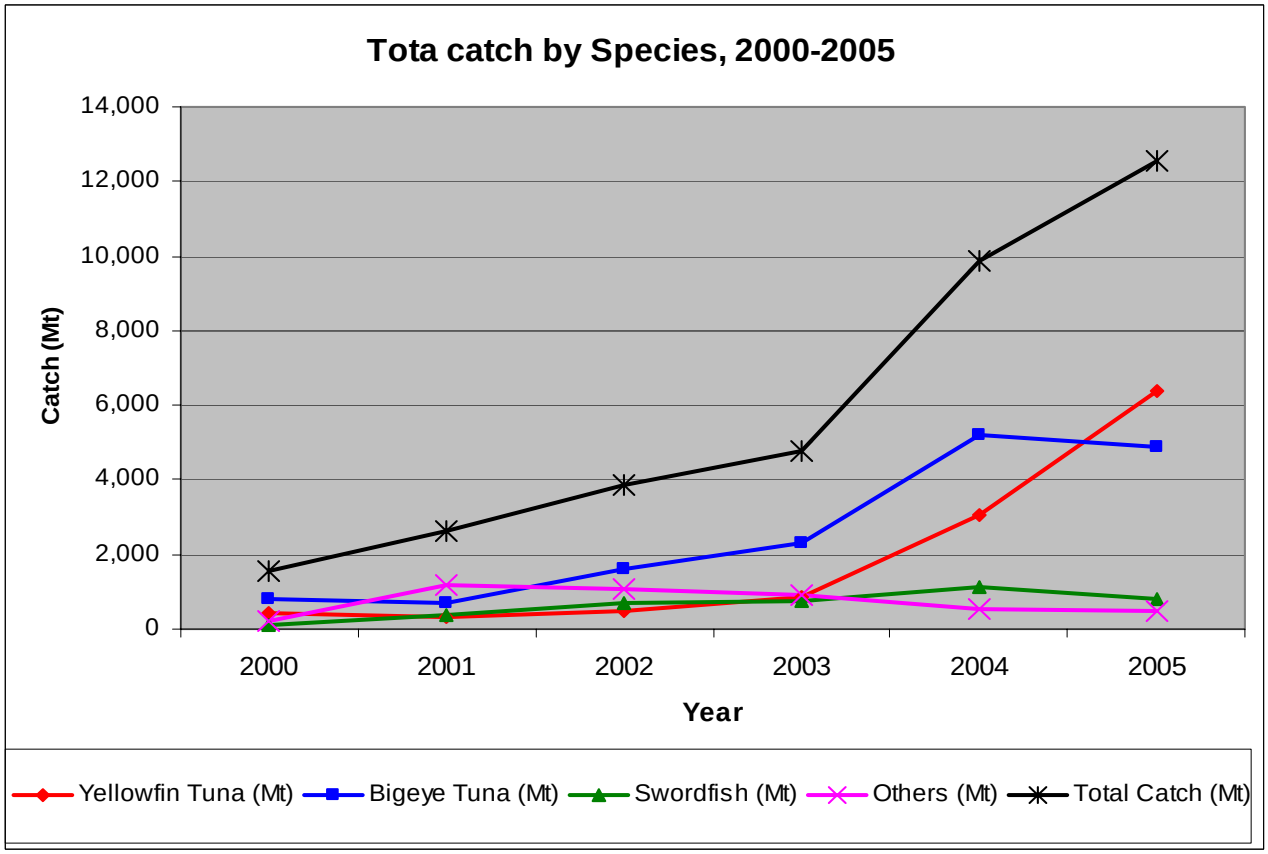

Figure 5. Species composition of revised catch data.

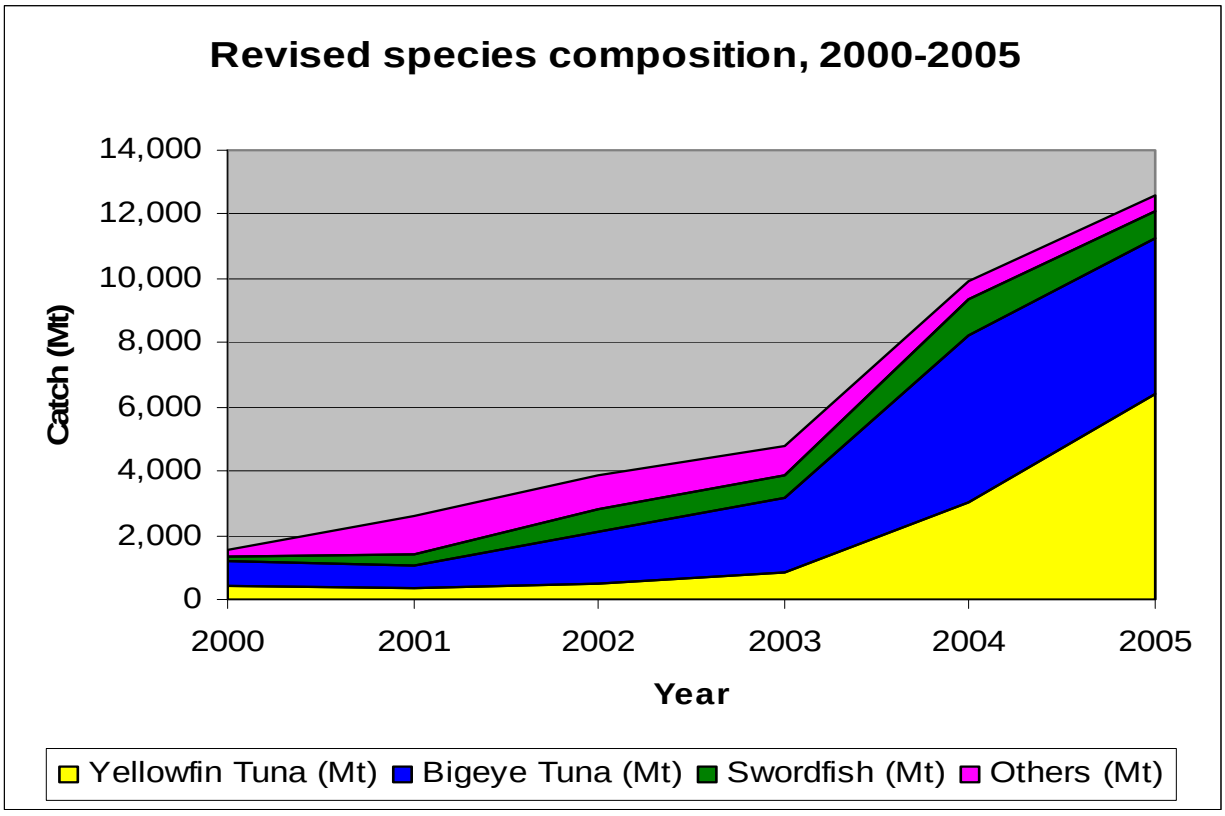

Figure 6a. Species composition of revised data.

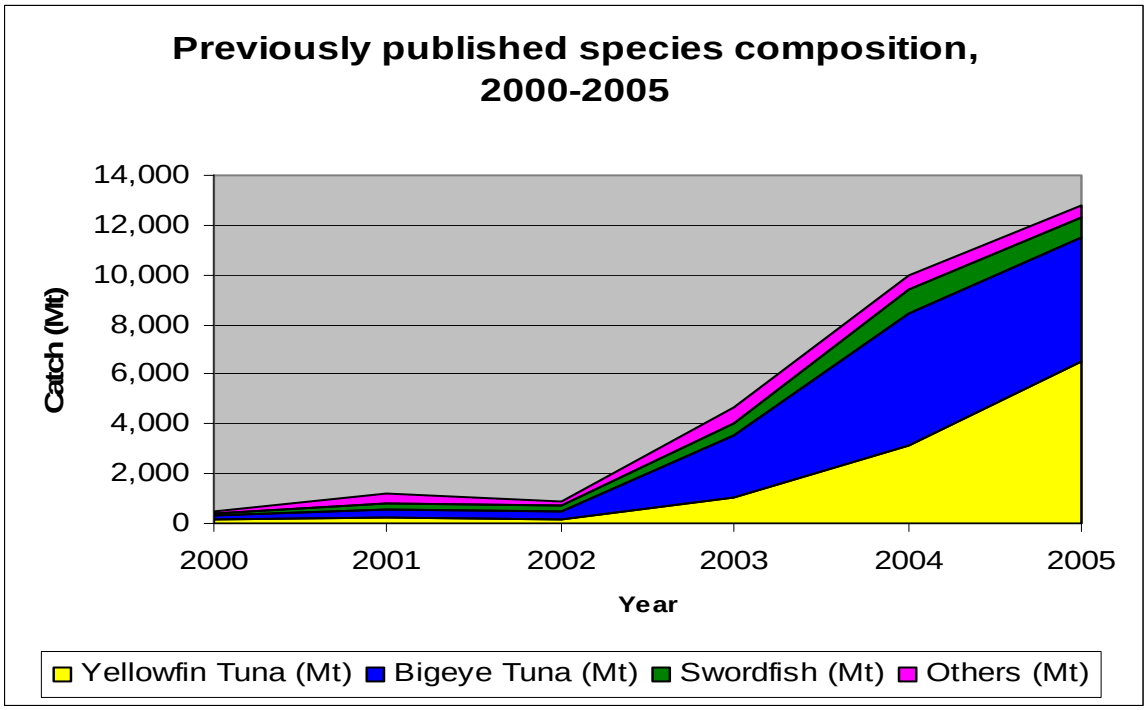

Figure 6b. Species composition of previously published data.

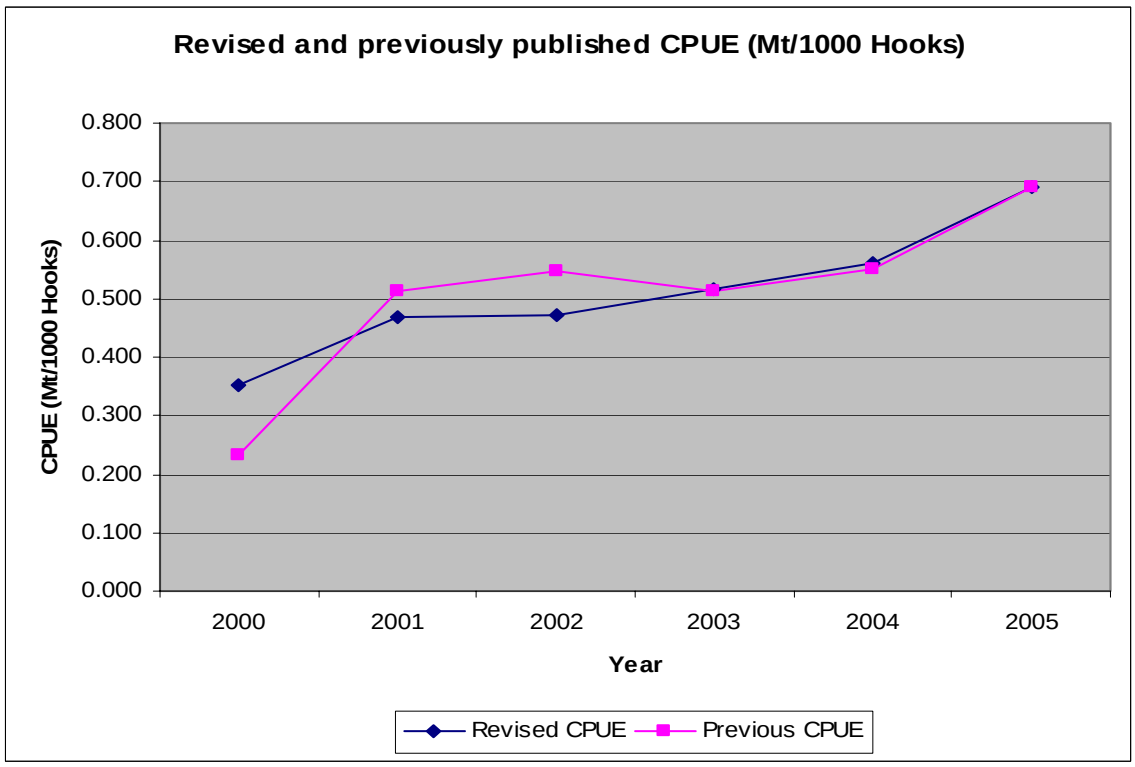

Figure 7. CPUE (Mt/1000 hooks) for the years 2000 to 2005

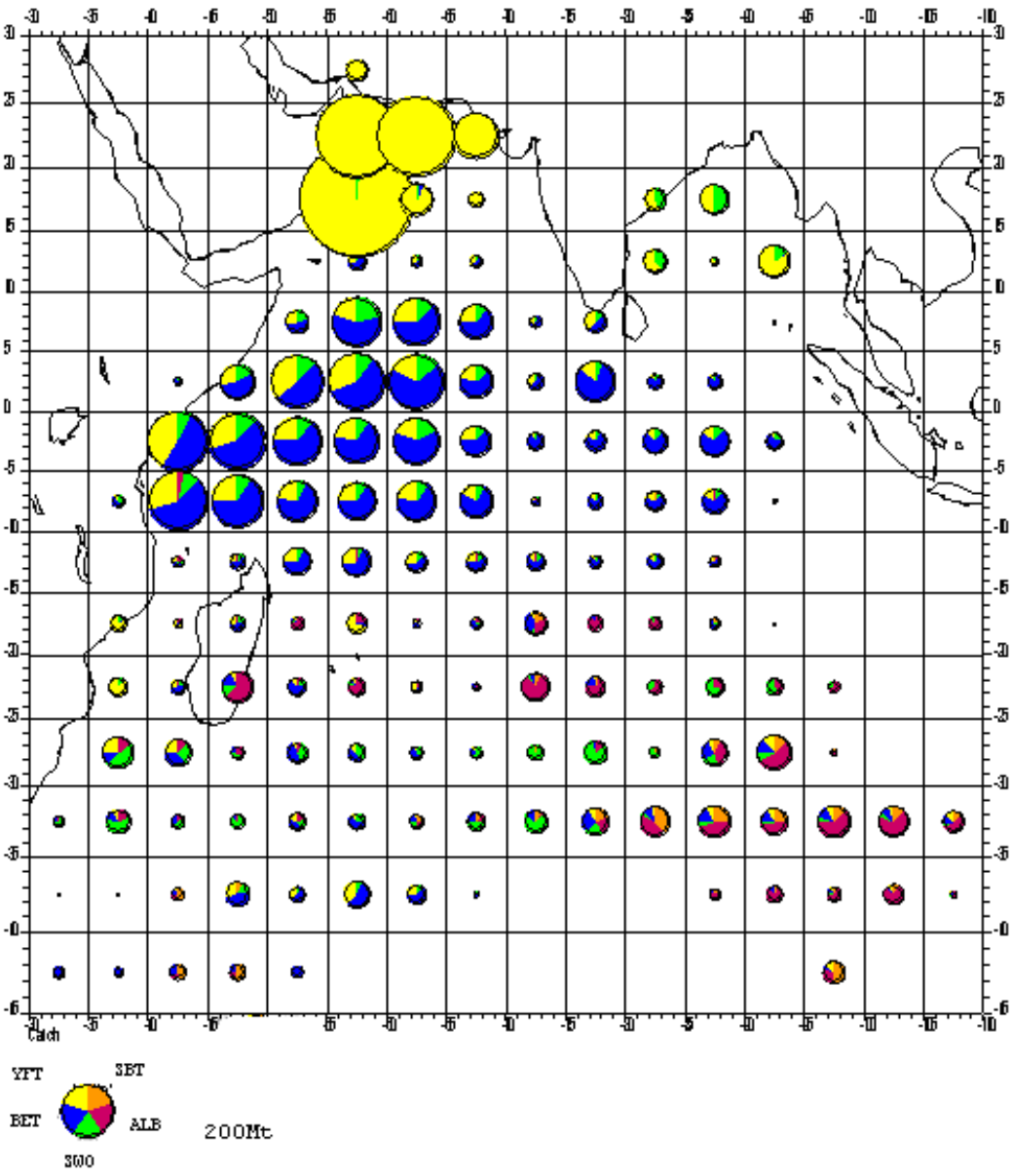

Figure 8. Average catches (Mt) per year for the years 2000 to 2005.

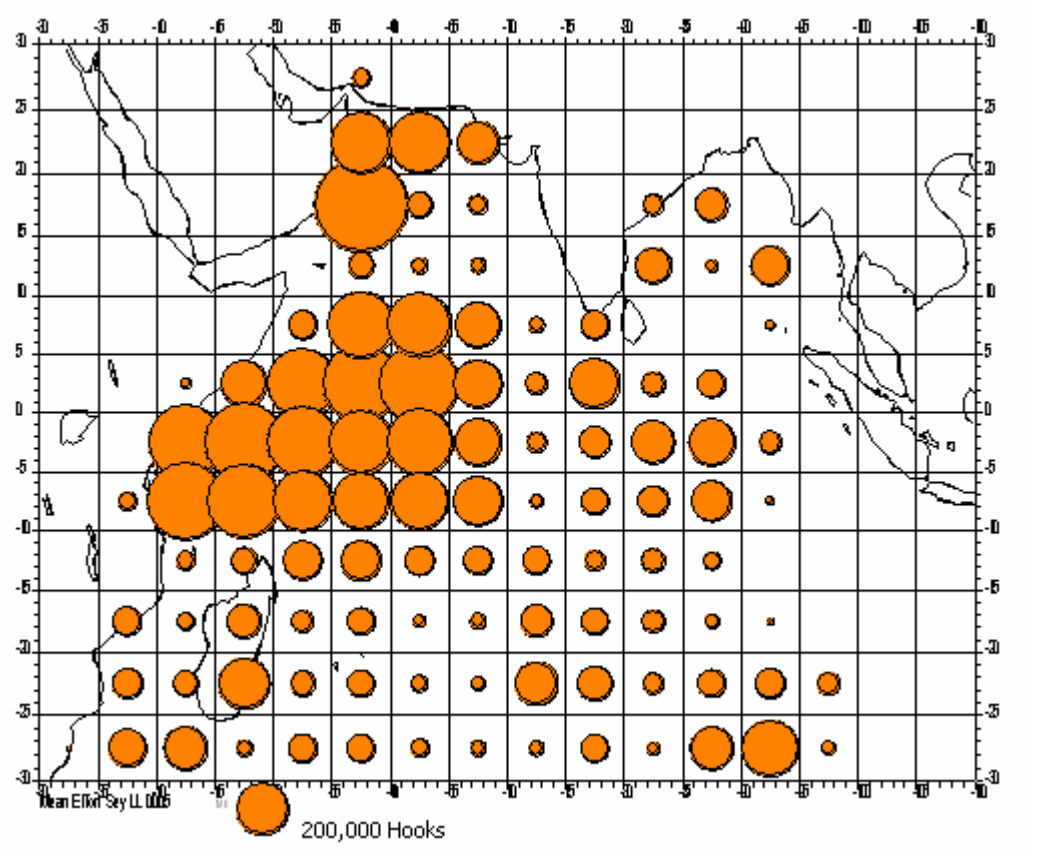

Figure 9. Average effort (1000 hooks) per year for the years 2000 to 2005.